EXTERNAL VALIDATION OF A NOVEL DIGITAL SIGNATURE IN CONTINUOUS CARDIORESPIRATORY MONITORING TO DETECT EARLY RESPIRATORY DETERIORATION OF ICU PATIENTS

Rachael A. Callcut¹, MD, MSPH, Yuan Xu², BS, Christina Tsai², BS, Andrea Villaroman², B.Eng., MTM, Anamaria J. Robles², MD, Douglas E Lake³,4, PhD, J Randall Moorman³,4, MD, Xiao Hu⁵, Ph.D., Matthew T Clark, Ph.D.⁶

University of California, Davis, Department of Surgery, Davis, CA, USA¹
University of California, San Francisco, Department of Surgery, San Francisco, CA, USA²
University of Virginia, UVa Center for Advanced Medical Analytics, Charlottesville, VA, USA³
University of Virginia, Cardiovascular Division, Charlottesville, VA, USA⁴
Duke University, School of Nursing⁵
Advanced Medical Predictive Devices, Diagnostics, and Displays, Charlottesville, VA⁶

CORRESPONDING AUTHOR:
Name: J Randall Moorman, MD
Affiliation: University of Virginia
Address: Box 800158, Charlottesville VA 22947
Email: rm3h@virginia.edu

EMAIL ADDRESSES

Rachael A Callcut: Rachael.Callcut@ucsf.edu
Yuan Xu: Yuan.Xu@ucsf.edu
Christina Tsai: Rachael.Callcut@ucsf.edu
Andrea Villaroman: Andrea.Villaroman@ucsf.edu
Anamaria Robles: Anamaria.Robles@ucsf.edu
Doug Lake: DEL2K@hscmail.mcc.virginia.edu
Matthew Clark: mclark@amp3d.biz
J Randall Moorman: rm3h@virginia.edu
Xiao Hu: Xiao.Hu@ucsf.edu

ADDRESS FOR REPRINTS: See corresponding author

RUNNING TITLE: Predicting emergent ICU intubation.
CONFLICTS OF INTEREST: Moorman and Clark are officers and own stock in Advanced Medical Predictive Devices, Diagnostics, and Displays
FINANCIAL DISCLOSURES: Dr. Callcut is supported by a Career Development Award from the NIH Big Data to Knowledge Initiative (NIH K01ES026834).
WORD COUNT: 2669 words

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ABSTRACT

The goal of predictive analytics monitoring is the early detection of patients at high risk of subacute potentially catastrophic illnesses. A good example of a target illness is respiratory failure leading to urgent unplanned intubation, where early detection might lead to interventions that improve patient outcome. Previously, we identified signatures of this illness in the continuous cardiorespiratory monitoring data of Intensive Care Unit patients and devised algorithms to identify patients at rising risk. Here, we externally validated 3 logistic regression models to estimate risk of emergency intubation that were developed in Medical and Surgical ICUs at the University of Virginia. We calculated the model outputs for more than 8000 patients in University of California – San Francisco ICUs, 240 of whom underwent emergency intubation as determined by individual chart review. We found that the AUC of the models exceeded 0.75 in this external population, and that the risk rose appreciably over the 12 hours prior to the event. We conclude that abnormal signatures of respiratory failure in the continuous cardiorespiratory monitoring are a generalizable phenomenon.

INTRODUCTION

Patients in intensive care unit (ICU) that experience respiratory failure leading to emergent intubation have significantly longer hospital length-of-stay and higher in-hospital mortality\(^1\)\(^-\)\(^3\). Earlier identification of patients that will require intubation may allow earlier intervention to reduce morbidity and mortality. Patients in early stages of
respiratory decompensation may be administered corticosteroids or bronchodilators, given supplemental oxygen, or placed on non-invasive positive-pressure ventilation\textsuperscript{4}. Patients for whom non-invasive positive-pressure ventilation is insufficient to prevent respiratory failure, or recently extubated patients that require re-intubation, may be intubated electively rather than emergently\textsuperscript{5,6}. Timely intervention aimed at better preparation for intubation, team coordination, proper intubation medication selection, and avoidance of peri-intubation hypotension have been shown to improve outcome of patients requiring emergent intubation.

Predictive analytics monitoring strives to identify high risk patients earlier by providing continuous risk estimates to clinical personnel in real-time. These early warning signals have potential to allow intervention that may alter the patient trajectory in a more favorable direction. In a prospective study using analogous predictive analytics for sepsis that were available at the bedside of preterm infants, mortality was reduced by more than 20\%. Similarly, in adults, predictive analytics for sepsis may reduce rates of septic shock and associated mortality by up to 50\%\textsuperscript{7}. There are candidate risk marker models for emergent intubation developed at single centers and based on static demographics, comorbidities, injury type, and intermittent vital signs\textsuperscript{3,8}. Models based on continuous cardiorespiratory dynamics from bedside physiological monitors also exist but are not yet validated in an external cohort of all comer ICU patients\textsuperscript{1,2,9}.

Before using a predictive analytical model for prospective clinical practice, it is important to validate that model across differing patient populations and institutions. It is
necessary to know if a model trained at one institution generalizes to provide early
detection of patient deterioration at another institution. Different care units and
institutions may have substantially different distributions of demographics, socio-
economic groups, admitting practices, and care patterns, all of which may degrade the
calibration and discriminatory performance of a model. This study tests the hypothesis
that predictive dynamic analytical models for emergent intubation using continuously
available data from ubiquitous physiologic monitors are well-suited for application at an
external center.
METHODS

Study Design

This is a retrospective cohort study of patients admitted to an ICU at the University of California San Francisco Medical Center (UCSF). We studied admissions to the two mixed medical/surgical ICUs, two neurological ICUs (NICUs), and coronary care (CCU) ICUs. Each ICU has continuous physiological monitoring archived by BedMaster (Excel Medical, Jupiter FL). The primary outcome was respiratory decompensation resulting in emergent intubation.

Study population and primary outcome

We studied consecutive ICU admissions from May 1, 2013 through April 30, 2015 and selected patients in which intubation occurred in the ICU. Intubation events prior to ICU admission (e.g., in the operating room or emergency department) were excluded. Patients with “Do Not Resuscitate” or “Do Not Intubate” orders (DNR/DNI) were excluded. Each intubation event was classified as planned (elective) or unplanned (emergent). Planned intubations included those done for procedures (such as endoscopy, interventional radiology procedures, or preceding elective operations) and those that were documented as elective in the medical record. All other intubations occurring in the ICU were considered unplanned. The procedure note and physician notes were examined to determine the reason for intubation. Timing of intubation was extracted from the medical record nursing/respiratory therapist documentation. The
times were checked for accuracy by two independent practitioners reviewing each potential case of emergent intubation

Identification of mechanically ventilated patients

When evaluating the performance of a model for emergent intubation it is important to exclude patients at times when they are not at risk. Patients are not at risk for emergent intubation at times when they are already mechanically ventilated. In order to exclude data during these times, it is necessary to know the time of all intubations and extubations (not only the emergent intubations). We used the flowsheet entries from respiratory therapists to identify periods of mechanical ventilation and merged the results with known times of emergent intubations.

We extracted the total ventilator respiratory rate flowsheet entry documented by RTs. Figure shows the probability density of time between RT documentation: most RT documentation of total ventilator respiratory rate occurs more frequently than q6-hours. Mechanical ventilation was defined as starting at the time of the first ventilator respiratory rate and ending at the time of the last ventilator respiratory rate. We split the period of mechanical ventilation when the time between consecutive measurements from a patient was larger than 16 hours, and identified the patient as not ventilated in the interim. Isolated measurements (i.e., those without another measurement within 16 hours) were used to define the start of a ventilation epoch with duration of 1 hour. For patients that were emergently intubated (i.e., for whom we know time of intubation based on individual chart review) we ensured that the epochs of mechanical ventilation started at the appropriate time. We split the ventilation epochs when an emergent
intubation occurred in the middle, defining the time of extubation as the time of the preceding ventilator respiratory rate and the time of intubation as the time of emergent intubation.

**Predictors of emergent intubation**

We calculated 3 risk estimates for emergent intubation\(^1,2\). The models were developed on data from SICU and MICU patients at University of Virginia (UVa) health system, and use only continuous cardiorespiratory dynamics calculated from the bedside physiological monitors. From Politano *et al.*\(^2\) we used the vital signs only model. Note that this model was developed on a subset of surgical ICU patients, excluding those on ancillary services. We also used the SICU and MICU models of Moss *et al.*\(^1\). The model of Politano was previously validated in the UVa for predicting upgrade from surgical intermediate care unit (IMU) to ICU, with and without intubation\(^10\). The SICU model of Moss was also validated at the same institution as part of a model for identifying low-risk patients at the time of surgical IMU and ICU discharge\(^11\).

Model inputs are were calculated in 30 minute windows with 50% overlap: means and standard deviations of, and cross-correlations between, vital signs (heart rate, respiratory rate, peripheral oxygen saturation, and blood pressure), as well as statistical measures of cardiac dynamics (slope of log RR interval variance versus log scale for detrended fluctuation analysis, coefficient of sample entropy, and the standard deviation of RR intervals)\(^12-15\). Missing data were imputed with the median values from the UVa development cohort. Features and models were calculated using CoMET® (AMP3D,
Charlottesville, VA). No features or models were available to the care team and all patients received standard of care. This study was approved by the UCSF institutional review board under a waiver of consent.

**Statistical analysis**

We evaluated the performance of the UVa urgent intubation risk models for identifying UCSF patients prior to emergent intubation. We used the continuous risk estimate and a binary response variable. The response was “1” for patients at times prior to emergent intubation (the event population), and “0” for those patients far from intubation and all other patients at all other times. We evaluated the performance of risk estimates for 4 to 24-hour windows prior to emergent intubation. Confidence intervals were determined by 200 bootstrap runs, resampled by hospital admission.
RESULTS

We studied 9,828 admissions for 8,434 patients to UCSF. There were 240 episodes of emergent intubation in 238 hospital admissions. There was continuous monitoring data before 225 (93%) of events.
Table 1 shows characteristics of the study population. Patients with emergent intubation were less likely to be white, more likely to be Asian, had 3-fold higher mortality and stayed 19 days longer in hospital. We calculated 105.5 patient-years of risk estimates: the 3 risk estimates were calculated for 3.7 million 15-minute epochs. We censored 32.3% of measurements that were for patient that were already intubated or for patients following DNI orders. The incidence of emergent intubation was 0.9 per 100 ICU days, significantly lower than in the Moss\textsuperscript{1} and Politano\textsuperscript{2} development cohorts (2.1 and 2.8 per 100 ICU days respectively).

Figure 2 (left) shows the number of events with continuous monitoring data leading up to the time of emergent intubation. Figure 2 (right) shows the time course of risk estimates for UCSF patients using the three UVa models for urgent unplanned intubation during the 48 hours preceding emergent intubation. Risk estimates doubled over the 12 to 24 hours prior to emergent intubation, from about 1.5 to more than 3. At each time we performed a signed rank test with the null hypothesis that risk estimates are equal to risk estimates from the same patient 12 hours prior, and a two-sided alternative hypothesis. White points identify times when we rejected the null hypothesis. Risk estimates were significantly higher beginning about 5 hours prior to emergent intubation.

Figure 3 shows the performance of the 3 models for emergent intubation as a function of the window before intubation that is considered the event. For a 24-hour detection window, for example, all data within 24 hours prior to emergent intubation were
identified as ‘event’ and all other data (from event patients far from the event, and from all non-event patients) was identified as control. The AUC rose from about 0.75 to about 0.78 for event windows 24 to 4 hours, respectively. The MICU model from Moss\textsuperscript{1} had slightly better performance than the SICU models from Moss\textsuperscript{2} and Politano\textsuperscript{2}.

Figure 4 shows the calibration of the 3 models using a 12-hour detection window. Well calibrated models have predicted risk equal to observed risk (dashed line). The SICU models from Moss\textsuperscript{1} and Politano\textsuperscript{2} exhibited excellent calibration in all but the highest risk patients, where they overestimated the risk. The MICU intubation model from Moss\textsuperscript{1} also overestimated risk for high risk patients and also deviated from the line of identity in the 2 lowest deciles of predicted risk.
DISCUSSION

This study evaluated the performance of 3 predictive analytical models for early detection of respiratory failure leading to emergent intubation at an external center. These models are based upon converting continuous physiologic monitoring data into a novel digital signature that can be easily displayed to the bedside practitioner. The existing models were originally developed to predict emergent intubation in the UVa SICU (2 models) and MICU (1 model), and this study attempted to externally validate the performance using an independent cohort of patients from all ICUs at the University of California San Francisco. The major findings were that the models accurately identified patients at risk for emergent intubation (AUC > 0.75 starting at 12 hours pre-intubation), risk estimates continued to rise up to the time of emergent intubation, and these predictive models were well-calibrated.

Whereas elective intubation in the operating room is an extremely safe procedure, emergent intubation occurring out of the operating room has been shown to be a significantly vulnerable time for patients with a high association of complications\textsuperscript{16-18}. These complications include hemodynamic compromise or severe hypoxemia leading to cardiac arrest, esophageal intubation, aspiration, and pneumothorax\textsuperscript{17, 20-21}. Cardiac arrest attributable to intubation complications occur between 2-4\% of all emergent intubations with few of these patients surviving to hospital discharge even when initially successfully resuscitated\textsuperscript{18-20}. Identifying patients earlier who are at risk of needing emergent intubation may be key in avoiding these poor outcomes\textsuperscript{18,21}.
Recent studies have linked poor outcome to a number of potentially mitigatable factors that are time-dependent. These include a worse outcome for patients intubated around change of nursing shift\textsuperscript{17}. Wardi et al. hypothesized this finding to be a result of potential staff fatigue, hand-off errors, and lack of familiarity with a patient which combined lead to a tendency to overlook subtle but important changes in patient condition\textsuperscript{17}. In each of these potential explanations, predictive algorithms such as those validated in this current study has theoretic potential to overcome these challenges at vulnerable moments for our critically ill patients. Essentially these algorithms convert complex data that is difficult for a clinician to assimilate in real-time into a patient specific digital signature. This signature serving as an early warning signal displayed at the bedside could aid in drawing attention to key changes in patient condition.

Additionally, our data support that for those patients who ultimately require an emergent intubation, the decline in their physiologic state is often a slowly progressive process over many hours extending out beyond the preceding 12-24 hours. This time frame creates an opportunity to perform interventions that would make emergent intubation safer, and even possibly avoided. Prior research has shown a pre-intubation checklist integrating interventions to treat hemodynamic instability led to fewer complications\textsuperscript{17,18,22}. This is consistent with investigations that have correlated the worse outcomes in emergent intubation in the Emergency Department (ED) occurring in those that develop hemodynamic compromise following intubation\textsuperscript{20,23}.

Commonly patients requiring emergent intubation are in shock with an increased shock index highly correlated with intubation associated cardiac arrest\textsuperscript{17,23}. Volume depletion,
vasodilation, acidosis, and reduced venous return resulting from increased positive pressure ventilation are all factors that contribute. Each of these could be addressed, albeit to differing degrees of success, in the lead time provided by an early warning signal. Trivedi et al. argue that the shock index which is calculating from a ratio of maximum heart rate to lowest systolic blood pressure is a helpful adjunct to make intervention in the immediate 60 minutes preceding intubation\textsuperscript{23}. However, they acknowledge that addressing the ‘dynamic changes in patient status in the ICU…require continuous monitoring and interpretation of data before the development of overt hypotension and cardiorespiratory collapse\textsuperscript{23}.’ Our study indicates that implementation and integration of digital signatures into clinical practice may provide even greater opportunity to induce timely clinical action as prediction of need for emergent intubation can be extended out to more than 12 hours pre-intubation. We can not definitively determine from this study if interventions during this window would avoid intubation, but it would likely make them safer. Both of these warrant further investigation in a prospective study.

This study was limited in several other ways. We relied on surrogate data to initially identify patients who had mechanical ventilation initiated during their ICU stay. It is possible that due to documentation errors, the location of intubation initiation could have misclassified. However, to avoid this, each potential case of ICU intubation was reviewed by two clinicians to verify location of intubation, reason for intubation, and timing of intubation. Similarly, we attempted to identify tracheostomy patients that may
go on and off mechanical ventilation and excluded those with mechanical ventilation initiation following tracheostomy.

Determining the exact timing of emergent intubation was also challenging in this study. Physician documentation through intubation procedure notes were not a reliable source of the timing as it was apparent in chart review that this documentation is often done in a delayed manner. This is understandable as the physician attention around the event is focused on providing bedside care. To address this, the current study utilized the first nursing documentation of intubation medications being administered cross referenced with the respiratory therapist documentation of initial ventilator settings to confirm intubation timing. In all cases, these times were within 5 minutes of each other.

In addition, we did not quantify the performance of predictive models in the context of known risk factors: diagnoses, demographics, or severity of illness. We note, however, that these known risk factors are static indicators; though they may identify high risk patients, they do not rise leading up to the time of lung failure. Finally, we did not evaluate other predictors of emergent intubation such as shock index from vital signs and laboratory measurements. Adding independent streams of information improves performance, though we found that models that are plug-and-play with bedside physiological monitors have excellent performance.
CONCLUSION

By leveraging the growth in digital technology paired with graphical visualization, early warning signals displayed at the bedside could aid in provider ability to identify subtle, but key changes in patient condition. Earlier identification of subtle changes in physiologic deterioration may provide a useful adjunct to clinicians to mitigate need for emergent intubation.
REFERENCES


AUTHOR CONTRIBUTIONS:
Study Design: Callcut, Clark, Moorman, Hu
Data Analysis: Callcut, Xu, Clark, Lake, Moorman, Hu
Data Interpretation: Callcut, Xu, Tsai, Villaroman, Robles, Lake, Clark, Moorman, Hu
Writing: Callcut, Clark
Critical Review: Callcut, Xu, Tsai, Villaroman, Robles, Lake, Clark, Moorman, Hu
FIGURE LEGENDS:

Figure 1: Probability density of the time between consecutive ventilator respiratory rate entries for all patients.

Figure 2: (left) Number of events with continuous monitoring data as a function of time leading up to emergent intubation. (right) Average time course of risk estimates over the 48 hours leading up to the time of emergent intubation. White points indicate that the risk estimates at that time are significantly higher (p < 0.05) than risk estimates 12 hours prior.

Figure 3: Area under the receiver operating characteristic as a function of the window size before emergent intubation defined as the event, from 4 to 24 hours. The 95% confidence interval is indicated by errorbars and was determined by 200 bootstrap runs resampled by admission.

Figure 4: Calibration curves for the 3 models for emergent intubation. The observed relative risk is shown as a function of the predicted risk. Each point represents 10% of the data, and the line of identity (perfect calibration) is shown as a dashed line.
### Table 1: Characteristics of the study population

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Event</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Count</strong></td>
<td>9828</td>
<td>240 (2.4%)</td>
<td>0.5296</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>4583 (46.6%)</td>
<td>107 (44.6%)</td>
<td>0.7583</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>61.0 (48.0 - 70.0)</td>
<td>61.5 (51.8 - 68.2)</td>
<td>0.7583</td>
</tr>
<tr>
<td><strong>Days hospital stay</strong></td>
<td>7.0 (4.0 - 12.0)</td>
<td>26.0 (15.0 - 41.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Mortality</strong></td>
<td>1394 (14.2%)</td>
<td>105 (43.8%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>5644 (57.4%)</td>
<td>115 (47.9%)</td>
<td>0.0026</td>
</tr>
<tr>
<td>Black</td>
<td>782 (8.0%)</td>
<td>24 (10.0%)</td>
<td>0.2364</td>
</tr>
<tr>
<td>Asian</td>
<td>1387 (14.1%)</td>
<td>45 (18.8%)</td>
<td>0.0367</td>
</tr>
<tr>
<td>Other</td>
<td>2015 (20.5%)</td>
<td>56 (23.3%)</td>
<td>0.2715</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>1351 (13.7%)</td>
<td>32 (13.3%)</td>
<td>0.8507</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>8104 (82.5%)</td>
<td>197 (82.1%)</td>
<td>0.8771</td>
</tr>
<tr>
<td>Unknown</td>
<td>373 (3.8%)</td>
<td>11 (4.6%)</td>
<td>0.5177</td>
</tr>
</tbody>
</table>

*Values are shown as mean (standard deviation) or count (percentile)*
Figure 1: Probability density of the time between consecutive ventilator respiratory rate entries for all patients.
Figure 2: (left) Number of events with continuous monitoring data as a function of time leading up to emergent intubation. (right) Average time course of risk estimates over the 48 hours leading up to the time of emergent intubation. White points indicate that the risk estimates at that time are significantly higher ($p < 0.05$) than risk estimates 12 hours prior.
Figure 3: Area under the receiver operating characteristic as a function of the window size before emergent intubation defined as the event, from 4 to 24 hours. The 95% confidence interval is indicated by errorbars and was determined by 200 bootstrap runs resampled by admission.
Figure 4: Calibration curves for the 3 models for emergent intubation. The observed relative risk is shown as a function of the predicted risk. Each point represents 10% of the data, and the line of identity (perfect calibration) is shown as a dashed line.