Annals of Medicine and Surgery 11 (2016) 52–57

Contents lists available at ScienceDirect

Annals of Medicine and Surgery

journal homepage: www.annalsjournal.com

Using electronic health record collected clinical variables to predict medical intensive care unit mortality

Jacob Calvert [a], Qingqing Mao [a], Jana L. Hoffman [a, *], Melissa Jay [a], Thomas Desautels [a], Hamid Mohamadlou [a], Uli Chettipally [b, c], Ritankar Das [a]

[a] Dascena Inc., Hayward, CA, USA
[b] Kaiser Permanente South San Francisco Medical Center, South San Francisco, CA, USA
[c] Department of Emergency Medicine, University of California San Francisco, San Francisco, CA, USA

HIGHLIGHTS

- Multi-dimensional analysis of clinical inputs used to generate mortality risk scores.
- AutoTriage 12 h mortality prediction achieves an AUROC of 0.88.
- Sensitivity of 80% at a specificity of 81% with diagnostic odds ratio of 16.
- Outperforms MEWS, SOFA and SAPS II for mortality prediction, with an accuracy of 80%.

ARTICLE INFO

Article history:
Received 19 May 2016
Received in revised form
2 September 2016
Accepted 4 September 2016

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

ABSTRACT

Background: Clinical decision support systems are used to help predict patient stability and mortality in the Intensive Care Unit (ICU). Accurate patient information can assist clinicians with patient management and in allocating finite resources. However, systems currently in common use have limited predictive value in the clinical setting. The increasing availability of Electronic Health Records (EHR) provides an opportunity to use medical information for more accurate patient stability and mortality prediction in the ICU.

Objective: Develop and evaluate an algorithm which more accurately predicts patient mortality in the ICU, using the correlations between widely available clinical variables from the EHR.

Methods: We have developed an algorithm, *AutoTriage*, which uses eight common clinical variables from the EHR to assign patient mortality risk scores. Each clinical variable produces a subscore, and combinations of two or three discretized clinical variables also produce subscores. A combination of weighted subscores produces the overall score. We validated the performance of this algorithm in a retrospective study on the MIMIC III medical ICU dataset.

Results: *AutoTriage* 12 h mortality prediction yields an Area Under Receiver Operating Characteristic value of 0.88 (95% confidence interval 0.86 to 0.88). At a sensitivity of 80%, *AutoTriage* maintains a specificity of 81% with a diagnostic odds ratio of 16.26.

Conclusions: Through the multidimensional analysis of the correlations between eight common clinical variables, *AutoTriage* provides an improvement in the specificity and sensitivity of patient mortality prediction over existing prediction methods.

1. Introduction

There is a need for accurate prediction of mortality risk and patient deterioration in the Intensive Care Unit (ICU) [1]. Advanced warning of patient deterioration is crucial for timely medical intervention and patient management, and accurate risk assessment aids in the allocation of limited ICU resources. Clinical Decision Support Systems (CDSS) have been used in the ICU for predicting patient outcome and to score the severity of patient condition [2–4]. The vast majority of prediction models currently in

* Corresponding author. Dascena Inc., 1135 Martin Luther King Drive, Hayward, CA 94541, USA.
E-mail address: jana@dascena.com (J.L. Hoffman).

http://dx.doi.org/10.1016/j.amsu.2016.09.002

use are based on aggregate baseline patient characteristics. These systems usually rely on a weighted linear combination of features, such as age, type of admission, and vital sign measurements. However, the most commonly used CDSS such as the Modified Early Warning Score (*MEWS*) [5], the Sequential Organ Failure Assessment (*SOFA*) [6], and the Simplified Acute Physiology Score (*SAPS II*) [7], have suboptimal specificity and sensitivity when applied to patient mortality prediction [2]. These CDSS assessments assume that risk factors are independent from one another, and, therefore, they are not sensitive to the underlying complex homeostatic physiologies of patients. Additionally, they do not account for variations in individual patient physiologies and trends in patient information.

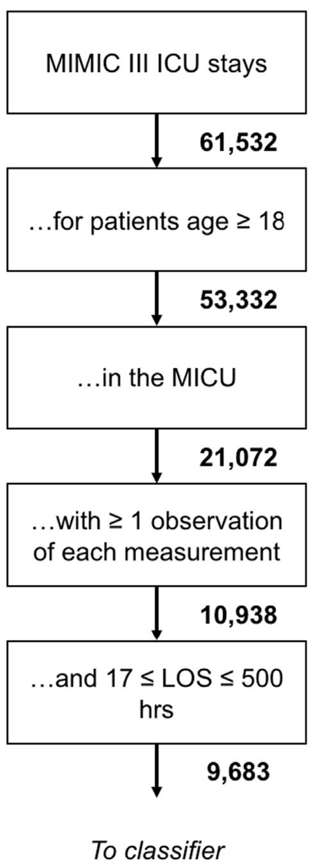

Fig. 1. Patient inclusion flowchart.

The increasing prevalence of Electronic Health Records (EHR) provides an opportunity to extract clinically relevant patient vital signs and laboratory results for increased predictive value in patient outcome [8]. In the ICU, a variety of relevant clinical measurements are available with high frequency and present a wealth of information regarding patient status and trends. Some recent studies have attempted to use these EHR data and trends to improve patient mortality predictions with computational algorithms, with some success [9–11]. In particular, analyses of time interval motifs have led to accurate predictions, and we build on this previous work in this study [12]. We present here a computational approach called *AutoTriage*, which not only utilizes patient clinical variables including vital signs, but also analyzes the correlations and trends between these measurements to provide information about patient stability. Using correlations among clinical variables allows us to achieve improved accuracy of patient stability prediction, using only eight very common measurements. *AutoTriage* provides an all-cause mortality prediction score 12 h in advance for ICU patients.

2. Methods

2.1. Data set

We used a dataset of 9683 patient records from the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) III database [13], which were selected according to the patient exclusion process depicted in Fig. 1. This subset consisted of anonymized clinical documentation of adult patients admitted to the Beth Israel Deaconess Medical Center (BIDMC) Medical Intensive Care Unit (MICU), with a variety of chief complaints (Table 1). The Institutional Review Boards of BIDMC and the Massachusetts Institute of Technology waived the requirement for individual patient consent, as the study did not impact clinical care and all data were deidentified.

Inclusion criteria for this study were:

I. Adult (i.e. age ≥18 years) admitted to the MICU.
II. Documented length-of-stay and survival for at least 17 h and fewer than 500 h following admission. The cutoff of 17-h observation was chosen to allow 12-h advance prediction based on at least 5 h of data. The limit of 500 h was chosen to reduce memory usage and the time cost of computations.

We utilized dynamic physiological measurements with a one hour timeresolution. Specifically, we used heart rate, pH, pulse

Table 1
Demographics of patient population over 18 years of age in the MICU of the MIMIC III database (20,108 total hospital admissions).

Demographic overview	Characteristic	Number of ICU stays	Percentage
Gender	Female	10,176	48.29%
	Male	10,896	51.71%
Age	18–29	984	4.67%
Median 64, IQR (51–78)	30–39	1328	6.30%
	40–49	2421	11.49%
	50–59	3717	17.64%
	60–69	4147	19.68%
	70+	8475	40.22%
Length of Stay (days)	0–2	13,646	64.76%
Median 2.1, IQR (1.2–4.1)	3–5	4057	19.25%
	6–8	1301	6.17%
	9–11	685	3.25%
	12+	1383	6.56%
Death During Hospital Stay	Yes	18,821	89.32%
	No	2251	10.68%