Improving Recurrent Neural Network Responsiveness to Acute Clinical Events

David Ledbetter, Eugene Laksana, Melissa Aczon, and Randall Wetzel

Abstract—Predictive models in acute care settings must be able to immediately recognize precipitous changes in a patient’s status when presented with data reflecting such changes. Recurrent neural networks (RNN) have become common for training and deploying clinical decision support models. They frequently exhibit a delayed response to acute events. New information must propagate through the RNN’s cell state memory before the total impact is reflected in the model’s predictions. This work presents input data perseveration as a method of training and deploying an RNN model to make its predictions more responsive to newly acquired information: input data is replicated during training and deployment. Each replication of the data input impacts the cell state and output of the RNN, but only the output at the final replication is maintained and broadcast as the prediction for evaluation and deployment purposes. When presented with data reflecting acute events, a model trained and deployed with input perseveration responds with more pronounced immediate changes in predictions and maintains globally robust performance. Such a characteristic is crucial in predictive models for an intensive care unit.

Index Terms—Recurrent Neural Network, Long Short-Term Memory, Acute Clinical Events, Electronic Medical Records.

1 INTRODUCTION

1.1 The Problem

Critical care environments require rapid decision making. To be meaningful, predictive models in these settings must immediately recognize precipitous changes in a patient’s state when presented with data reflecting such changes [1], [2]. Newly acquired information must be integrated quickly with the existing understanding of the patient’s state to inform clinical decisions.

Long Short-Term Memory (LSTM) models [3] are a type of Recurrent Neural Network (RNN) architecture and have become increasingly popular for modeling tasks in clinical settings, including Intensive Care Units (ICUs) [4], [5], [6], [7], [8], [9], [10], [11]. It is occasionally noted that RNNs, including LSTM models, exhibit some lag after acquiring new information [12]. This characteristic is frequently observed in stock prediction tasks where the apparent predictive capability of a model is comparable to a model that predicts the last observed stock price [13], [14]. However, the intersection of deep learning and healthcare, though ever widening, has a paucity of literature on this phenomenon. Commonly reported metrics such as Area Under the Receiver Operating Characteristic Curve, sensitivity, specificity, and mean absolute errors mask certain behaviors of RNN based models.

RNN models require computational cycles to incorporate new data into their memory cell [15]. There are two primary observable effects of this, summarized briefly here. First, what we refer to as a pipe-up period, requires the initial input to propagate through the memory cell to overcome the initialized cell state before that input significantly influences the predictions. Second, during all subsequent times, new information only slightly changes the overall prediction of a model at the time of input and requires propagation through the memory cell before fully augmenting the model’s prediction. The first phenomenon is a special case of the second. These two effects are further illustrated in Sections 5 and 6.

1.2 Proposed Solution

This work aims to enable more pronounced changes in RNN predictions when the model is presented with data indicating acute clinical events by perseverating the input. Instead of giving newly available input data to an RNN model only once, the input is replicated and given to the model multiple times, during both training and deployment, with only the prediction of the final perseveration maintained and made visible to the end-user. This approach provides additional computational cycles for new data to be incorporated into the model’s internal states before a prediction is broadcast. We hypothesized that the resultant model would be able to react to acute events more quickly than traditionally trained LSTMs and still maintain overall model performance.

2 RELATED WORKS

Since the seminal paper by R.E. Kalman in 1960 describing what is now called the Kalman Filter [16], engineers have been aware of the trade off between integrating historical data and responding to new data. Most training techniques used in modern deep learning affect the balance between relying on historical trends and responding to new information; such techniques include dropout, optimizers, and activation functions [17], [18], [19], [20]. Generating an appropriate target vector (e.g., using changes in values between consecutive time points instead of predicting raw values) is another training technique that can prevent generation of an auto-correlation model [12].
Attention mechanisms [21] are sometimes used to manage the balance between historical and new data by appending the entire input sequence to the final hidden layer of an RNN. Doing so affords the model another opportunity to learn weights which expose the moments in time series data most relevant to the predictions. Attention networks were originally developed for post-hoc sequence-to-sequence modeling tasks such as image captioning and neural machine translation which permit access to the entire input sequence when making predictions [21], [22]. An attention mechanism was applied by Zhang, et al. to predict risk of future hospitalization using medical records from a fixed observation window. [23].

3 DATA

3.1 Clinical Data Sources

Data for this work were extracted from de-identified observational clinical data collected in Electronic Medical Records (EMR, Cerner) in the Pediatric Intensive Care Unit (PICU) of Children’s Hospital Los Angeles (CHLA) between January 2009 and February 2019. The CHLA Institutional Review Board (IRB) reviewed the study protocol and waived the need for IRB approval. A patient record included static information such as gender, race, and discharge disposition at the end of an ICU episode, defined as a contiguous admission in the PICU. A patient may have multiple episodes. The EMR for an episode also contained irregularly, sparsely and asynchronously charted measurements of physiologic observations (e.g., heart rate, blood pressure), laboratory results (e.g., creatinine, glucose level), drugs (e.g., epinephrine, furosemide) and interventions (e.g., intubation, oxygen level). Episodes without discharge disposition were excluded, leaving 12,826 episodes (9,250 patients) in the final dataset.

Prior to any of the computational experiments described here, the episodes were randomly partitioned into three sets: a training set for deriving model weights, a validation set for optimizing hyper-parameters, and a holdout test set for measuring performance. To minimize bias in the performance evaluation metrics, partitioning was done at the patient level, i.e. all episodes from a single patient belonged to only one of these sets: 60% in the training set, 20% in the validation set, and 20% in the test set. No other stratification was applied. Table 1 displays basic characteristics of the resulting data partitions.

Preprocessing steps described in previous work [24] converted each episode data to a matrix structure amenable to machine learning model development. A row of values in this matrix represented measurements (recorded or imputed) of different variables at one particular time, while a column contained values of a single variable at different times. A complete list of variables used as model inputs can be found in Appendix A. Note that diagnoses, while available in the EMR, were not used as input features.

3.2 Target Outcome

ICU mortality was chosen as the target outcome because it is a simple, unambiguous outcome. Importantly, risk of mortality is used commonly as a proxy for severity of illness (Mortality is used commonly as a proxy for severity of illness.

| Table 1: Basic demographics of data used in this study for each data partition. |
|------------------|------------------|------------------|------------------|
| **Number of episodes** | **training** | **validation** | **test** | **overall** |
| **Number of patients** | 5551 | 1850 | 1850 | 9251 |
| **Mortality rate** | 0.035 | 0.037 | 0.043 | 0.037 |
| **Fraction male** | 0.56 | 0.55 | 0.58 | 0.56 |
| **Age mean (years)** | 7.99 | 8.23 | 8.09 | 8.06 |
| **std dev** | 6.45 | 6.51 | 6.35 | 6.44 |

Fig. 1: The perseverating recurrent neural network (PRNN) is an RNN with repeating inputs. Each input vector associated with a particular time is replicated \( k \) times, and only the output from the last replication is considered as the prediction for that time. The diagram above illustrates the process for \( k = 3 \).

The overall mortality rate of the data was 3.7% (Table 1).

4 RNN MODELS

Many-to-many recurrent neural network models, consisting of stacked Long Short-Term Memory (LSTM) layers followed by a dense layer, were trained to predict ICU mortality of each patient episode. All the models output a probability of survival at each distinct time point where an observation or measurement of the patient was made, generating a trajectory of scores that reflect the changing status of a patient during their ICU episode.

The baseline RNN model was trained in the traditional manner: when the model acquires new data \( x(t_i) \), it generates and immediately broadcasts a prediction \( y(t_i) \). Other models which share the architecture of the baseline RNN model were trained using input perseveration: the same \( x(t_i) \) vector was repeatedly given as input to the RNN model \( k \) times, where \( k \) is a controllable parameter. While all \( k \) outputs – corresponding to the \( k \) times that \( x(t_i) \) was given to the model – are used for optimization during training, only the last one is maintained and broadcast as the prediction at time \( t_i \) during performance assessment and deployment. We call these models perseverating recurrent neural networks (PRNN). Figure 1 illustrates a PRNN model with \( k = 3 \). Note that the baseline RNN can be considered as a PRNN with \( k = 1 \). Input perseveration provides the internal cell state memory of the RNN additional computational cycles.
Fig. 2: Overview of a standard RNN model with a causal attention layer.

TABLE 2: Hyperparameters for all permutations of the PRNN.

| Hyperparameter                  | Value                          |
|--------------------------------|--------------------------------|
| Number of LSTM Layers          | 3                              |
| Hidden Units in LSTM Layers    | [128, 256, 128]                 |
| Batch Size                     | 32                             |
| Initial Learning Rate          | 1e-5                           |
| Loss                           | binary cross-entropy            |
| Optimizer                      | rmsprop                        |
| Dropout                        | 0.2                            |
| Recurrent Dropout              | 0.2                            |
| Regularizer                    | 0.0001                         |
| Output Activation              | sigmoid                         |

5 Model Assessments

Standard metrics such as the Area Under the Receiver Operating Characteristic Curve (AUROC), precision, and recall scores for a binary classification task can capture a model’s overall predictive performance. However, they can mask the lag phenomenon displayed by LSTMs that, although known, is rarely commented on in the literature. The trajectory of probability of survival predictions should reflect the evolving state of a patient which can have instantaneous changes. Metrics were designed to quantify a model’s pipe-up behavior and its ability to capture rapid changes resulting from clinically adverse events. AUROC was used to compare overall predictive performance for the main task – ICU mortality prediction.

5.1 Pipe-Up Behavior

A model’s prediction at the first time step is a function of its weights, initial state of its memory cells and the first input. Of these, only the first input varies across different patient episodes; therefore, the distribution of a model’s prediction at the first time point of all episodes indicates the model’s level of reliance on the first input data. We refer to this period as the pipe-up period.

The mean and standard deviation of the distribution of all \( \hat{y}_p(t_0) \) predictions, where \( \hat{y}_p(t_0) \) represents the first prediction for patient episode \( p \), were computed for all survivors and non-survivors in the test set. These metrics were computed for all six RNN models described in the previous section and used to compare their responsiveness to the first available information about a patient episode.

5.2 Responsiveness to Acute Events

The models were also assessed for their instantaneous responses to clinically adverse events using an average variation metric computed from the predictions during such periods. In consultation with clinicians, the times when any of the following occurred were defined as acute time points or intervals:

1) substantial increases in creatinine levels or inotrope score between two consecutive recordings;
2) substantial decreases in heart rate, mean arterial pressure, Glasgow Coma Score, blood oxygen level (SpO2), arterial blood gas (ABG) pH, and venous blood gas (VBG) pH between two consecutive recordings;
3) a patient’s heart rate goes to 0.

We quantified substantial changes as those that were in the top or bottom \( X \) percentile \( (X \in \{5.0, 1.0, 0.5\}) \) of inter-measurement changes. For example, a time point \( t_i \) was considered acute in terms of creatinine when the increase in creatinine level from \( t_{i-1} \) to \( t_i \) was in the top \( X \) percentile of all creatinine level percent changes between any two consecutive time points in the test dataset. Similarly, acute time points for heart rate were those times when the decrease in heart rate was in the bottom \( X\% \) of all heart rate changes between any two consecutive time points of all test set episodes. Table 3 shows the specific percent change thresholds indicating acute changes in a patient. Thus, a 32%
TABLE 3: Thresholds of percent change used to define acute changes in patient state. Negative thresholds correspond to the bottom $p\%$ of changes and positive thresholds correspond to the top $p\%$ of changes, for $p \in [5, 1, 0.5]$.

| Metric                | 5%     | 1%     | 0.5%   |
|-----------------------|--------|--------|--------|
| Heart Rate            | -18.6  | -31.6  | -37.1  |
| Mean Arterial Pressure| -23.5  | -36.1  | -44.2  |
| Glasgow Coma Score    | -50.0  | -72.7  | -72.7  |
| Pulse Oximetry        | -5.6   | -20.2  | -35.1  |
| ABG pH                | -1.5   | -2.9   | -3.5   |
| VBG pH                | -1.4   | -2.6   | -3.1   |
| Inotrope Score        | 66.7   | 150.0  | 250.0  |
| Creatinine            | 33.3   | 60.5   | 81.8   |

drop in heart rate between two consecutive measurements would be in the top 1 percentile, while a 70% increase in inotrope score would be in the top 5 percentile.

Applying three percentile-based thresholds to define acute patient state changes meant a total of 25 acuity definitions (2 * 3 increases in measured variables, 6 * 3 decreases in measured variables, and zero heart rate) used to assess models’ responses when such events occur. It is important to note that we are not proposing these as general definitions of clinical acuity. They were designed to capture events that indicate precipitous changes in a patient’s state with high specificity to assess the magnitude of the instantaneous change in predicted probability of survival.

Each previously defined acute event, denoted by $E$, identified a set of acute time points, denoted by $S_E(p)$, for each patient episode $p$. Changes in a model’s prediction at these time points were evaluated through an average temporal variation metric given by:

$$V_E(p) = \frac{100}{|S_E(p)|} \sum_{t_i \in S_E(p)} |\hat{y}_p(t_i) - \hat{y}_p(t_{i-1})|,$$

where $\hat{y}_p(t_i)$ is the prediction for patient episode $p$ at time point $t_i$, and $|S_E(p)|$ is the number of time points in $S_E(p)$. This metric measures how much the predicted probability of survival (scaled to $[0, 100]$) changed, on average, at the defined acute time points of an episode. Model $A$ having a higher $AV(p)$ than model $B$ means that when both models were presented with data reflecting a precipitous change in a patient’s state, model $A$’s prediction underwent a greater change than model $B$’s prediction, indicating that model $A$ had a more pronounced response to the acute event. The average of this metric across $N_p$ episodes gives a measure of the model’s overall responsiveness to acute event $E$ of those episodes:

$$V_E = \frac{1}{N_p} \sum_p V_E(p).$$

5.3 Overall Predictive Performance

Model predictions at ICU admission and at 1, 3, 6, 12, and 24 hours after ICU admission were evaluated for predictive performance via the AUROC. Performance was only measured for episodes lasting at least 24 hours to maintain a consistent cohort across time slices ($n = 2130$, mortality rate = 0.046).

6 RESULTS

Figure 3 illustrates trajectories of probability of survival predictions from two models, the baseline RNN ($k = 1$) and a PRNN with $k = 5$, for two non-surviving episodes. In the first episode (top), the patient had a rapid increase of creatinine level, which indicates kidney dysfunction, about 500 hours after ICU admission. When presented with this defined acute event, the baseline model’s prediction dropped from 0.89 to 0.88, while the PRNN model dropped from 0.75 to 0.71. In the second episode (bottom), the patient’s Glasgow Coma Score decreased from 7 to 4. When presented with this defined acute event, the baseline model’s probability of survival prediction dropped from 0.76 to 0.70, while the PRNN model dropped from 0.69 to 0.50. In either case, the PRNN model’s immediate response to an acute event was more pronounced than that of the baseline RNN model. Further, the trajectory of the baseline RNN’s predictions after either acute event appears to lag behind that of the PRNN by about 1-2 hours.

The remainder of this section describes the results from aggregating the assessment metrics described in Section 5 across episodes in the test dataset.

6.1 Pipe-Up Behavior

Figure 4 shows the distribution of predictions at the first observation. The mean prediction for survivors increased with the perseveration parameter $k$: from 0.70 ($k = 1$) to 0.90 ($k = 5$). For both survivors and non-survivors, the standard deviation increased with $k$. The increase of standard deviation was more pronounced in the non-survivors, going from from 0.04 at $k = 1$ to 0.13 at $k = 5$.

6.2 Responsiveness to Acute Events

Figure 5 displays the average variation as a function of $k$ and acuity definitions. Two trends are apparent. First is the behavior of average variation as a function of percentile change in a given physiologic observation or intervention. For each model, the resulting average variation increased as the change in a physiologic or intervention variable became more severe (i.e. from 95.0 to 99.5 percentile). Second, a clear positive correlation between $k$ and average variation is evident across all definitions of acuity. The attention layer generally demonstrated lower average variation than the $k = 1$ model.

6.3 Overall Predictive Performance

Table 4 summarizes the AUROC of model predictions at different times after ICU admission of all episodes in the test set that lasted at least 24 hours. The AUROC gains due to input perseveration were larger at the earlier times of prediction ($t \leq 6$ hours), increasing at the first hour from 0.77 when $k = 1$ to 0.83 when $k = 5$. By the 12th hour, the AUROC did not significantly change with the perseveration parameter $k$. Performance increases for all models as additional hours of observation are available. Adding an attention layer displayed similar performance as the baseline model ($k = 1$) at all evaluation times.
7 Discussion

Real-time predictions for ICU mortality are proxies for severity of illness [28], [32], [33] and should reflect acute changes in a patient. The examples in Figure 3 show that the mortality model trained with input perseveration (PRNN with \( k = 5 \)) responded more pronouncedly and immediately than the traditionally trained model (\( k = 1 \)) when both were presented with data reflecting acute changes. Subsequent to the acute changes, the responses of the traditional RNN appeared to lag behind the PRNN’s. Standard metrics such as AUROCs for classification or mean absolute errors for continuous regression often mask this predictive lag and other deleterious behaviors that can be detrimental in critical or intensive care settings where rapid recognition and response are crucial.

Input perseveration provides the RNN’s internal cells additional computational cycles to incorporate the current state of the patient into the final prediction. The effect of this method was measured by metrics designed to capture a model’s immediate responses to newly acquired data.

The variation metric described in Section 5.2 compares the models’ immediate responses to data indicating precipitous changes in a patient’s status. These changes require quick reaction time from the care team, therefore capturing them in the predictions is important. The comparison of variation metrics from the different models (Figure 5) shows that the LSTM became more responsive to acute clinical events as the level of perseveration, \( k \), increased. When \( k \) increased from 1 to 5, the variation metric corresponding to the defined precipitous events increased by a factor of 2-3 times. They also show that a given model’s responsiveness increased with more acute events (i.e. those belonging to higher percentile changes for a given physiologic or intervention variable). This result is consistent with expectations about the variation metric. For example, one would expect the change in predictions of those with larger drops in blood gas level to be greater than those of less severe drops.

Figure 4 compares the distributions of initial predictions generated by the different models. Since the first prediction is a function of both the initial input (which varies across episodes) and the initial cell state memory (which is fixed), a wider distribution of these predictions across episodes indicates a higher reliance on the initial input. This is important because children admitted to the PICU have different severities of illness [25], [27]. Increasing \( k \) resulted in a wider distribution as measured by the standard deviation. The increase was greater for the non-surviving population (\( \sigma = 0.04 \) when \( k = 1 \) to \( \sigma = 0.13 \) when \( k = 5 \)) relative to the surviving population. This is again consistent with clinical expectations.

Increasing the first prediction’s reliance on the initial measurement – as achieved by the PRNN models with higher \( k \) – also resulted in higher AUROC at ICU admission (Table 4). As the models observed the patient longer, their AUROCs increased as expected. Importantly, the increase of AUROC due to perseveration was greater at the early hours (\( t \leq 3 \)) when information is most scarce, with the 1-hour AUROC increasing from 0.77 when \( k = 1 \) to 0.83 when \( k = 5 \). This means that predictive clinical models relying

**Table 4:** Test set AUROCs for the mortality prediction at admission and after 1, 3, 6, 12, and 24 hours of observation. Only the patient episodes lasting at least 24 hours of ICU time were evaluated.

| \( k \) | 0hr | 1hr | 3hr | 6hr | 12hr | 24hr |
|-------|-----|-----|-----|-----|------|------|
| 1     | 0.716 | 0.766 | 0.859 | 0.895 | 0.914 | 0.952 |
| 2     | 0.729 | 0.804 | 0.871 | 0.902 | 0.916 | 0.956 |
| 3     | 0.731 | 0.819 | 0.875 | 0.902 | 0.916 | 0.955 |
| 4     | 0.735 | 0.821 | 0.876 | 0.900 | 0.915 | 0.954 |
| 5     | 0.736 | 0.832 | 0.883 | 0.905 | 0.919 | 0.955 |
| attention | 0.723 | 0.734 | 0.853 | 0.892 | 0.910 | 0.949 |
on scarce data for early detection could benefit from the perseveration approach.

Adding a causal attention layer to the baseline ($k = 1$) model had no apparent performance improvement in any of the metrics. The attention network theoretically can put more weight to the most recent state than to the previous ones. The results indicate that this mechanism did not improve on what the baseline LSTM’s gates were already doing, but persistently giving the same input to the model – i.e. perseverating the input – did.

There are limitations to the perseverated input approach. Although LSTMs are theoretically good at understanding temporal trends through their memory cells, there remain practical limitations to how long prior information can be maintained to inform future predictions [34], [35]. The PRNN has the potential to exacerbate these algorithmic deficiencies because of the memory cell’s prolonged exposure to the same data. A comprehensive assessment of the impact of a potential reduction in temporal memory was not performed. However, the basic data perseveration technique can easily be generalized to any sequential data.

Additionally, perseveration increases the number of sequences requiring computation. Perseveration is used during both training and inference, and the compute time scales linearly for both, commensurate with the level of perseveration. However, the time required for inference is on the order of $30 ms$ for $k = 1$ (baseline) and $150 ms$ for $k = 5$ on an NVIDIA Titan RTX. Since the recording frequency is approximately every 15 minutes, these computational burdens do not hinder deployment.

Finally, this proof-of-concept study used only a single clinical outcome (ICU mortality) and data from a single
center. Future work will examine the effect of perseverating the data input on other important clinical tasks such as risk of desaturation, sepsis, and renal failure.

8 CONCLUSION
This work demonstrates that perseverated data input increases the responsiveness of LSTM models to a variety of acute changes to patient state and also significantly increases predictive performance in the early hours following admission. The PRNN is a simple solution to the predictive lag exhibited by standard LSTMs when encountering an acute event and may enable more rapid responses to critical conditions.

ACKNOWLEDGEMENTS
This work was supported by the L. K. Whittier Foundation. The authors would like to thank Alysia Flynn for help aggregating the data and Mose Wintner for reviewing the manuscript and providing valuable feedback.
### Appendix A
**EMR Variables**

| Demographics and Vitals | Labs |
|-------------------------|------|
| **Demographics** | **Laboratory Tests** |
| Age | ABG Base excess | CBG Base excess |
| race_Asian/Indian/Pacific Islander | CBG PCO2 | CBG PCO2 |
| Abdominal Girth | CBG PO2 | CBG PO2 |
| Activity Level | CBG TCO2 | CBG TCO2 |
| Bladder pressure | CBG pH | CBG pH |
| Capillary Refill Rate | ABG PCO2 | ABG PCO2 |
| Central Venous Pressure | ABG O2 sat | ABG O2 sat |
| Cerebral Perfusion Pressure | ABG PO2 | ABG PO2 |
| Diastolic Blood Pressure | ABG TCO2 | ABG TCO2 |
| EtcO2 | ABG pH | ABG pH |
| Extremity Temperature Level | ALT | ALT |
| Eye Response Level | AST | AST |
| FLACC Pain Activity | Albumin Level | Albumin Level |
| FLACC Pain Consolability | Amylase | Amylase |
| FLACC Pain Cry | Anti-Xa Heparin | Anti-Xa Heparin |
| Height | B-type Natriuretic Peptide | B-type Natriuretic Peptide |
| Hemofiltration Fluid Output | BUN | BUN |
| Intracranial Pressure | Bands % | Bands % |
| Left Pupil Size After Light | Basophils % | Basophils % |
| Middle Pupil Size Before Light | Bicarbonate Serum | Bicarbonate Serum |
| Motor Response Level | Bilirubin Conjugated | Bilirubin Conjugated |
| Nasal Flaring Level | Bilirubin Total | Bilirubin Total |
| Near-Infrared Spectroscopy SO2 | Bilirubin Unconjugated | Bilirubin Unconjugated |
| Nutritional Index | Blaasts % | Blaasts % |
| Oxygen Index | C-Reactive Protein | C-Reactive Protein |
| Patient Mood Level | CBG Base excess | CBG Base excess |
| Pulse Oximetry | CBG Fio2 | CBG Fio2 |
| Quality of Pain Level | CBG HCO3 | CBG HCO3 |
| Right Pupil Size Before Light | CBG O2 sat | CBG O2 sat |
| Right Pupil Size Before Light | CBG Fio2 | CBG Fio2 |
| Respiratory Effort Level | CBG HCO3 | CBG HCO3 |
| Respiratory Rate | CBG O2 sat | CBG O2 sat |
| Right Pupil Size Before Light | CBG PO2 | CBG PO2 |
| Right Pupil Size Before Light | CBG TCO2 | CBG TCO2 |
| Systolic Blood Pressure | CBG pH | CBG pH |
| Sedation Scale Level | CBG PO2 | CBG PO2 |
| Skin Turgor Edema | CBG TCO2 | CBG TCO2 |
| Skin Turgor Turgor | CBG Fio2 | CBG Fio2 |
| Temperature | CBG O2 sat | CBG O2 sat |
| Verbal Response Level | CBG HCO3 | CBG HCO3 |
| WAT1 Total | CBG PO2 | CBG PO2 |
| Weight | CBC | CBC |

**Table 5**: EMR variables (demographics, vitals and labs) in patient episode matrix. Demographics such as gender and race/ethnicity were encoded as binary variables.
| Drugs                          | Interventions                                                                 |
|-------------------------------|-------------------------------------------------------------------------------|
| Acetaminophen/Codeine_inter   | Abdominal X Ray                                                               |
| Acetaminophen/Hydrocodone_inter | ECMO Hours                                                                   |
| Acetaminophen_inter           | CT Abdomen Pelvis                                                            |
| Acetazolamide_inter           | CT Brain                                                                      |
| Acyclovir_inter               | CT Chest                                                                      |
| Albumin_inter                 | Central Venous Line Site                                                     |
| Albuterol_inter               | Chest Tube Site                                                              |
| Allopurinol_inter             | Chest X Ray                                                                   |
| Alleplase_inter               | Comfort Response Level                                                        |
| Amikacin_inter                | Continuous EEG Present                                                       |
| Aminophylline_inter           | Diversional Activity                                                         |
| Aminophylline_inter           | Diversional Activity                                                          |
| Amodipine_inter               | Diversional Activity                                                          |
| Amoxicillin/clavulanic acid_inter | Diversional Activity                                                        |
| Amoxicillin_inter             | Diversional Activity                                                          |
| Ampicillin/Sulbactam_inter    | Diversional Activity                                                          |
| Ampicillin_inter              | Diversional Activity                                                          |
| Aspirin_inter                 | Diversional Activity                                                          |
| Atropine_inter                | Diversional Activity                                                          |
| Azathioprine_inter            | Diversional Activity                                                          |
| Azthromycin_inter             | Diversional Activity                                                          |
| Baclofen_inter                | Diversional Activity                                                          |
| Basiliximab_inter             | Diversional Activity                                                          |
| Budesonide_inter              | Diversional Activity                                                          |
| Bumetanide_inter              | Diversional Activity                                                          |
| Calcium Chloride_cont         | Diversional Activity                                                          |
| Calcium Chloride_inter        | Diversional Activity                                                          |
| Calcium Gluconate_inter      | Diversional Activity                                                          |
| Carbamazepine_inter           | Diversional Activity                                                          |
| Cefazolin_inter               | Diversional Activity                                                          |
| Celiprole_inter               | Diversional Activity                                                          |
| Cefotaxime_inter              | Diversional Activity                                                          |
| Cefotixin_inter               | Diversional Activity                                                          |
| Ceftriaxone_inter             | Diversional Activity                                                          |
| Cephalixin_inter              | Diversional Activity                                                          |
| Chloral Hydrate_inter         | Diversional Activity                                                          |
| Chlorothiazide_inter          | Diversional Activity                                                          |
| Ciprolloxac HCL_inter         | Diversional Activity                                                          |
| Cicatrascrubin_cont           | Diversional Activity                                                          |
| Clindamycin_inter             | Diversional Activity                                                          |
| Drugs                         | Drivational Activity                                                          |
| Clonazepam_inter              | ECLS Hours                                                                    |
| Ipratropium Bromide_inter     | ECLS Set Rate                                                                 |
| Ketamine_cont                 | ECLS Activity Level Completed                                                 |
| Ketamine_inter                | ECLS Hours                                                                    |
| Lactobacillus_inter           | ECLS Activity Level Completed                                                 |
| Lansoprazole_inter            | ECLS Hours                                                                    |
| Levatoral_inter               | ECLS Activity Level Completed                                                 |
| Levetiracetam_inter           | ECLS Activity Level Completed                                                 |
| Oxacillin_inter               | ECLS Activity Level Completed                                                 |
| Dexamethasone_inter           | ECLS Activity Level Completed                                                 |
| Hydrocortisone_inter          | ECLS Activity Level Completed                                                 |
| Nitric oxide_inter            | ECLS Activity Level Completed                                                 |
| Nifedipine_inter              | ECLS Activity Level Completed                                                 |
| Nitroprusside_inter           | ECLS Activity Level Completed                                                 |
| Norepinephrine_inter          | ECLS Activity Level Completed                                                 |
| Nystatin_inter                | ECLS Activity Level Completed                                                 |
| Octreotide Acetaecont         | ECLS Activity Level Completed                                                 |
| Olanzapine_inter              | ECLS Activity Level Completed                                                 |
| Ondansetron_inter             | ECLS Activity Level Completed                                                 |
| ECLS Rate                      | ECLS Activity Level Completed                                                 |
| ECLS Set Rate                 | ECLS Activity Level Completed                                                 |
| ECLS Activity Level Completed  | ECLS Activity Level Completed                                                 |
| Nitrofurantoin_inter          | ECLS Activity Level Completed                                                 |
| Oxygen Mode_inter             | ECLS Activity Level Completed                                                 |
| O2 Flow Rate                  | ECLS Activity Level Completed                                                 |
| Oxygen Therapy                | ECLS Activity Level Completed                                                 |
| PEEP                          | ECLS Activity Level Completed                                                 |
| Peak Inspiratory Pressure     | ECLS Activity Level Completed                                                 |
| Peritoneal Dialysis Type      | ECLS Activity Level Completed                                                 |
| Pharmacological Comfort Measures Given | ECLS Activity Level Completed                          |
