Rapid Response Systems

Individualizing and optimizing the use of early warning scores in acute medical care for deteriorating hospitalized patients

Muge Capan a,e, Julie S. Ivy a, Thomas Rohleder b,d, Joel Hickman b,d, Jeanne M. Huddleston b,c,d,∗

a Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Campus Box 7906, Raleigh, NC 27695-7906, USA
b Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Rochester, MN 55905, USA
c Division of Hospital Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, MN 55905, USA
d Division of Health Care Policy and Research, Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA
e Health Care Delivery Science, Value Institute, Christiana Care Health System, Newark, DE 19718, USA

A R T I C L E   I N F O
Article history:
Received 10 February 2014
Received in revised form
24 November 2014
Accepted 11 December 2014

Keywords:
Markovian
Deterioration
Prediction

A B S T R A C T
Aim: While early warning scores (EWS) have the potential to identify physiological deterioration in an acute care setting, the implementation of EWS in clinical practice has yet to be fully realized. The primary aim of this study is to identify optimal patient-centered rapid response team (RRT) activation rules using electronic medical records (EMR)-derived Markovian models.

Methods: The setting for the observational cohort study included 38,356 adult general floor patients hospitalized in 2011. The national early warning score (NEWS) was used to measure the patient health condition. Chi-square and Kruskal Wallis tests were used to identify statistically significant subpopulations as a function of the admission type (medical or surgical), frailty as measured by the Braden skin score, and history of prior clinical deterioration (RRT, cardiopulmonary arrest, or unscheduled ICU transfer). Results:

Results: Statistical tests identified 12 statistically significant subpopulations which differed clinically, as measured by length of stay and time to re-admission (P < .001). The Chi-square test of independence results showed a dependency structure between subsequent states in the embedded Markov chains (P < .001). The SMDP models identified two sets of subpopulation-specific RRT activation rules for each statistically unique subpopulation. Clinical deterioration experience in prior hospitalizations did not change the RRT activation rules. The thresholds differed as a function of admission type and frailty.

Conclusions: EWS were used to identify personalized thresholds for RRT activation for statistically significant Markovian patient subpopulations as a function of frailty and admission type. The full potential of EWS for personalizing acute care delivery is yet to be realized.

© 2015 The Authors. Published by Elsevier Ireland Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction
Hospitalized patients are at risk of unexpected clinical deterioration, usually characterized by a disturbance in physiology. Early detection of deterioration with appropriate treatment can reduce the risk of undesired clinical outcomes.1,2 Early recognition for the right patient in the right setting and providing care with the right provider are critical to deliver safe and effective acute care the first time.

Care providers have long embraced the use of early warning scores (EWS) to assess a patient’s clinical condition and enable timely response to APD.3,4 The primary aim of an EWS is to provide a simple, uniform method for categorizing a patient’s condition and guidance to indicate when a patient may require additional attention. EWS commonly rely on a system that scores physiological measurements obtained at admission or by regular monitoring during hospitalization. Widely used EWS such as the modified early warning score (MEWS), VitalPAC early warning score (VIEWS), and national early warning score (NEWS) not only aim to standardize patients’ clinical assessments, but also provide guidelines for clinical decisions.5–7 Specifically, an aggregate score above a certain threshold suggests the need to activate a clinical team with acute-care competencies, e.g., a rapid response team (RRT). However, the recommended EWS thresholds in literature are not personalized. For example, a cumulative NEWS value above

http://dx.doi.org/10.1016/j.resuscitation.2014.12.032
0300-9572/© 2015 The Authors. Published by Elsevier Ireland Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
7 suggests immediate emergency assessment by RRT and transfer of the patient to a higher level of care if necessary without considering patient characteristics.1 The performance of EWS-based RRT intervention may depend on the patient type and prior hospitalization experience. Capan and Ivy9 compared the performance of EWS in predicting future deterioration events using two vital sign related indicators. Their study indicated the predictive decision tool’s performance varied significantly depending on patient characteristics.

Although a number of studies have either analyzed the impact of medical emergency teams on deterioration events during hospitalization, or examined the use of track and trigger systems, such as EWS, to predict undesirable outcomes there are gaps in the literature associated with resuscitation decision making.9–13 While a number of EWS systems are currently used for detection and response to APD, there is a need to capture the uncertainty in the deterioration and recovery process throughout a patient’s hospitalization, due to the patient’s health, the occurrence of deterioration events, and the timing of these events.

Every EWS-based resuscitation decision requires adaptation, based on the care provider’s judgment and the individual patient’s characteristics. Hence, a natural next step in acute medical care is to use EWS for real-time decision support. The primary aim of this study is to identify optimal patient-centered RRT activation rules using electronic medical records (EMR)-derived Markovian models. To the best of our knowledge, there are no models for optimizing RRT activation considering patient characteristics and personnel resource utilization. We propose an analytical model for identifying optimal EWS-based RRT activation thresholds in an acute medical care setting. This study also identifies medically relevant, statistically significant subpopulations, and identifies subpopulation-based RRT activation thresholds.

2. Methods

2.1. Setting and patient population

Following approval from Mayo Clinic’s Institutional Review Board, we used a retrospective observational cohort of 38,356 adult (>18 years) general floor medical and surgical patients admitted to a single center (Mayo Clinic, Rochester, MN) from January through December 2011. Exclusion criteria included age at admission (<18 years) and care location (only patients admitted to the general floor, and observations collected only during general floor episodes are included).

We define “encounter” as a hospitalization, and “episode” as a period of time in the general care ward. An encounter includes one or more episodes. Patient subpopulations were identified by: (i) deterioration event in a previous episode within the same encounter, (ii) the risk of frailty at admission measured by BSS, and (iii) admission type (medical or surgical). The deterioration events considered included RRT activation, Code45 activation, or an unscheduled transfer to the ICU. Code45 is an emergent event where an individual experiences cardiac arrest or acute respiratory compromise requiring intubation. An unscheduled transfer to the ICU indicates the patient deteriorated to a point where monitoring and treatment intensity provided in the general ward were insufficient. The episodes without deterioration events were used to derive the model input for “no prior deterioration event” models, and episodes after an event were used for “prior deterioration event” models. Medical patients (87.2%) were characterized as those who were being treated non-operatively, e.g., they had a primary medical diagnosis. Surgical patients (12.8%) underwent an operative procedure during the hospitalization. Patients could change categories from medical to surgical during the course of a hospitalization. However, patients could not convert from surgical to medical. Once

<table>
<thead>
<tr>
<th>Table 1: Patient demographics.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Total number admitted</td>
</tr>
<tr>
<td>Gender (% female)</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Min.</td>
</tr>
<tr>
<td>25th percentile</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>75th percentile</td>
</tr>
<tr>
<td>Max.</td>
</tr>
<tr>
<td>Preexisting conditions</td>
</tr>
<tr>
<td>Myocardial infarct (%)</td>
</tr>
<tr>
<td>Congestive heart failure (%)</td>
</tr>
<tr>
<td>Peripheral vascular disease (%)</td>
</tr>
<tr>
<td>Cerebrovascular disease (%)</td>
</tr>
<tr>
<td>Dementia (%)</td>
</tr>
<tr>
<td>Chronic pulmonary disease (%)</td>
</tr>
<tr>
<td>Ulcer (%)</td>
</tr>
<tr>
<td>Mild liver disease (%)</td>
</tr>
<tr>
<td>All liver disease (%)</td>
</tr>
<tr>
<td>Diabetes (%)</td>
</tr>
<tr>
<td>Diabetes with organ damage (%)</td>
</tr>
<tr>
<td>Hemiplegia (%)</td>
</tr>
<tr>
<td>Moderate/severe renal disease (%)</td>
</tr>
<tr>
<td>Moderate/severe liver disease (%)</td>
</tr>
<tr>
<td>Metastatic solid tumor (%)</td>
</tr>
<tr>
<td>Other cancer (%)</td>
</tr>
<tr>
<td>Aids (%)</td>
</tr>
<tr>
<td>Rheumatologic disease (%)</td>
</tr>
<tr>
<td>Admission surrogate measure of frailty (Braden skin score)</td>
</tr>
<tr>
<td>Low risk for frailty (%)</td>
</tr>
<tr>
<td>Moderate risk for frailty (%)</td>
</tr>
<tr>
<td>High risk for frailty (%)</td>
</tr>
<tr>
<td>Previous hospitalization experience</td>
</tr>
<tr>
<td>Patients who experienced a RRT event (%)</td>
</tr>
<tr>
<td>Patients who experienced a Code45 event (%)</td>
</tr>
</tbody>
</table>

an operative procedure occurred, the patient was considered a surgical patient for the remainder of the hospitalization. Frailty is a multi-dimensional symptom which captured the physiological loss of reserve capacity.14,15 We utilized the Braden skin score (BSS) as a surrogate for frailty. We classified patients as low risk for frailty (BSS ≥ 23), moderate risk for frailty (12 ≤ BSS ≤ 22), or high risk for frailty (BSS ≤ 11). During the study period, 4833 deterioration events occurred (1941 RRT events, 211 Code45 events, and 2681 unscheduled transfers to the ICU). A summary of the patient demographics is presented in Table 1.

2.2. Measurement of physiological deterioration

NEWS was used as a measure of health condition and deterioration level because: (i) it has been shown to most accurately predict adverse outcomes,7,16 and (ii) it considers extreme values of a single component whereas many EWS only use the aggregate score for clinical evaluation. Table 2 presents the NEWS components and the weights associated with each physiological measure.

2.3. Statistical analysis

This study aimed to identify clinically relevant and statistically significant subpopulations as a function of patient-specific covariates (defined in Section 2.1). Using SAS Version 9.2 (SAS Institute, Inc., Cary, NC) patient subpopulations were identified using two nonparametric statistical tests: (i) the Chi-square test was used to test the homogeneity of a patient’s health condition evolution, and (ii) the Kruskal Wallis test was used to test the homogeneity of the holding times. The health condition evolution is the stochastic path
that a patient’s NEWS follows over time. Holding time is the time spent in a given health condition (or NEWS score). We conducted 30 hypothesis tests to determine if the transition probabilities are identical for the subpopulations and subsets of the subpopulations created using the three factors (defined in Section 2.1) in six different order combinations. For each of the 30 hypotheses tested, we tested the null hypothesis that holding times are identical as well. All test results provided sufficient evidence to reject the hypotheses that the considered subpopulations are identical. The Markovian property was tested using the Chi-square test of independence. The null hypothesis was independence. The results suggested a dependency structure between two subsequent model states in the embedded Markov chains.

2.4. Optimization

We developed infinite-horizon semi-Markov decision process (SMDP) models to represent the uncertainty in a patient’s health (as measured by NEWS) progression and identify optimal patient-specific, NEWS-based RRT triggers. The RRT-trigger (or threshold) identifies the NEWS value above which the patient may benefit from the expertise provided by the RRT. Markov decision process models (MDPs) are used to control stochastic systems. MDPs have been widely applied in healthcare to support screening, diagnosis, and treatment decisions, as well as patient flow and hospital operations optimization due to their simplicity and interpretability. An MDP allows the decision maker to influence the behavior of a probabilistic system through his/her actions as the system evolves through time.

A SMDP is an optimization model for multi-stage stochastic decision problems where the successive state occupancies are determined by the transition probabilities of a Markov process, and the holding time in any state is a random variable. SMDP models have been used to study the recovery process of patients with acute leukemia, coronary patients, end-stage renal disease patients, and to model the flow of pregnant patients through hospital units.

In the SMDP presented in this paper, the evolution of a patient’s health during an episode, as defined by the NEWS score, is characterized by a Markov process and described by the state transition probability matrix. A semi-Markov process allows for the time the patient spends in a given health state to follow any distribution. The “decision process” of a SMDP corresponds to the optimal selection of actions that can influence the patient’s natural health evolution (e.g., a call to RRT may change the probability that a patient moves to a worse NEWS state). The optimization seeks to minimize (or maximize) a metric which captures the effect of the action taken on the state of the SMDP. In this context we select actions (call RRT or wait) with the goal of minimizing the total expected time to stabilization.

For the infinite-horizon SMDP models, the NEWS guidelines were used to inform the state definition for the semi-Markov process. Expert opinion was used to aggregate the NEWS scores into states. Specifically, for the Markov chain, the patient health states, \( s \in \{5, 4, 3, 2, 1, 0\} \), correspond to NEWS values \( 0, 0, [1-4] \) or a single extreme value, \( [5-6], \geq 7 \), end of episode), respectively. State 5 represents the healthy condition. State 4 corresponds to a stabilized patient after a critical care intervention. The health states \( \{3, 2, 1\} \) correspond to the slightly concerning, concerning, and critical distress conditions. State 0 represents end of episode, i.e., discharge from the general floor and corresponds to: discharge alive, hospice, death, and transfer to ICU. There are two types of ICU admissions in the model: (i) a transfer to the ICU within the same episode ends the episode and is represented as an absorbing state (state 0) in the Markov chain, (ii) an ICU admission in a prior episode within the same encounter is defined as a deterioration event for the “prior deterioration event” subpopulation. Fig. 1 illustrates possible health state transitions for a sample patient during an episode. The patient has a NEWS value of 0 (state 5) at the beginning of the episode, and the episode ends when the patient leaves the general ward (i.e., transitions to state 0). State transition matrices and holding time distributions for each subpopulation are derived from the EMR using maximum likelihood estimates.

For the SMDP models, a decision epoch corresponds to the time period during which a decision is made, i.e., the time a bedside provider team enters a patient’s room during routine hospital rounding. There are two possible actions at each decision epoch: wait, or to initiate RRT. The holding times are assumed to be exponentially distributed as derived from EMR. The objective of the SMDP model is to minimize the total expected resource use, time to stabilization (TTS) and failure to rescue (FTR) times by choosing the best RRT initiation policy. TTS is defined as the time from the start of an APD episode (identified by NEWS >0 for \( \geq 30 \) min) until the patient was considered stabilized (NEWS value of 0 for \( \geq 1 \) h). FTR...
is defined as a patient’s NEWS remaining above 7 for ≥1 h without RRT activation.

For each subpopulation, the SMDP model incorporates changes in the health condition, the time spent in a given health condition, the nurse resource time (nursing classification data was used to determine nurse resource requirements based on the patient severity), and the RRT resource time (time from RRT arrival to departure from the bedside). The SMDP model trade-offs the costs associated with false negative (e.g., TTS due to delayed response) and false positives (e.g., unnecessary RRT calls) to identify the optimal actions. The cost associated with a state is defined in terms of time with elements that depend on the state in a NEWS category (TTS and TTR), and state-specific elements (nurse and RRT resource time). Costs were informed by EMR and nursing classification data. These inputs were used to derive the optimal policy resulting from the SMDP model algorithm. The episodes with no prior deterioration events (i.e., ICU admission, Code45, or RRT call in a prior episode within the same encounter) were used to derive the model input for “no prior deterioration event” subpopulation models, and episodes after a deterioration event were used for “prior deterioration event” subpopulation models. The EMR in our study indicated that a higher NEWS was associated with increased nurse and RRT utilization.

3. Results

The Chi-square test for the homogeneity of the evolution of patient health for each subpopulation and the Kruskal Wallis test for the homogeneity of holding times identified 12 statistically significant patient subpopulations (P < .001). The Markovian hypothesis tests indicated a significant first-order dependency between two subsequent health states (P < .001). As shown in Fig. 2, the statistically different subpopulations behaved differently in clinical practice as measured by length of stay (LOS) during an episode. Fig. 2a and b shows the LOS for medical patients and surgical patients, respectively, by their risk of frailty and prior adverse event experience. Note, a medical patient will be reclassified as surgical after surgery. The time periods of care in the general ward before and after surgery represent two episodes with distinct LOS. Fig. 2 indicates that surgical patients had longer episodes within encounters compared to medical patients. Further, BSS level at admission affected LOS. A higher BSS level for both medical and surgical patients was associated with a longer LOS.

Fig. 3 shows expected total cost measured in hours for each patient subpopulation and health state. For a given subpopulation, the total expected cost (i.e., the time required to care for the patient) increases as a patient’s condition declines. Furthermore, the rate of increase differs by subpopulation. Surgical patients require more resource intensity than medical patients, and higher BSS is associated with increased resource intensity.

Fig. 4 shows the subpopulation-specific optimal RRT thresholds. Fig. 4 indicates there are two categories of patients with distinct RRT thresholds. For example, a highly frail surgical patient without previous deterioration events would benefit from RRT activation when the NEWS is [1–4], whereas the threshold is NEWS ≥ 7 for a moderately frail medical patient. Fig. 4 suggests the prior deterioration event experience does not change the optimal threshold, i.e., medical and moderate risk patients have the same threshold (state 1) regardless of prior deterioration event experience. However, admission type and BSS affect the optimal thresholds.

Fig. 2. Cumulative percentage of LOS in days for medical and surgical patients. (A) Patients with low risk of frailty and without prior deterioration event. (B) Patients with moderate risk of frailty and without prior deterioration event. (C) Patients with high risk of frailty and without prior deterioration event. (D) Patients with low risk of frailty and with history of prior deterioration event. (E) Patients with moderate risk of frailty and with history of prior deterioration event. (F) Patients with high risk of frailty and with history of prior deterioration event.

Fig. 3. Total expected time-based costs in hours by subpopulation including resource intensity, time to stabilization (TTS) and failure to rescue (TFR). (A) Medical patient with low BSS at admission and without prior deterioration event. (B) Medical patient with moderate BSS at admission and without prior deterioration event. (C) Medical patient with low BSS at admission and with prior deterioration event. (D) Medical patient with moderate BSS at admission and with prior deterioration event. (E) Surgical patient with moderate BSS at admission and with prior deterioration event. (F) Medical patient with high BSS at admission and with prior deterioration event. (G) Surgical patient with high BSS at admission and with prior deterioration event. (H) Surgical patient with low BSS at admission and without prior deterioration event. (I) Surgical patient with moderate BSS at admission and without prior deterioration event. (J) Medical patient with high BSS at admission and without prior deterioration event. (K) Surgical patient with high BSS at admission and with prior deterioration event. (L) Surgical patient with low BSS at admission and with prior deterioration event.
4. Discussion

We hypothesized that the patient population is heterogeneous and the personalized use of EWS for the RRT activation may improve acute care delivery. The Chi-square and Kruskal Wallis tests identified 12 subpopulations that exhibited clinically different behavior during hospitalization as measured in LOS. These results highlight the need for patient centricity in EWS systems – also a focus in the literature. Identifying the patient characteristics, which are readily available at admission and throughout hospitalization, to create medically coherent and statistically different patient groups, will enable the establishment of frameworks for the individualized implementation of EWS in the acute care environment. The SMDP model is a first step toward such framework and the results show that while a previous deterioration event does not impact the optimal RRT activation policies, the RRT thresholds change as a function of admission type and risk of frailty. This translates into a simple two decision rules to personalize the RRT activation decision. The provider team can identify a patient’s group using the three covariates, and follow the optimal RRT activation rules for that group.

The SMDP models also provide insight regarding hospital personnel resource utilization. The total expected resource intensity, including nurse and RRT utilization, TTS, and FTR, increases as a patient’s condition deteriorates. Further resource intensity differs by patient type and risk of frailty, i.e., surgical patients require more resource intensity than medical patients, and increased risk of frailty at admission was associated with higher resource intensity. These results align with studies that have shown more frail patients are at increased risk of deterioration and provide insight for the hospital management to plan for the resource needs of patients as a function of the health condition. Our findings highlight the potential for improvement in bedside resuscitation by integrating EMR-based models and clinician judgment. Further, we identify key factors for classifying statistically significant subpopulations to personalize acute medical care. Hospitals and other care delivery systems can benefit from our methodology for incorporating large-scale EMR into clinical practice to support real-time resuscitation decision making.

Our study has some limitations. Foremost, we relied on data collected retrospectively define patient states. Specifically, we assumed that the call to activate the RRT, or transfer the patient to an ICU, were for clinically valid reasons. We did not second-guess the clinical decisions made at the time, nor as a result, did it matter which RRT criteria were met leading up to the trigger. This may have significantly inflated the estimated number of deteriorations because not all RRT calls may have been valid clinical deteriorations. Another limitation was data from a single institution; thereby impacting the generalizability of these results. Third, our analysis requires documentation of vital signs. The TTS and FTR times may be incorrectly calculated because there are times when the vital signs are manually assessed but not recorded. Fourth, we focused on general care floor patients with the goal of informing clinical providers at the point of care. Therefore we cannot extrapolate the findings to patients in cardiac monitored settings. However, our approach could be extended by including data from additional care levels in a hospital, as well as from other medical systems to test generalizability.

Another limitation is that we utilized BSS as a surrogate for risk of frailty. Practically, patients at risk for pressure sores seem more likely to be frail. From a literature perspective, the BSS contains two of the major constructs used to define frailty (mobility and nutrition status). We may have incorrectly categorized frailty risk in these patients making interpretation of actions at different states difficulty. Likewise, we utilized NEWS to represent a patient’s health condition. NEWS is a convenient choice from the bedside providers’ perspective because the components of the score are readily available. NEWS also provides a more realistic view of the acute care decision process because the score includes a single, physiological measure’s extreme value in addition to the aggregate score to initiate interventions, whereas other EWS commonly rely only on the aggregate score. The model states are created based on real-time score categories and the proposed approach can be easily applied to other EWS systems. However, changing a patient’s EWS classification real-time adds significant complexity for the bedside provider. If implemented, it would not be as simple as remembering the score components. The provider would also have to consider the BSS and timing of any surgical procedures during the hospitalization. Thus, implementation of personalized models may result in cognitive error with misclassification of patient groups. Such implementation may require bedside clinical decision support.

We present a model to inform real-time dynamic decision support for caregivers to support patient-centered clinical decisions during hospitalization for recognizing and responding to APD. Comparison and improvement of EWS-based clinical decision-aids to support bedside resuscitation represent a promising future research area. The potential of EWS in personalized acute medical care delivery is yet to be fully realized.
Conflict of interest statement

All authors disclose that the Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery and North Carolina State University, Raleigh, NC do not have any input into study design, data analysis and result interpretation.

Acknowledgements

This research was performed in collaboration with Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery and North Carolina State University, Raleigh, NC.

References