

# Dynamic and Personalized Risk Forecast in Step-Down Units

## Implications for Monitoring Paradigms

Lujie Chen<sup>1</sup>, Olufunmilayo Ogundele<sup>2</sup>, Gilles Clermont<sup>3</sup>, Marilyn Hravnak<sup>4</sup>, Michael R. Pinsky<sup>3</sup>, and Artur W. Dubrawski<sup>1</sup>

<sup>1</sup>Auton Laboratory, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania; <sup>4</sup>Department of Acute and Tertiary Care, University of Pittsburgh School of Nursing, Pittsburgh, Pennsylvania; <sup>3</sup>Department of Critical Care Medicine, School of Medicine, University of Pittsburgh, Pennsylvania; and <sup>2</sup>LifeBridge Critical Care, Sinai Hospital Baltimore, Baltimore, Maryland

ORCID ID: 0000-0002-7185-8405 (L.C.).

### Abstract

**Rationale:** Cardiorespiratory insufficiency (CRI) is a term applied to the manifestations of loss of normal cardiorespiratory reserve and portends a bad outcome. CRI occurs commonly in hospitalized patients, but its risk escalation patterns are unexplored.

**Objectives:** To describe the dynamic and personal character of CRI risk evolution observed through continuous vital sign monitoring of individual step-down unit patients.

**Methods:** Using a machine learning model, we estimated risk trends for CRI (defined as exceedance of vital sign stability thresholds) for each of 1,971 admissions (1,880 unique patients) to a 24-bed adult surgical trauma step-down unit at an urban teaching hospital in Pittsburgh, Pennsylvania using continuously recorded vital signs from standard bedside monitors. We compared and contrasted risk trends during initial 4-hour periods after step-down unit admission, and again during the 4 hours immediately before the CRI event, between cases (ever had a CRI) and control subjects (never had a CRI). We further explored heterogeneity of risk escalation patterns during the 4 hours before CRI among cases, comparing personalized to nonpersonalized risk.

**Measurements and Main Results:** Estimated risk was significantly higher for cases (918) than control subjects (1,053;  $P \leq 0.001$ ) during the initial 4-hour stable periods. Among cases, the aggregated nonpersonalized risk trend increased 2 hours before the CRI, whereas the personalized risk trend became significantly different from control subjects 90 minutes ahead. We further discovered several unique phenotypes of risk escalation patterns among cases for nonpersonalized (14.6% persistently high risk, 18.6% early onset, 66.8% late onset) and personalized risk (7.7% persistently high risk, 8.9% early onset, 83.4% late onset).

**Conclusions:** Insights from this proof-of-concept analysis may guide design of dynamic and personalized monitoring systems that predict CRI, taking into account the triage and real-time monitoring utility of vital signs. These monitoring systems may prove useful in the dynamic allocation of technological and clinical personnel resources in acute care hospitals.

**Keywords:** instability; finite mixture model; machine learning; early warning score; physiologic monitoring

(Received in original form November 18, 2016; accepted in final form December 28, 2016)

Supported by National Institutes of Health grants NR13912 and HL122478, and National Science Foundation grant 1320347.

**Author Contributions:** L.C. conducted computerized experiments, analyzed and interpreted the data, and wrote the manuscript; L.C. and A.W.D. ideated machine learning models used; O.O., G.C., M.H., M.R.P., and A.W.D. contributed to the concept and design of the study, collected the primary data, interpreted the data, and revised the manuscript for intellectual content.

Correspondence and requests for reprints should be addressed to Michael R. Pinsky, M.D., 606 Scaife Hall, 3550 Terrace Street, Pittsburgh, PA 15261. E-mail: pinskymr@upmc.edu

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org)

Ann Am Thorac Soc Vol 14, No 3, pp 384–391, Mar 2017

Copyright © 2017 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201611-905OC

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

Cardiorespiratory insufficiency (CRI) is the external manifestation of loss of normal cardiorespiratory reserve and portends a bad outcome. It is manifest in a variety of

ways, depending on the underlying disease (e.g., chronic obstructive lung disease, trauma, heart failure, pneumonia), associated treatments (e.g., surgery,

pain and sedation medication, anti-hypertensive medication, supplemental oxygen), and acute associated processes (e.g., bronchospasm, airway occlusion,

retained secretions, decreased mental status, recurrent blood loss). Missed or delayed recognition of CRI among hospitalized patients often leads to missed opportunities for early intervention and compromised quality of care (1).

Step-down units represent a unique subset of hospitalized patients with elevated propensity for CRI, combined with continuous and intermittent noninvasive vital sign monitoring. In common practice, however, the experience level of bedside clinicians (primarily nurses) plays a major role for early CRI recognition, which is an increasing challenge, given common shortages in personnel (2). We hypothesized that the large amounts of data collected from bedside monitors could provide both a useful view of a patient's current state of risk for CRI and its possible future evolution.

Various existing risk-scoring systems use mortality as the primary targeted outcome. The Acute Physiology and Chronic Health Evaluation IV (3) and the Simplified Acute Physiology Score II (4) predict intensive care unit mortality, whereas the Rothman Index (5) predicts mortality in ward patients. We believe that CRI risk prediction would greatly augment the early-warning alert landscape, especially if such alerts could be generated far upstream from a mortality outcome. Resulting models, meant to aid in CRI diagnosis, could potentially enable earlier diagnosis and the application of preemptive interventions to limit or even prevent overt CRI, minimizing end-organ injury and its subsequent morbidity. A system that could predict unstable vital sign states preceding CRI using continuous vital sign monitoring would introduce a novel risk-monitoring paradigm, by helping identify not only patients who are at increased CRI risk, but also when the CRI will occur in the future.

As a first step in exploring this paradigm, we hypothesized that many patients who eventually develop CRI show a discernible elevated risk, even at the time of admission, and that the status of those who develop CRI evolves along various heterogeneous trajectories of risk over time that differ from both those stable patients who never experience CRI, and from other patients whose CRI trajectories are different. To accomplish this, we leveraged machine learning techniques to construct a robust CRI detection, forecasting, and characterization capability for step-down

unit patients, and to prove the concept using a large annotated vital sign database.

We conducted this proof-of-concept study in several sequential stages, as explained subsequently here. First, we describe the methods to derive an instantaneous and continuously evolving CRI risk score for tracking patients' cardiorespiratory state over time. The results of this analysis reflect a relative risk (RRisk) of developing CRI as compared with patients who will never develop CRI. This is done without making any assumptions on heterogeneity of personal baseline values. Then we repeat the analysis using the patients' own initial vital sign values to define their baseline states, from which their CRI risk evolves, referred to as a personalized RRisk (PRRisk) estimate. Second, we analyze risk trends among cases (patients who will develop CRI) and control subjects (patients who never develop CRI) in the initial period of their step-down unit stay, as well as during periods immediately preceding CRI events, to identify any potential patient-specific trajectories. Finally, we group these trajectories into classes of equivalence to identify a manageably small yet diverse set of heterogeneous phenotypes of risk escalation among cases using a group-based modeling approach.

## Methods

The method used in this work is an extension of the modeling framework that we previously introduced in 2015 based on classic machine learning approaches (6). In that study, we extracted many numeric features from time series vital sign data (e.g., heart rate, respiratory rate, blood pressure, and pulse oximeter O<sub>2</sub> saturation). We then used the resulting large dynamic feature dataset to estimate risk scores using supervised machine learning techniques by comparing cases with control subjects from a small cohort of thoroughly characterized step-down unit patients. We then aimed to discover phenotypes of risk escalation trajectories among the patients who would subsequently develop CRI.

Although insightful, that previous analysis lacked personalized parameters, because we considered all patients to live within a common range of normal and abnormal vital sign ranges. In this study, we

take a further step to define the personalized risk measure to add to the nonpersonalized version of it and compare, contrast, and analyze the risk trajectories for these two types of risk measures (e.g., RRisk and PRRisk) to shed new insight on the step-down unit patient CRI risk evolution.

In addition, we also create a triage view of potential future CRI using only the first 4 hours of data since step-down unit admission to ascertain if an initial characterization of CRI risk upon step-down unit admission can viably separate cases from control subjects, and then determine with which phenotypical risk group each case is most likely aligned. Our previous work limited the data view to 30 minutes preceding the onset of CRI. This work marks a major extension of our predictive analytical approach.

## Data

Noninvasively recorded vital signs from the bedside monitors of a 24-bed adult surgical trauma step-down unit at a large urban teaching hospital in Pittsburgh, Pennsylvania were collected during 15 months (May 2007 to September 2008) in an institutional review board-approved protocol. Data included continuous recordings of heart rate, respiratory rate (bioimpedance signaling), and peripheral oxygen saturation as measured by pulse oximetry (SpO<sub>2</sub>) at 1/20 Hz, and systolic and diastolic blood pressure recorded at least every 2 hours.

CRI events were flagged whenever any of the vital signs exceeded a predefined threshold consistent with our Medical Emergency Team calling criteria (heart rate < 40 or > 140 min<sup>-1</sup>, respiratory rate < 8 or > 36 min<sup>-1</sup>, systolic blood pressure < 80 or > 200 mm Hg, diastolic blood pressure > 110 mm Hg, SpO<sub>2</sub> < 85% [7]) and persisting for at least 80% of the time over a 3-minute window. Using the method described in our previous work (8, 9), we identified a subset of clinically important CRI events to be used as the predictive endpoint in this study.

We used the first 3 months of data to build the CRI risk model using 15 minutes of vital sign measurements immediately preceding CRI, as previously described (6). We then applied this model to moving time windows during the periods before the CRI events on complete vital sign histories from patients in the remaining 12 months of data

(the validation set) to compute patient-specific CRI risk trajectories.

### Risk Model

Computed features informing the CRI risk model include various statistical parameters of vital sign time series and signal entropy metrics (10, 11) (see Table E1 in the online supplement). This feature set was replicated for time windows of 5, 10, and 15 minutes, each ending at the same time stamp of interest, resulting in 126 features, to see if shorter time series durations retained their predictive power. We used it to train a variant of the random forest classification model (12) using nonrandom splits.

### Risk Trajectories

We featurized the raw data at every minute for all patients in the validation set and applied the trained CRI risk model to compute time series of risk scores (Figure 1).

We extracted two types of risk subtrajectories. The first type reflected the first 4 hours of each patient's step-down unit stay ("stable period") and the second reflected the 4 hours immediately preceding onset of the leading CRI event ("event period") for each case patient. For the second subtrajectory type among control subjects, the event period was chosen such that the distribution of hypothetical onset times roughly matched distribution of the actual CRI case onset times within their step-down unit stays. We excluded 4% of cases whose CRI events occurred within 4 hours of admission.

We defined absolute risk as the raw prediction score from the random forest classification model: a number between

0 (low risk, similar to control subjects) and 1 (high risk, similar to cases at CRI onset). We further defined a nonpersonalized or RRisk as the ratio of the absolute risk and the mean absolute risk for control subjects computed over the 4-hour stable period. RRisk expresses the particular patient's risk normalized against an average control subject's risk level. It follows that control patients' RRisk is approximately 1.0 on average. In addition, we defined each patient's PRRisk as the nonpersonalized RRisk risk normalized by each patient's own RRisk computed during their first 4 hours in the step-down unit.

The RRisk scores reflect generic risk relative to population-based vital sign norms, whereas the PRRisk accounts for the individual vital sign ranges when defining subsequent CRI events. By tracing these risk scores over time, we obtained risk trends for each patient and analyzed them first according to whether they belonged to cases or control subjects. Finally, we summarized trends of case patients using group-based modeling (13) to observe and reflect heterogeneity of risk evolution trajectories.

## Results

### Patient Demographics

The training set had 322 CRI events (cases) from 138 step-down unit admissions (134 patients) and 370 control windows from 68 patients who never experienced CRI, randomly selected from the first 6 hours of their step-down unit stay. The validation set included 918 CRI events (cases) from 918 step-down unit admissions (863 patients)

and 1,053 control subjects from 1,053 step-down unit admissions (1,017 patients). Table 1 details the patients' demographics.

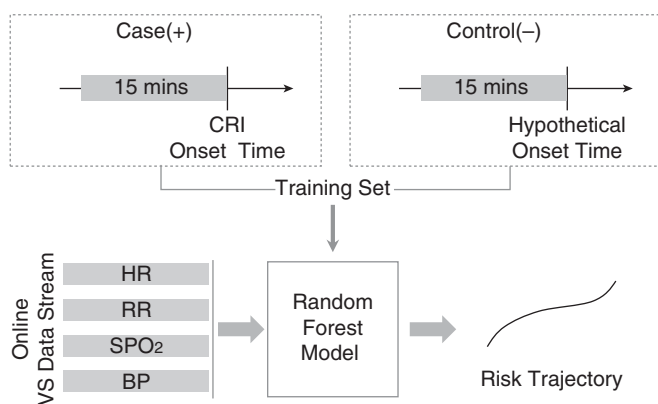
### Discrimination of CRI

The random forest classification model learned from the training set discriminated between case and control events at the mean CRI onset time with a 10-fold cross validation area under the receiver operating characteristic curve (AUC) of 0.94. For comparison, we also fitted logistic regression (LR) and regularized Lasso LR models (14). However, the 10-fold cross-validation AUC scores obtained with LR (0.7) and Lasso LR (0.82) were inferior to the random forest classification model, suggesting that this model is more suitable in this context, possibly due to its ability to reflect complex relationships inherent in these data. When applying the trained model to the validation set at various time points before CRI onset in cases, the random forest classification model AUC initially remained constant, followed by an increasing trend, with AUCs rising from 0.57 to 0.89 during the 4 hours immediately preceding events, whereas, during the 4-hour stable period of these same cases, the AUCs were consistently low (at a level of 0.58–0.60, 95% CI; Figure E1).

### Risk Analysis

Figure 2 displays the nonpersonalized RRisk subtrajectories aggregated separately for cases and control subjects during the 4-hour stable period and the 4 hours immediately preceding the event period. During the initial 4-hour stable period after admission, the RRisk for control subjects was close to 1.0 by design, whereas, for cases, the RRisk was 1.18 ( $\pm 0.02$ ), suggesting that even shortly after step-down unit admission, the average risk is slightly, but significantly, higher for cases than for control subjects (shown in the *P* value panels of Figure 2). As time progresses toward CRI, the risk for the cases begins increasing around 2 hours before the CRI event, whereas, for control patients, it remains unchanged.

Figure 3 shows a personalized view of risk evolution. After normalization of each patient's risk to their own baseline computed within the first 4 hours of step-down unit admission, the PRRisk was similar and close to 1.0 during the early period for both cases and control subjects (left panel). The PRRisk for cases starts to rise around 2 hours before the event



**Figure 1.** Workflow for estimating individual risk trajectories. BP = blood pressure; CRI = cardiorespiratory insufficiency; HR = heart rate; RR = respiratory rate; SPO<sub>2</sub> = pulse oximetry; VS = vital signs.

Table 1. Patient demographics

Characteristics	Training Set			Validation Set		
	Cases	Control Subjects	P Values	Cases	Control Subjects	P Values
Step-down unit admissions, n	138	68		918	1,053	
Patients, n	134	68		863	1,017	
Events, n	322	370		918	1,053	
Male %	57	61	0.371	58%	60%	0.299
Race, white, %	74	66	0.065	73%	73%	0.444
Age, yr, mean $\pm$ SD	59.24 $\pm$ 18.20	54.11 $\pm$ 20.13	<0.001	61.69 $\pm$ 18.81	57.90 $\pm$ 19.96	<0.001
Charlson Deyo index, median (IQR)	0 (0–1)	0 (0–1)	0.166	1 (0–3)	1 (0–2)	<0.001
Step-down unit length of stay, d, median (IQR)	7 (4–9)	3 (2–4)	<0.001	4 (2–6)	2 (1–4)	<0.001
Hospital of stay, d, median (IQR)	11 (6–17)	4 (2–9)	<0.001	8 (4–14)	4 (2–9)	<0.001
CRI events vital types, %						
Heart rate	10			8		
Respiratory rate	37			27		
Peripheral Sp <sub>O</sub> <sub>2</sub>	43			48		
Blood pressure	10			17		

Definition of abbreviations: IQR = interquartile range; CRI = cardiorespiratory insufficiency; Sp<sub>O</sub><sub>2</sub> = oxygen saturation as measured by pulse oximetry.

(right panel), but the difference is not significant until about 90 minutes before the event. Of note, an upward drift in PRRisk was observed in both cohorts as the event period approached, owing to less risky control subjects having been discharged from the step-down unit at that point, such that the residual control cohort was comprised of relatively sicker patients with slightly higher risk trends.

### Group-Based Modeling Results

Next, we further stratified CRI cases according to heterogeneous risk escalation patterns during the event period using the group-based modeling method (13). This was performed independently for RRisk (Figure 4, left) and PRRisk (Figure 4, right). Three phenotypical groups (group 1 [red], group 2 [blue], and group 3 [green]), accounting for 66.8% of the cohort, exhibited a late risk escalation pattern, though each with slightly different risk levels initially (group 1, low; group 2, medium; group 3, high). Interestingly, the initial risk level for group 1 was even lower than for control subjects, and this subset would have been considered as low risk until very close to the event. Group 4 (purple) patients, accounting for 18.6% of cases, exhibited a gradual and close-to-linear-with-time escalation pattern, which started with very low risk and gradually increased until the CRI event period was reached. For this phenotype, the earliest sign of an upward trend occurred approximately 3 hours before the event.

Group 5 (orange), accounting for 14.6% of cases, characterized a persistently high-risk type well before the event, and could have evolved as early as during the stable period.

Figure 4 (right) shows risk escalation phenotypical patterns estimated from PRRisk as compared with RRisk. Prevalence of patterns varies between these different views, suggesting that some trajectories, which ascribed to a particular pattern in the nonpersonalized risk view (RRisk), migrated to a different pattern in their personalized risk view (PRRisk).

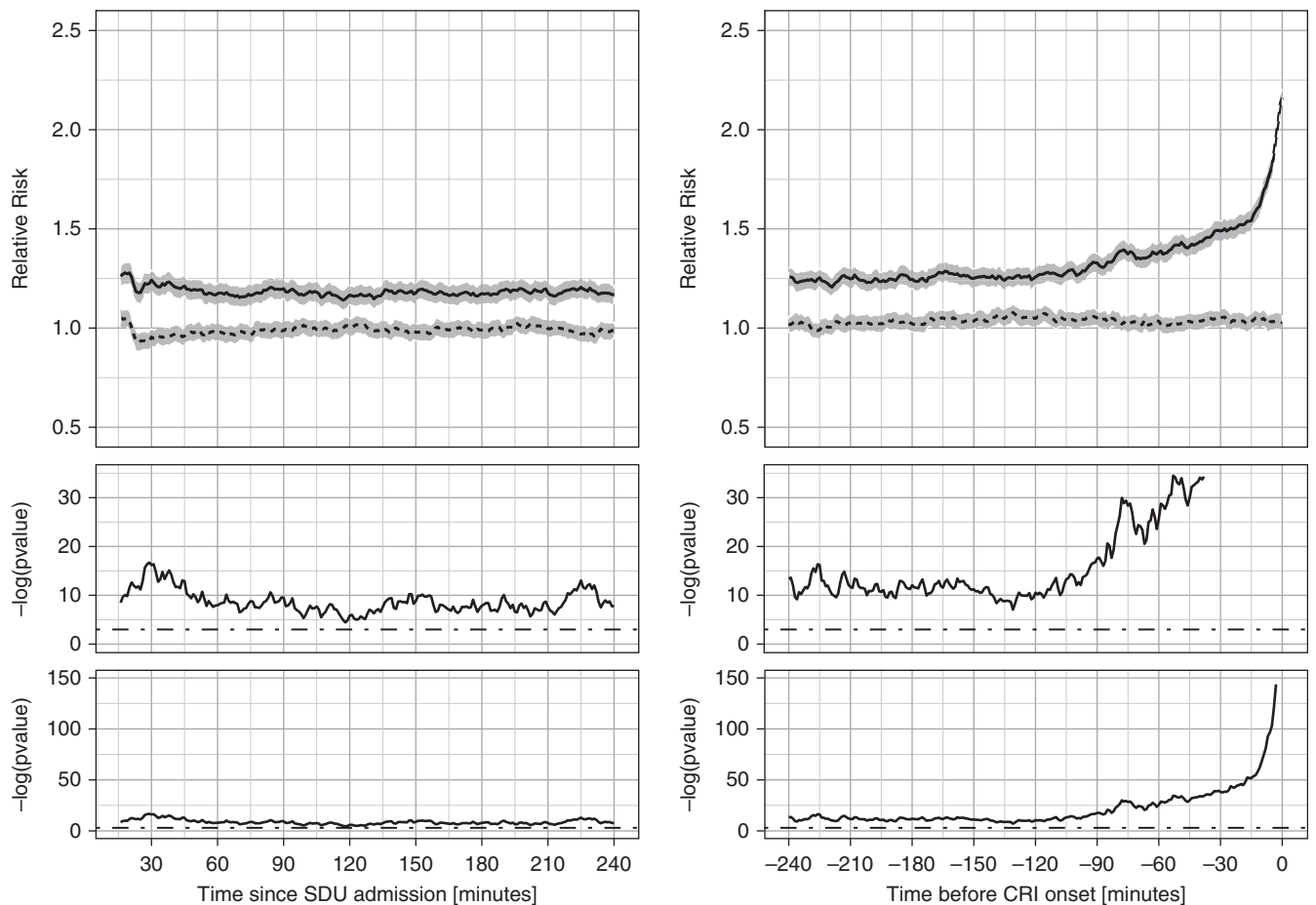
The matrix in Table 2 summarizes these migrations. For ease of interpretation, we combined the three late-onset groups into one combined group named “late onset.” The off-diagonal counts reflect trajectories that changed their group assignment between the RRisk and PRRisk views. The most salient migration pattern is that 110 (89%) of the persistently high-risk patients from the RRisk view move to the late-onset type in the PRRisk view, which isolates a subgroup of patients whose risk level was high even at the time of admission, but whose personalized risk was not elevated until soon before CRI. This shows a prognostic utility unique to the personalized risk scores.

### Discussion

Different from the commonly used mortality risk-based scoring system, our interest is in both diagnosis (nowcasting)

and prognosis (forecasting) of clinically important CRI events to serve as an aid in triage and real-time monitoring for CRI. We created an analytic framework to characterize evolution of CRI risk in step-down unit patients based on risk trends estimated from moving time windows of continuously streaming vital sign data. This framework uses established supervised machine learning tools and simple statistical featurization of the streaming multiparameter vital sign data. Our results provide proof of concept that a CRI diagnostic aid based on current and future risk of CRI is possible. We validated performance of this approach on a substantial cohort of step-down unit patients in a large tertiary care center. We derived several insights from these data that have important implications in support of assessing personalized continuous instability risk monitoring in acutely ill patients, both at the time of step-down unit admission and in a prognostic setting, updated dynamically throughout their step-down unit stay.

First, we observed that the continuous vital sign recordings from the initial stable period immediately after step-down unit admission contain useful information to discriminate patients susceptible to future vital sign instability from patients who will never become unstable, answering the question of *who* will likely become unstable. Interestingly, as shown in the left panels of Figures 2 and 3, the first 30 minutes after step-down unit admission do not appear to provide a stable enough signal for our



**Figure 2.** Relative risks. Relative cardiorespiratory insufficiency (CRI) risk subtrajectories aggregated separately for case (*solid lines*) and control (*dotted lines*; *gray areas* represent 95% confidence intervals) patients for the 4-hour stable period (*left*) and for the 4 hours immediately before CRI onset (*right*). Time index 0 corresponds to time of step-down unit (SDU) admission (*left*) or CRI onset time (*right*). In the *bottom plots*, we show the *P* value series computed from two-sample *t* tests at each time point shown on the logarithmic scale. The *middle plots* are zoomed-in versions of the *bottom plots*. The cutoff threshold of 5% significance is shown with *dot-dash straight lines*.

analyses. This may not be a surprise clinically, as patients typically require some time to settle after being transferred to the step-down unit and receiving initial admission care. Regardless, the triaging utility of this type of early baseline provides the first layer of risk stratification in patient monitoring, as it could potentially guide initial allocation of monitoring resources, as well as optimization of nurse-to-patient ratios, levels of nurse expertise required for individual patients, and frequency of needed personal encounters.

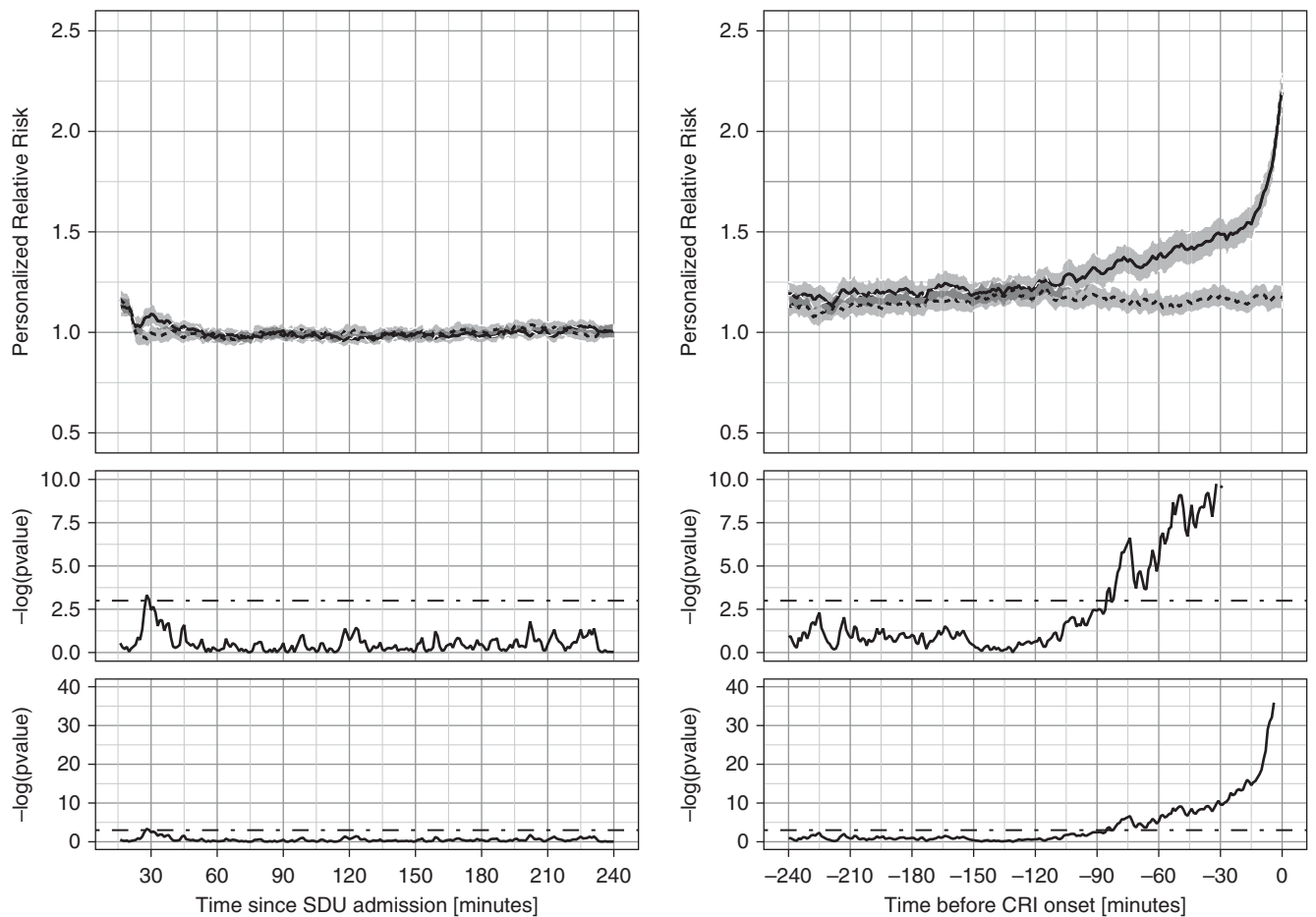
Second, we observed that, for the majority of patients, CRI development can be forecast by new trends in risk evolution during the hours immediately preceding CRI, answering the question of *when* they will become unstable. Referring to the event period plots in Figures 2 and 3 (*right*

*panels*), there were obvious upward risk trends for cases at an aggregated level. This was true from both the RRisk and PRRisk perspectives, suggesting that advance notice of personally forecasted CRI based on risk trends observed for an individual is possible. This constitutes the basis for the second layer of stratification that is adaptive to the patients real-time risk evolution.

Third, we observed that the group-based risk trajectory analysis provides further details of risk escalation patterns as patients progress toward CRI. By stratifying the individual trajectories into multiple groups, each with a distinct phenotypic escalation pattern, we confirmed our hypothesis that risk progression toward CRI is not only dynamic, but also heterogeneous. In a sense, this simultaneously addresses both the questions of *who* and *when* at a

fine-grained level of vital sign data analysis. We envision that a CRI risk notification system based upon this insight could stratify patients further into unique phenotypes for more specific and accurate CRI forecasting than what is currently available in practice. For example, patients who exhibit persistently high risk of CRI may require different monitoring resources and treatments than those who do not demonstrate escalation until just before onset. In addition, even the patients who rapidly and unexpectedly progress toward CRI may benefit from preemptive interventions applied earlier, which is usually not possible in our current clinical model of reactive practice.

Fourth, by comparing the heterogeneous trajectories of RRisk and PRRisk (Table 2), we were able to



**Figure 3.** Personalized relative risks (PRRisk). PRRisk subtrajectories aggregated separately for case (solid lines) and control (dotted lines; gray areas represent 95% confidence intervals) patients for the 4-hour stable period (left) and for the 4 hours immediately before cardiorespiratory insufficiency (CRI) onset (right). Time index 0 corresponds to time of step-down unit (SDU) admission (left) or CRI onset time (right). In the bottom plots, we show the *P* value series computed from two-sample *t* tests at each time point shown on the logarithmic scale. The middle plots are zoomed-in versions of the bottom plots. The cutoff threshold of 5% significance is shown with dot-dash straight lines.

isolate a group of patients who were at a relatively high risk even during the initial part of their stay, and remained in a “stable high-risk” state most of the time, and only escalated their personalized risk very close to the event period. This implies that these individuals at an initial high-risk level deserve special attention, or transfer to an environment with higher-intensity monitoring and caregiver density, such as an intensive care unit, because their decompensation would likely proceed quickly.

Fifth, the above insights from this study provide guidance on building a truly proactive and personalized CRI risk alert system that could produce a more preemptive action than is available from the existing early warning systems for mortality. One important difference is its potential to predict at various lead times, while reflecting

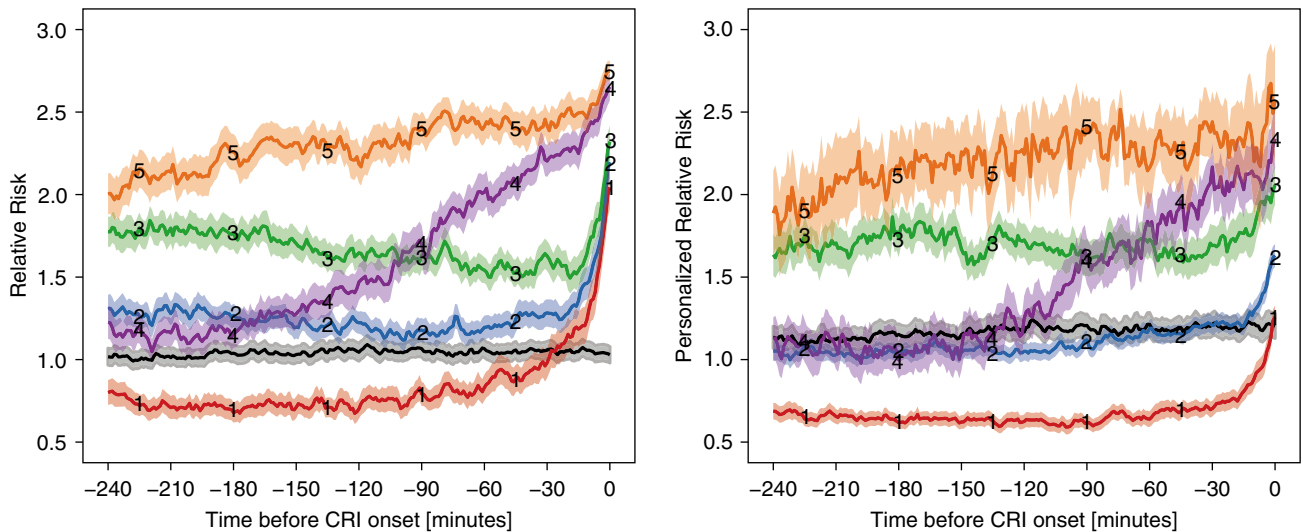
a heterogeneous character of the risk escalation trajectories. It is possible that the heterogeneity of these phenotypal patterns may be due to the underlying differences in the pathophysiologic causes leading to the unstable state.

Although patients in specific groups tend to have slightly different mean ages, their admission diagnoses did not appear to be similarly grouped. Thus, simple explanations for patient grouping based on preadmission data do not seem feasible. This point warrants further analysis, as it could lead to pathophysiology-based phenotypes of risk evolution at a more protean level. Another important difference between the system we propose and the existing early warning systems which predict risk of mortality, is that our approach predicts occurrence

of instability far upstream of possible death after instability. Thus, such a system could be amenable to preemptive, as well as proactive, approaches and treatments.

### Limitations

Despite having been externally validated on a large group of patients, this study has several limitations. It focuses on a retrospective analysis of data from a single center of mostly postsurgical and trauma patients, and further work is necessary to incorporate our insights into a fully operational predictive algorithm and to evaluate it in a rigorous, prospective framework (15) (e.g., an online predictive algorithm to alert the future CRI risk evolution given vital sign features) in a more varied patient case mix.



**Figure 4.** Estimated trajectories of stratified relative risk (RRisk) groups and stratified personalized RRisk (PRRisk) groups. Estimated group trajectories for RRisk (*left*) and PRRisk (*right*). Time index 0 corresponds to cardiorespiratory insufficiency (CRI) onset time for both plots. Groups are color coded. The *solid lines* reflect mean trajectories estimated from all raw trajectories for the representative groups, as determined by the maximum posterior probability from the group-based model. *Shaded areas* depict the associated 95% confidence intervals. Separate groups that include sampled trajectories from all control subjects are plotted in *black*. 1 (*red*) = late onset low risk; 2 (*blue*) = late onset medium risk; 3 (*green*) = late onset high risk; 4 (*purple*) = gradual increase in risk; 5 (*orange*) = persistently high risk.

The model was limited to noninvasive vital sign data. Incorporation of other data domains to the model, such as demographics, medications, and biomarkers, might further improve prediction accuracy and risk group assignment. Importantly, the frequency of acquisition of vital sign data was highly granular, exceeding typical technical availability in most hospital environments. It appears, however, that such granularity is very important in the development of highly predictive analytics, raising the possibility that even higher granularity, such as beat-to-beat and waveform characteristics of vital sign signals, might further improve performance.

Finally, the AUC trend presented here is an aggregated measurement of discrimination during the hours before CRI events, which is primarily driven by the heterogeneous risk evolution patterns of our study patients. The relatively poor discrimination several hours before the event could be attributable to the fact that only a small subset of patients exhibits phenotypes of risk evolution (specifically, persistently high risk and early onset) that are conducive to early CRI detection.

**Conclusions**

This study represents a proof of the concept that analysis of noninvasive vital sign time series data, routinely available in the monitored units of most acute care

hospitals, can be used to accurately predict who will eventually become unstable, approximately when cardiovascular instability will occur, and into which general category of physiologic response each patient will be placed. Importantly, the initial vital sign data upon admission can effectively triage patients into those who never become unstable and those who will likely develop CRI during an index hospitalization. Furthermore, the personalization of each patient’s own vital sign normal trends, referred to as PRRisk, increases the specificity of identifying CRI. Finally, these data show that such machine learning–based analyses can categorize patients who will eventually develop CRI into specific performance groups, the group behavior of which is remarkably similar across patients.

Although these machine learning models were tested on retrospectively collected patient vital sign data, such analyses can be performed in real time, providing the bedside clinical and hospital managers data-driven options for optimizing care delivery and, potentially, preventing CRI from occurring, or at the least, minimizing its deleterious effects on the patient. ■

**Table 2.** Group membership migration patterns between stratified group memberships with respect to relative risk and personalized relative risk

Relative Risk Groups	Personalized Relative Risk Groups		
	Late Onset	Early Onset	Persistently High Risk
Late onset	550	33	30
Early onset	105	45	21
Persistently high risk	110	4	20

“Late Onset” group combines groups 1, 2, and 3 (seen in Figure 4) into a single group. “Early Onset” group is group 4, and “Persistently High Risk” group is group 5.

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

## References

- 1 Buist MD, Moore GE, Bernard SA, Waxman BP, Anderson JN, Nguyen TV. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. *BMJ* 2002;324:387–390.
- 2 Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA* 2002;288:1987–1993.
- 3 Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006;34:1297–1310.
- 4 Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993;270:2957–2963.
- 5 Rothman MJ, Rothman SI, Beals J IV. Development and validation of a continuous measure of patient condition using the electronic medical record. *J Biomed Inform* 2013;46:837–848.
- 6 Chen L, Dubrawski A, Clermont G, Hravnak M, Pinsky MR. Modelling risk of cardio-respiratory instability as a heterogeneous process. *AMIA Annu Symp Proc* 2015;2015:1841–1850.
- 7 Devita MA, Bellomo R, Hillman K, Kellum J, Rotondi A, Teres D, Auerbach A, Chen WJ, Duncan K, Kenward G, et al. Findings of the first consensus conference on medical emergency teams. *Crit Care Med* 2006;34:2463–2478. [Published erratum appears in *Crit Care Med* 34:3070.]
- 8 Hravnak M, Chen L, Dubrawski A, Bose E, Clermont G, Pinsky MR. Real alerts and artifact classification in archived multi-signal vital sign monitoring data: implications for mining big data. *J Clin Monit Comput* 2016;30:875–888.
- 9 Chen L, Dubrawski A, Wang D, Fiterau M, Guillame-Bert M, Bose E, Kaynar AM, Wallace DJ, Guttendorf J, Clermont G, et al. Using supervised machine learning to classify real alerts and artifact in online multisignal vital sign monitoring data. *Crit Care Med* 2016;44:e456–e463.
- 10 Lake DE, Richman JS, Griffin MP, Moorman JR. Sample entropy analysis of neonatal heart rate variability. *Am J Physiol Regul Integr Comp Physiol* 2002;283:R789–R797.
- 11 Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci USA* 1991;88:2297–2301.
- 12 Breiman L. Random forests. *Machine Learning* 2001;45:5–32.
- 13 Nagin DS. Group-based modeling of development. Cambridge, MA: Harvard University Press; 2005.
- 14 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.
- 15 Jiang X, Cooper GF, Neill DB. Generalized AMOC curves for evaluation and improvement of event surveillance. *AMIA Annu Symp Proc* 2009;2009:281–285.