Computers in Biology and Medicine 75 (2016) 74–79

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

A computational approach to mortality prediction of alcohol use disorder inpatients

Jacob Calvert [a], Qingqing Mao [a], Angela J. Rogers [b], Christopher Barton [c], Melissa Jay [a], Thomas Desautels [a], Hamid Mohamadlou [a], Jasmine Jan [d], Ritankar Das [a,*]

[a] Dascena, Inc, United States
[b] Department of Medicine, Stanford University, United States
[c] Department of Emergency Medicine, University of California, San Francisco, United States
[d] Department of Bioengineering, University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 14 March 2016
Received in revised form
17 May 2016
Accepted 22 May 2016

Keywords:
Alcohol use disorder
Clinical decision support systems
Mortality prediction
Electronic health records
Medical informatics

ABSTRACT

Background: Health information technologies can assist clinicians in the Intensive Care Unit (ICU) by providing additional analysis of patient stability. However, because patient diagnoses can be confounded by chronic alcohol use, the predictive value of existing systems is suboptimal. Through the use of Electronic Health Records (EHR), we have developed computer software called *AutoTriage* to generate accurate predictions through multi-dimensional analysis of clinical variables. We analyze the performance of *AutoTriage* on the Alcohol Use Disorder (AUD) subpopulation in this study, and build on results we reported for *AutoTriage* performance on the general population in previous work.
Methods: AUD-related ICD-9 codes were used to obtain a patient population from MIMIC III ICU dataset for a retrospective study. Patient mortality risk score is generated through analysis of eight EHR-based clinical variables. The score is determined by combining weighted subscores, each of which are obtained from singlets, doublets or triplets of one or more of the eight continuous-valued clinical variable inputs. A temporally updating risk score is computed with a continuously revised 12-hour mortality prediction.
Results: Among AUD patients, in a non-overlapping test set, *AutoTriage* outperforms existing systems with an Area Under Receiver Operating Characteristic (AUROC) value of 0.934 for 12-h mortality prediction. At a sensitivity of 90%, *AutoTriage* achieves a specificity of 80%, positive predictive value of 40%, negative predictive value of 89%, and an Odds Ratio of 36.
Conclusions: For mortality prediction, *AutoTriage* demonstrates improvements in both the accuracy and the Odds Ratio over current systems among the AUD patient population.

1. Introduction

Clinical Decision Support Systems (CDSS) can be used to help assess patient conditions and predict patient mortality risk. Accurate predictions in the ICU are needed for timely medical attention and for the allocation of limited ICU resources [1,2]. Several existing prediction models such as Modified Early Warning Score (MEWS) [3], Sepsis-Related Organ Failure Assessment (SOFA) [4], and Simplified Acute Physiology Score (SAPS II) [5], rely on weighted linear combinations of basic patient characteristics such as age, type of admission and vital sign measurements. The high false positive rate of alerts resulting from such classifiers often leads to alarm fatigue in the clinical setting [6]. Desensitization from an overexposure to false alarms has widely been documented to increase response times and decrease receptivity to correct alerts [7]. For the reasons described below, the difficulty of assessing stability is exacerbated in patients suffering from alcohol-use dependence (AUD) and calls for improvement.

AUD patients encompass 1 out of 10 critical care admissions and are up to 8% more likely to experience unplanned re-hospitalization within 30-days of discharge [8,9]. This is because the standard techniques for screening patient stability can be confounded by chronic alcohol use. In particular, signs of acute hypotension, which can be indicative of life-threatening

Abbreviations: ICU, Intensive Care Unit; MICU, Medical Intensive Care Unit; EHR, Electronic Health Records; AUD, Alcohol Use Disorder; ICD-9, International Statistical Classification of Diseases version 9; MIMIC III, Multiparameter Intelligent Monitoring in Intensive Care version III; ROC, Receiver Operating Characteristic; AUROC, Area Under Receiver Operating Characteristic; CDSS, Clinical Decision Support Systems; MEWS, Modified Early Warning Score; SOFA, Sepsis-Related Organ Failure Assessment; SAPS II, Simplified Acute Physiology Score; WBC, White Blood Cell count; HIPAA, Health Insurance Portability and Accountability Act; BIDMC, Beth Israel Deaconess Medical Center; PPV, Positive Predictive Value; NPV, Negative Predictive Value; LR+, Positive Likelihood Ratio; LR-, Negative Likelihood Ratio; OR, Odds Ratio

* Corresponding author.

http://dx.doi.org/10.1016/j.compbiomed.2016.05.015
0010-4825/

homeostatic failures like sepsis, anaphylaxis and renal failure [4,10,11] may be masked among patients in the AUD subpopulation, because AUD patients often suffer from chronic hypertension [12]. Chronic alcohol use can also confound the value of a Leukocyte Differential (WBC) Count, a common lab test used in the diagnosis of a variety of medical conditions [13]. The higher risk of unplanned rehospitalization, in conjunction with poorer diagnostic screening performance, underscore the need for improved risk scoring systems for AUD patients.

The use of Electronic Health Records (EHR) in hospitals provides an opportunity to improve predictive value from clinical data and provide clinical decision support. Recent studies have used various patient measurements and patient trends to improve mortality predictions, leading to incremental progress [14–16]. To further improve the quality of mortality prediction, we have developed *AutoTriage*, an algorithm which interrogates trends among clinical variables and also analyzes inter-correlations [17]. Using only eight common clinical measurements and analyzing inter-correlations reduces the chances that real-time data unavailability challenges affect algorithm performance. As homeostasis is governed by multi-organ feedback regulation, these variable correlations uncover useful patterns across organ systems. Accurate and early identification of deteriorating patients with assistance from a CDSS tool like *AutoTriage* has the potential to significantly decrease the number of life-threatening situations arising in the critical care wards of the hospital, and in turn lead to reductions in patient mortality rates. In this study, we demonstrate the application of *AutoTriage* on the difficult-to-predict AUD subpopulation to demonstrate the ability of the algorithm to overcome factors that are confounded by traditional diagnostic analysis.

2. Methods

2.1. Data set

Fig. 1 depicts the patient exclusion process used to select 3054 patient records from the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) III database [18]. The records consist of de-identified clinical information of adults admitted to the Beth Israel Deaconess Medical Center (BIDMC) medical intensive care unit (MICU). Since the study did not impact patient safety and all data were in accordance with the HIPAA Privacy Rule, the requirement for patient consent was waived by the Institutional Review Boards of BIDMC and the Massachusetts Institute of Technology.

Inclusion criteria for this study were:

I. Age of at least 18 years and admission to the MICU.
II. Documented length-of-stay and survival for at least 17 h following admission. A 17-h minimum accounts for a 12-hour advance warning after 5 h of patient monitoring using *AutoTriage*. Documented AUD-related ICD-9 code (291.X, 291.XX, 303.XX, 305.XX, 357.5, 425.5, 535.3X, 571.2, and 571.3, where X denotes a wildcard).

The eight physiological measurements utilized with 1-h resolution were heart rate, pH, pulse pressure, respiration rate, blood oxygen saturation, systolic blood pressure, temperature, and white blood cell count (WBC). With the exception of pH and WBC, these measurements are frequently sampled in the ICU, thus ensuring our predictions are broadly applicable. In addition, MEWS, SAPS II, and SOFA utilize similar measurements for prediction, allowing us to compare performance with these systems.

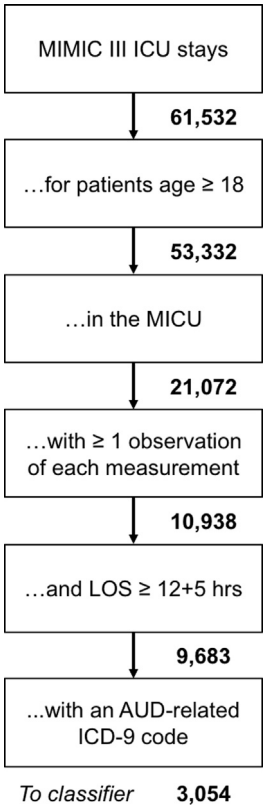

Fig. 1. Patient inclusion flowchart.

2.2. Gold standard

In-hospital mortality was our gold standard. The time of last measurement available was chosen as the time of death of the patient. By this definition, 392 patients were flagged as having in-hospital deaths and 2662 as survivors, resulting in a 12.84% mortality prevalence.

2.3. Binning, feature construction, and score assignment

Our classifiers use a set of four non-linear input features designed to capture empirical risk as a function of the patient measurements. Each individual type of measurement was associated with a non-linear function mapping from value of the measurement to approximate empirical risk of gold standard-defined mortality. These non-linear functions were typically polynomials of degree four or five. The sum of all such functions was taken as the first input to our classifier. Trends for observed patient measurements were also calculated using time-parameterized sequences of measurements. For example, a change in respiratory rate was calculated for each set of two consecutive timestamps with respiratory rate measurements (imputed to the nearest hour) and used as a stand-alone input feature [17]. During hours with no updated entries, the measurement value was approximated as the most recent value available and trends were calculated from this value. In the same fashion as with the measurement values, the empirical risk of mortality as a function of each individual trend was approximated with a polynomial and used as an input feature. The sum of all of these risk approximating functions of the trend values was taken as the second feature input of our classifier. Finally, combinations of two or three trends were related to empirical risk of mortality. In this procedure, the trends were binned by value, where bin edges were assigned heuristically. The empirical risk of mortality for each combination of two or three

binned trends was calculated and stored. The resulting non-linear functions from trend values to approximate empirical mortality risk were summed, with a sum of pair-derived features, and a separate sum of triplet-derived features. These two sums comprised the third and fourth inputs to our classifiers.

We trained our logistic regression classifiers using this set of features. These classifiers were trained to assess patient stability 12 h prior to patient discharge or death. A window of five hours prior to this time was used to gather data, calculate the measurement and trend values, and generate the patient mortality score. All of our performance measures are the results of a four-fold cross validation procedure, where assignment to each fold was randomized. We utilized custom MATLAB (MathWorks, Natick, MA, R2015a) scripts to implement all of the operations described in this section (Fig. 2).

2.4. Generating MEWS, SAPS II, and SOFA scores for comparison

Software provided alongside the MIMIC III database [19] was used to produce *SAPS II* and *SOFA* scores. Similarly, the *MEWS* heuristic table was used to generate *MEWS* scores [3].

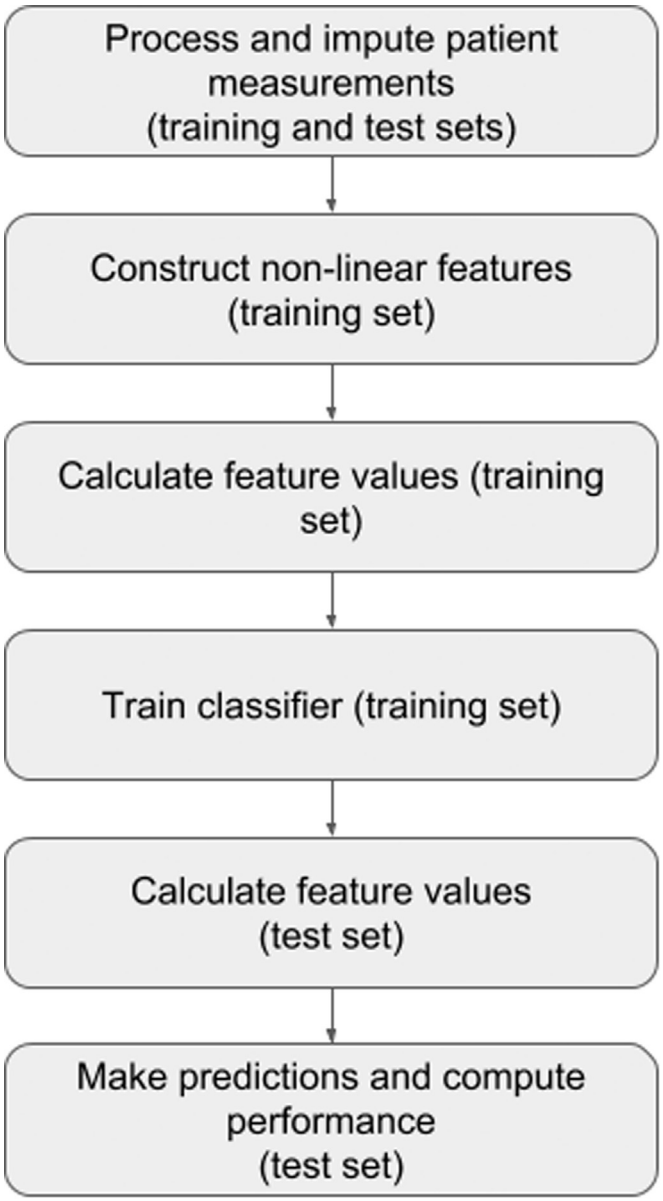

Fig. 2. *AutoTriage* algorithm flowchart.

3. Results

Receiver Operating Characteristic curves (ROC) for 12-h mortality prediction on the non-overlapping test set is shown in Fig. 3. *AutoTriage* yields a higher average Area Under ROC (AUROC) value of 0.934 compared to AUROC of 0.7597, 0.7657, and 0.7913, by *MEWS, SAPS II* and *SOFA,* respectively.

AutoTriage score ≥ -2.64 achieves an improved sensitivity of 90% while maintaining a higher specificity of 80% when compared to the 66% sensitivity and 74% specificity for *MEWS* ≥ 3 [20]. At this 90% sensitivity, *AutoTriage* demonstrates significant improvements in accuracy, specificity and predictive value over existing methods. The generalizability of these results was confirmed by 4-fold cross validation.

AutoTriage also achieves a higher accuracy of 81% than existing methods. The increased positive likelihood ratio (LR+) and decreased negative likelihood ratio (LR-) both demonstrate increased accuracy and confidence in mortality and stability assessment. Furthermore, *AutoTriage* achieves an Odds Ratio of 36.46, providing more than a six-fold increase in confidence of correct mortality prediction compared to MEWS ≥ 3. This relatively high

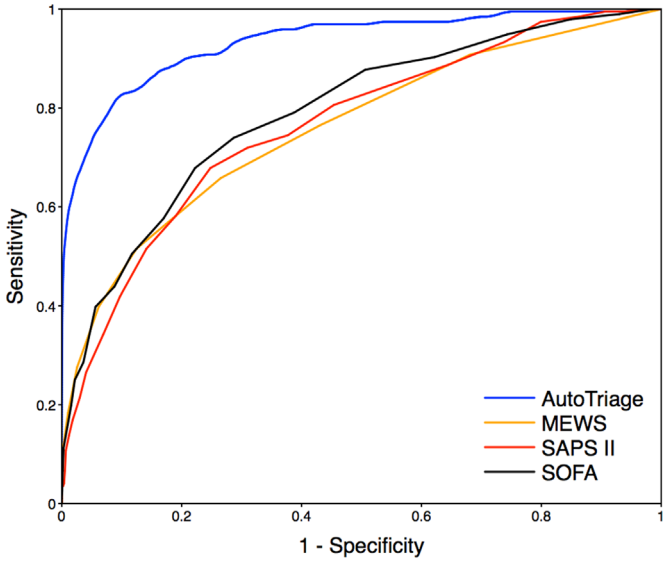

Fig. 3. *AutoTriage*, Modified Early Warning Score (*MEWS*), Sequential Organ Failure Assessment (*SOFA*), and Simplified Acute Physiology Score (*SAPS II*) Receiver Operating Characteristic (ROC) curves for 12-hour mortality prediction in the medical intensive care unit.

Table 1
Statistical outcomes of 12-h mortality prediction using *AutoTriage* as compared to other commonly used severity of illness scoring systems. The F1 score is the harmonic mean of precision and recall.

	AutoTriage (≥ -2.64)	MEWS (≥ 3)	MEWS (≥ 1)	SAPS II (≥ 17)	SOFA (≥ 5)
Sensitivity	0.9031	0.6582	0.9082	0.8878	0.9031
Specificity	0.7964	0.7348	0.3186	0.3546	0.3779
PPV	0.3951	0.2677	0.1641	0.1684	0.1761
NPV	0.8915	0.6825	0.7763	0.7596	0.7959
F1	0.5497	0.3806	0.2779	0.2832	0.2948
OR	36.4557	5.3356	4.6258	4.3474	5.6615
LR+	4.4357	2.4819	1.3328	1.3756	1.4517
LR−	0.1217	0.4652	0.2881	0.3164	0.2564
Accuracy	0.8101	0.7250	0.3943	0.4230	0.4453

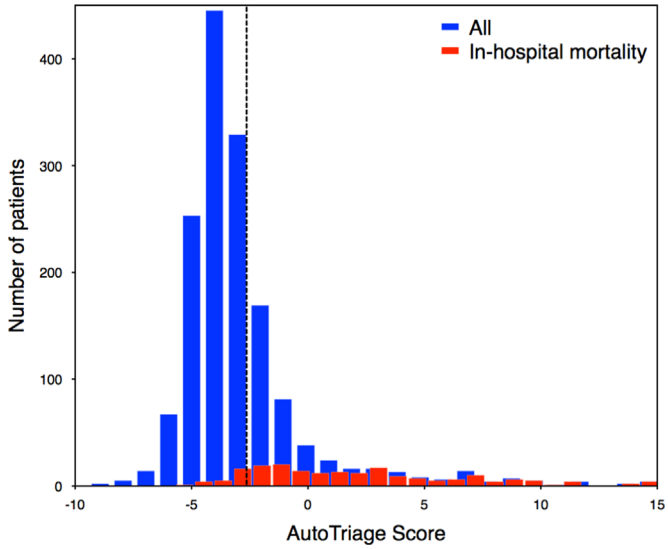

Fig. 4. AUD patient distribution across *AutoTriage* score values among survivors and non-survivors. The dotted vertical line represents an *AutoTriage* threshold of −2.64, corresponding to a sensitivity of 90% and specificity of 81%.

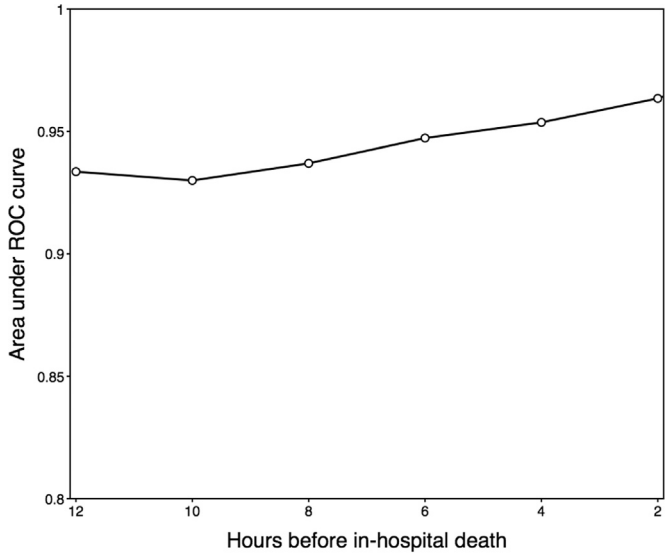

Fig. 5. *AutoTriage* AUROC over time preceding in-hospital death in the MICU.

performance of *AutoTriage* demonstrates the ability to reduce false alarms, increase confidence in the need for medical intervention and increase confidence in assessing patient stability. Performance of *AutoTriage* and of existing systems in the non-overlapping test set is summarized in Table 1.

Fig. 4 shows the distribution of survivors and non-survivors across the *AutoTriage* score range in the non-overlapping test set. The vertical line of *AutoTriage* = −2.64 indicates a sensitivity of 90% and specificity of 81%. Lowering this threshold would indicate more patients at risk but increase the number of false positive cases.

AUROC is visualized in Fig. 5 in the non-overlapping test set, as it increases over time (prior to in-hospital death). As time-of-death approaches the method improves in predictive power as the patient condition becomes more obviously unstable.

4. Discussion

Risk scoring systems like *MEWS*, *SAPS II*, and *SOFA* evaluate patients by combining weights associated with value-ranges of vital signs and lab tests. This tabulation approach benefits from its simple implementation and reliance on only the most common patient measurements, which makes it more likely that the necessary score components are available. However, these methods fail to take full advantage of the information made available in EHR, which can contribute valuable time series data to the prediction-making process. *AutoTriage* pairs the strengths of existing scoring systems with multidimensional analysis, including trend assessment [17].

Fig. 3 demonstrates the substantial benefits of employing multidimensional analysis in prediction-making. Note that, while *AutoTriage* ultimately incorporates many more features than the comparison methods, it too uses only a handful of measurements drawn from the same pool of widely available vital signs and lab tests as data inputs. *AutoTriage* does not make any assumptions about the input data or place any restrictions, except that included patients have at least one observation of each of the relevant measurements. Specifically, *AutoTriage* uses 8 input measures consisting mostly of vital signs, while *MEWS, SAPS II, and SOFA* use 5, 17 and 10 input measures respectively, consisting in large part of laboratory tests for the latter two. The additional features used in *AutoTriage* arise from correlating combinations and trends of multiple input measurements with an outcome of interest, which are pieces of information not considered by the other methods in Fig. 3. For an individual prediction, *AutoTriage* is able to deliver this performance with sub-second computational cost on a desktop computer.

Like *MEWS*, *SAPS II*, and *SOFA, AutoTriage* generates an all-cause mortality score and does not discriminate between mortality causes. The higher-order correlations and trends between clinical measurements are designed to detect the loss of homeostasis that occurs when biological feedback loops break down, prior to the signs of patient instability that can be detected from single vital sign analysis. Therefore the *AutoTriage* score is generated independently of underlying disease state, instead reflecting the overall loss of homeostasis of the individual.

Since the prevalence of in-hospital mortality (12.84%) for AUD patients is low relative to the number of survivors, many of the common metrics for assessing the quality of a predictor's performance are also affected (tending to be lower). Despite this, the *AutoTriage* metrics shown in Table 1 are encouraging, and more than double that of several of the other methods' metrics. Due to the low prevalence positive class, the F1 score is particularly relevant, as it does not reward the number of survivors that were predicted to be survivors (of which there are many). The F1 score thus emphasizes the predictions made for the in-hospital death cases, which should be prioritized in order to align with clinical priorities. As is the case with most of the other metrics, *AutoTriage* roughly doubles the F1 score over existing methods.

Fig. 4 illustrates how the choice of *AutoTriage* score threshold made in Table 1 classifies patients with respect to their distribution over patient scores. Unlike the sharp peak in the distribution over all AUD patient scores, the in-hospital death score distribution is broad, flat, and situated mostly across positive *AutoTriage* scores. This enables the score threshold of −2.64 to capture most of the in-hospital death patients, while misclassifying relatively few survivors. However, the score threshold can be chosen to emphasize different prediction metrics, based on a health system's or even an individual clinician's priorities. For example, the score could be lowered further to capture a greater fraction of in-hospital deaths or raised to improve specificity. This threshold tuning could be motivated by a desire to reduce dangerous alarm fatigue associated with CDSS systems [21]. In the case that a higher specificity is chosen, it is possible that the concomitant decrease in sensitivity would be balanced out over time, because a prospective application of *AutoTriage* would be functional immediately at

admission and would predict mortality longitudinally with increasing quality over time (Fig. 5), without specifying a look-ahead period. Contrastingly, we fix the look-ahead period to 12 h in this retrospective study via inclusion criterion (ii), in order to concisely present the performance of the *AutoTriage* algorithm.

Limitations of our study include a retrospective design and the use of a single center cohort. Our performance may be overestimated because the training and testing were executed on partitions of the same data set. In the future we will test the performance of this algorithm on data from different medical centers. Also in future studies we plan to validate the *AutoTriage* algorithm prospectively in multiple hospitals. Demographic and institutional differences could result in *AutoTriage* performance variability. However, it is more important that *AutoTriage* performs well after retraining on retrospective data at each site, rather than that a single training of *AutoTriage* perform well across all institutions. This is because the clinical implementation process may include a preliminary stage of model customization to a particular site, if necessary.

It is important to note that a prediction of mortality for a given patient by *AutoTriage* or other methods like *MEWS*, *SAPS II*, or *SOFA*, does not imply a medical intervention can be made that will prevent or even delay mortality, or if such an intervention is possible, that it would be desirable to the clinical team, the patient, or the patient's family members. However, several studies have shown that the use of screening systems like *MEWS* leads to improved patient outcomes, reduced in-hospital mortality, and reduced code blue events [22–25]. These results suggest that, overall, patients would benefit from *AutoTriage* predictions, which would identify more at-risk patients than *MEWS* with fewer false alarms.

5. Conclusion

Inspired by existing mortality prediction systems, *AutoTriage* manipulates widely available data to generate risk scores. *AutoTriage* is distinct from other methods, however, through its use of multi-dimensional combinations of these measurements and their trends. These differences allow *AutoTriage* to outperform existing methods for the scoring of in-hospital mortality of AUD patients, achieving a sensitivity of 90% and a specificity of 80%, 12 h before in-hospital mortality. These results suggest that *AutoTriage* has the potential to add value in a clinical setting, even for difficult-to-predict subpopulations, and indicate that variations of this method could address other medical problems of interest through inpatient clinical decision support.

Funding source

This material is based upon work supported by the National Science Foundation under Grant no. 1549867. The funder had no role in the conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, and approval of the manuscript; and decision to submit the manuscript for publication.

Conflict of interest

None declared

Acknowledgments

We acknowledge Dr. Jana Hoffman for valuable contributions to the writing of this manuscript.

References

[1] G.C. Siontis, I. Tzoulaki, J.P. Ioannidis, Predicting death: an empirical evaluation of predictive tools for mortality, Arch. Intern. Med. 171 (19) (2011) 1721–1726, http://dx.doi.org/10.1001/archinternmed.2011.334.
[2] K. Reini, M. Fredrikson, A. Oscarsson, The prognostic value of the modified early warning score in critically ill patients: a prospective, observational study, Eur. J. Anaesthesiol. 29 (3) (2012) 152–157, http://dx.doi.org/10.1097/EJA.0b013e32835032d8.
[3] C.P. Subbe, A. Slater, D. Menon, L. Gemmell, Validation of physiological scoring systems in the accident and emergency department, Emerg. Med. J. 23 (11) (2006) 841–845, http://dx.doi.org/10.1136/emj.2006.035816.
[4] J.L. Vincent, R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, C. K. Reinhart, P.M. Suter, L.G. Thijs, The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure, Intensive Care Med. 22 (7) (1996) 707–710, http://dx.doi.org/10.1007/BF01709751.
[5] J.R. Le Gall, S. Lemeshow, F. Saulnier, A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study, JAMA 270 (24) (1993) 2957–2963, http://dx.doi.org/10.1001/jama.270.24.2957.
[6] J.M. Blankush, R. Freeman, J. McIlvaine, T.L. Tran, S.J. Nassani, C. Sullivan, I. M. Leitman, Watchful eye or ineffective nuisance: do automated mews systems incorporating end-tidal capnography provide effective extended coverage or more false alarms? J. Am. Coll. Surg. 22 (4) (2015) e70–e71, http://dx.doi.org/10.1016/j.jamcollsurg.2015.08.084.
[7] S. Sendelbach, M. Funk, Alarm fatigue: a patient safety concern, AACN Adv. Crit. Care 24 (4) (2013) 378–386, http://dx.doi.org/10.1097/NCI.0b013e3182a903f9.
[8] W.A. Baldwin, B.A. Rosenfeld, M.J. Breslow, T.G. Buchman, C.S. Deutschman, R. D. Moore, Substance abuse-related admissions to adult intensive care., CHEST J. 103 (1) (1993) 21–25, http://dx.doi.org/10.1378/chest.103.1.21.
[9] P.K. Lindenauer, S.L. Normand, E.E. Drye, Z. Lin, K. Goodrich, M.M. Desai, D. W. Bratzler, W.J. O'Donnell, M.L. Metersky, H.M. Krumholz, Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia, J. Hosp. Med. 6 (3) (2011) 142–150, http://dx.doi.org/10.1002/jhm.890.
[10] J.L. Vincent, D.C. Zapatero, The role of hypotension in the development of acute renal failure, Nephrol. Dial. Transplant. 24 (2) (2009) 337–338, http://dx.doi.org/10.1093/ndt/gfn679.
[11] Australia Wide First Aid, Anaphylaxis and Hypotension. ⟨https://www.australiawidefirstaid.com.au/hypotension/⟩, 2003 (accessed 15.08.03).
[12] L.J. Beilin, Alcohol and hypertension, Clin. Exp. Pharmacol. Physiol. 22 (3) (1995) 185–188, http://dx.doi.org/10.1111/j.1440-1681.1995.tb01977.x.
[13] H. Parry, S. Cohen, J.E. Schlarb, D.A. Tyrrell, A. Fisher, M.A. Russell, M.J. Jarvis, Smoking, alcohol consumption, and leukocyte counts, Am. J. Clin. Pathol. 107 (1) (1997) 64–67, http://dx.doi.org/10.1093/ajcp/107.1.64.
[14] L.W. Lehman, R.P. Adams, L. Mayaud, G.B. Moody, A. Malhotra, R.G. Mark, S. Nemati, A physiological time series dynamics-based approach to patient monitoring and outcome prediction, IEEE J. Biomed. Health Inf. 19 (3) (2015) 1068–1076, http://dx.doi.org/10.1109/JBHI.2014.2330827.
[15] S. Ghose, J. Mitra, S. Khanna, J. Dowling, An improved patient-specific mortality risk prediction in ICU in a random forest classification framework, in: A. Georgiou, H. Grain, L.K. Schaper (Eds.), Driving Reform: Digital Health is Everyone's Business: Selected Papers from the 23rd Australian National Health Informatics Conference (HIC 2015), IOS Press, 2015, pp. 56–61, http://dx.doi.org/10.3233/978-1-61499-558-6-56.
[16] R. Pirracchio, M.L. Petersen, M. Carone, M.R. Rigon, S. Chevret, M.R. van der Laan, Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study, Lancet Respir. Med. 3 (1) (2015) 42–52, http://dx.doi.org/10.1016/S2213-2600(14)70239-5.
[17] J.S. Calvert, D.A. Price, C.W. Barton, U.K. Chettipally, R. Das, Discharge recommendation based on a novel technique of homeostatic analysis, J. Am. Med. Inf. Assoc. (2016), http://dx.doi.org/10.1093/jamia/ocw014.
[18] M. Saeed, M. Villarroel, A.T. Reisner, G. Clifford, L.W. Lehman, G. Moody, T. Heldt, T.H. Kyaw, B. Moody, R.G. Mark, Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): a public-access intensive care unit database, Crit. Care Med. 39 (5) (2011), http://dx.doi.org/10.1097/CCM.0b013e31820a92c6.
[19] MIMIC Code Repository: Code shared by the research community for the MIMIC-III database. ⟨https://github.com/MIT-LCP/mimic-code⟩. Archived at: ⟨http://www.webcitation.org/6ffJC48Mj⟩.
[20] D.G. Altman, J.M. Bland, Diagnostic tests. 1: Sensitivity and specificity, Br. Med. J. 308 (1994) 1552, PMID:8019315.
[21] P.S. Roshanov, N. Fernandes, J.M. Wilczynski, B.J. Hemens, J.J. You, S. M. Handler, R. Nieuwlaat, N.M. Souza, J. Beyene, H.G. Van Spall, A.X. Garg, R. B. Haynes, Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials, BMJ 346 (2013) f657, http://dx.doi.org/10.1136/bmj.f657.

[22] C. Mathukia, W. Fan, K. Vadyak, C. Biege, M. Krishnamurthy, Modified Early Warning System improves patient safety and clinical outcomes in an academic community hospital, Community Hosp. Intern. Med. Perspect. 5 (2) (2015) 26716, http://dx.doi.org/10.3402/jchimp.v5.26716.
[23] B.L. Albert, L. Huesman, Development of a modified early warning score using the electronic medical record, Dimens. Crit. Care Nurs. 30 (5) (2011) 283–292, http://dx.doi.org/10.1097/DCC.0b013e318227761d.
[24] A.L. Green, A. Williams, An evaluation of an early warning clinical marker referral tool, Intensive Crit. Care Nurs. 22 (5) (2006) 274–282, http://dx.doi.org/10.1016/j.iccn.2006.04.004.
[25] J.M. Maupin, D.J. Roth, J.M. Krapes, Use of the modified early warning score decreases code blue events, Jt. Comm. J. Qual. Patient Saf. 35 (12) (2009) 598–603, PMID:20043499.