Uncovering Clinical Significance of Vital Sign Dynamics in Critical Care

Li-wei H. Lehman[1], Shamim Nemati[2], George B. Moody[1], Thomas Heldt[1], Roger G. Mark[1]
[1] Massachusetts Institute of Technology, Cambridge, MA
[2] Harvard School of Engineering and Applied Sciences, Cambridge, MA

Abstract

Vital-sign time series of heart rate (HR) and blood pressure (BP) exhibit complex patterns of fluctuations, reflecting the underlying pathological and physiological states of patients. In this work, we adopt a switching vector autoregressive process framework to learn a shared global library of "phenotypic" vital-sign dynamical behaviors from HR and BP time series of a patient cohort. Using HR and BP time series of over 450 adult ICU patients, we demonstrate that the fluctuation patterns in HR and BP are significantly correlated with several laboratory measurements routinely monitored in the ICU, and can potentially be used to reveal the underlying patho-physiological states of the patients. We demonstrate that the bivariate dynamics of HR/BP alone achieve similar performance in sepsis detection in comparison to the SAPS I scores, which use age and the most extreme values of 13 physiological variables. Further, combining the bivariate dynamics of HR/MAP time series and SAPS I provides a significantly more accurate (p=0.02) assessment of patients' risks in sepsis than using SAPS I alone (AUC 0.70 [0.63, 0.78] vs 0.78 [0.70, 0.85]), suggesting that the dynamics of the interaction between HR and BP may contain additional predictive value beyond that contained in the SAPS I scores.

1. Introduction

Vital sign time series of heart rate (HR) and blood pressure (BP) can exhibit complex dynamical patterns as a result of internally and externally-induced changes in the state of the underlying control system. Understanding these dynamical behaviors is both of physiological and clinical importance, and can potentially yield insights into the disease process of patients. In this work, we use a switching vector autoregressive process (SVAR) framework to systematically discover "phenotypic" dynamic behaviors shared across multivariate vital sign time series of a patient cohort. We model the changing dynamics of nonlinear and non-stationary vital sign time series via Markov transitions among a collection of simpler linear dynamical systems (or modes). Our previous work has demonstrated the utility of such a framework in discovering dynamic behaviors with prognostic values in hospital mortality prediction [1].

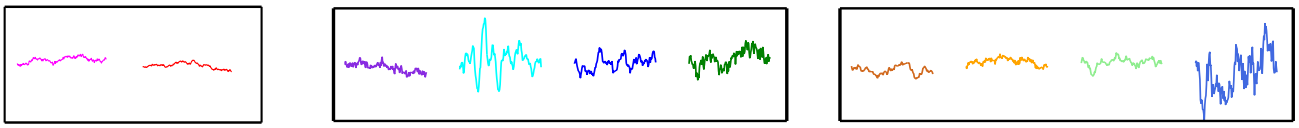
In this paper, we adopt a data-driven approach to explore the physiological and clinical significance of the discovered dynamic behaviors. Using minute-by-minute HR and BP time series from a cohort of ICU patients in the MIMIC II database [2], we investigate the correlations between different dynamic behaviors and a wide range of physiological and lab variables routinely monitored in an ICU. Finally, we investigate the classification performance of our approach in identifying patients with sepsis.

2. Materials and Methods

This study included 453 adult patients from the MIMIC II waveform database (Version 2) [2] with clinical information, and with at least eight hours of continuous minute-by-minute HR and invasive BP measurements during the first 24 hours of their ICU stays. Approximately 15% (67 out of 453) of patients in this cohort died in the hospital.

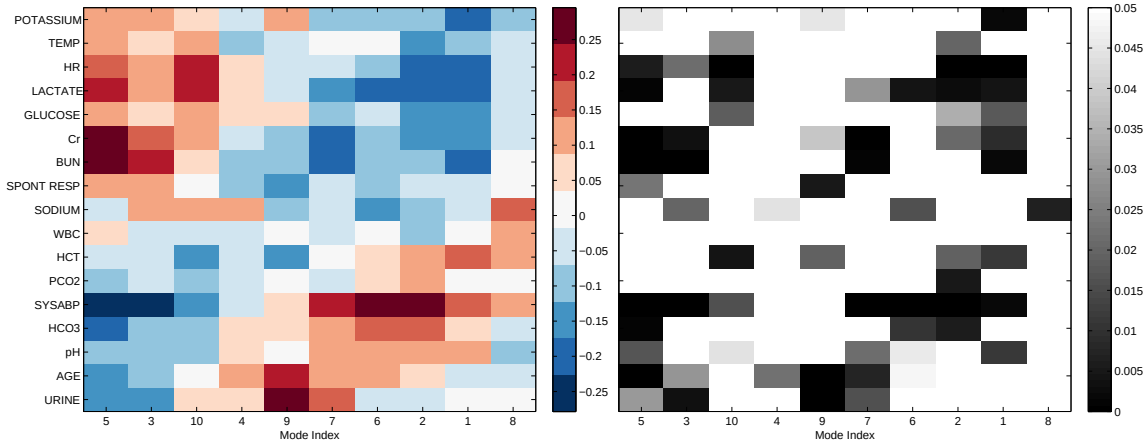
We employed a switching vector autoregressive processes framework to model physiological time series via Markov transitions between a collection of simpler linear dynamical systems [3]. For the n-th patient ($n = 1 \cdots N$), let $y_t^{(n)}$ be a $M \times 1$ vector of observed values of the vital signs at time t ($t = 1 \cdots T^{(n)}$). We assume that there exists a library of K possible dynamics or *modes*, a set of multivariate autoregressive model coefficient matrices $\{A_p^{(k)}\}_{k=1}^K$ of size $M \times M$, with maximal time lag $p = 1 \cdots P$, and the corresponding noise covariances $\{Q^{(k)}\}_{k=1}^K$. Let s_t be a switching variable, indicating the active dynamic mode at time t, and evolving according to a Markovian dynamic with initial distribution $\pi^{(n)}$ and a $K \times K$ transition matrix Z. Following these definitions, an AR-HMM is defined by $y_t = \sum_{p=1}^P A_p^{(z_t)} y_{t-p} + Q^{(z_t)}$. A collection of related time series can be modeled as switching among these dynamic behaviors which describe a locally coherent linear model that persists over a segment of time.

Minute-by-minute HR and mean arterial blood pressure (MAP) time series from MIMIC II were modeled as a switching AR(3) process. The number of dynamic modes (K=20) was determined using the Bayesian Information

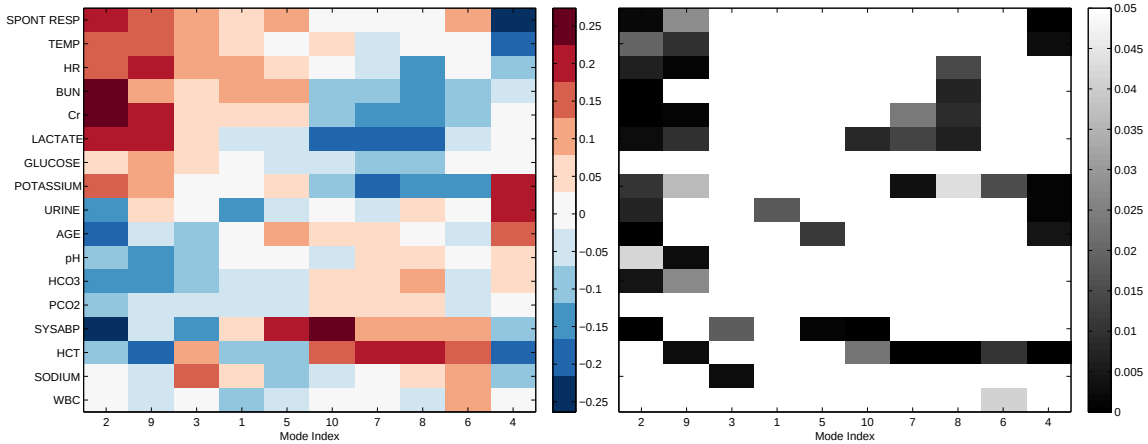


(a)High-risk modes 3 and 5 (b)Low-risk modes 1, 9, 7, and 2 (c)Neutral modes 10, 8, 4, and 6

Figure 1. The top 10 MAP dynamic modes from the first 24-hours in the ICU grouped by their associations with hospital mortality (N=453 patients). All modes were simulated from their AR coefficients and covariances, and plotted with the same time duration (150 minutes) and amplitude scale ([-40, 40] centered around zero).



(a)Correlation: MAP dynamics and physiological/lab variables (b)Significance (p <0.05 in gray/black)



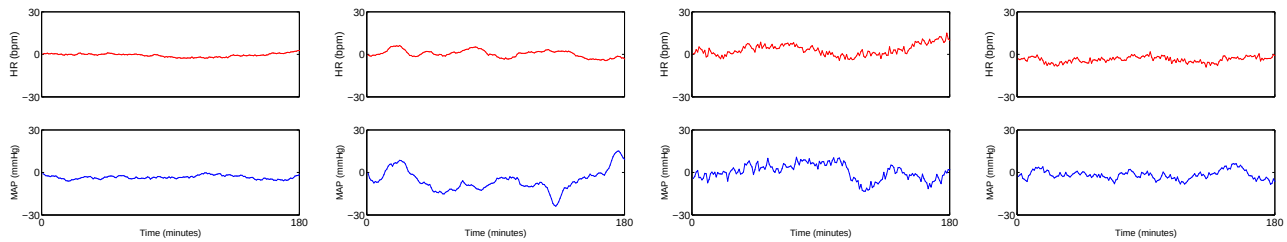
(c)Correlation: HR/MAP dynamics and physiological/lab variables (d)Significance (p <0.05 in gray/black)

Figure 2. Correlation between dynamic mode proportions and other physiological variables from the first 24-hours in the ICU. Cr-creatinine, HCT-hematocrit, WBC-white count. In parts (b) and (d), $p >= 0.05$ are in white.

Criterion (BIC) [4]. We define a *mode proportion* $MP_k^{(n)}$ as the proportion of time the n-th patient spends within the k-th mode [1]. We characterize each time series with its corresponding mode proportion (a $1 \times K$ feature-vector), and use a logistic regression classifier to make predictions about the outcome variables of interest. Test of statistical significance was based on p-values after correcting for the false discovery rate [5]. Comparison of AUCs was based on the method described in [6].

3. Results

Figure 1 shows the top ten most common MAP dynamic modes learned using the SVAR algorithm. Dynamic modes are labeled as high-risk, low-risk, or neutral based on their associations with hospital mortality from logistic regression analysis (see [1]). Note that the high-risk modes appear to contain less variability. To investigate the clinical significance of different dynamic modes, we



(a)Mode 2 (b)Mode 9 (c)Mode 7 (d)Mode 8

Figure 3. Example HR (red) and MAP (blue) bivariate dynamic modes significantly associated with sepsis in the ICU. Modes 2 and 9 are high-risk modes, and modes 7 and 8 are low-risk modes in sepsis. All modes were simulated from their AR coefficients and covariances.

Mode	P-Val	OR(95%CI)	Adjusted P-Val	Adjusted OR(95%CI)	HL PVAL
2	**0.0000**	**1.52 (1.30 1.79)**	**0.0046**	**1.30 (1.08 1.55)**	0.22
9	**0.0085**	**2.14 (1.21 3.78)**	0.5230	1.24 (0.64 2.43)	0.11
7	**0.0006**	**0.20 (0.08 0.50)**	**0.0063**	**0.30 (0.13 0.71)**	0.68
8	**0.0068**	**0.38 (0.19 0.77)**	0.1197	0.56 (0.27 1.16)	0.04
1	0.0525	1.26 (1.00 1.59)	0.0777	1.28 (0.97 1.67)	0.67
3	0.0527	1.33 (1.00 1.77)	0.4533	1.13 (0.82 1.55)	0.63
10	0.1396	0.56 (0.26 1.21)	0.0919	0.46 (0.18 1.14)	0.16
4	0.1876	0.87 (0.70 1.07)	0.5664	0.93 (0.73 1.19)	0.51
5	0.7003	1.07 (0.76 1.52)	0.6067	1.11 (0.74 1.66)	0.08
6	0.7312	1.07 (0.74 1.53)	0.4350	1.17 (0.79 1.75)	0.37

Table 1. Associations between the bivariate dynamics of HR/MAP and the risk of sepsis. In multivariate logistic regression model (columns 4-6), we report the p-values and OR of the dynamic mode proportion variables after adjusting for APACHE IV. The Hosmer-Lemeshow (HL) p values were reported to assess the model fit.

Features	MAP Dynamics	HR Dynamics	HR/MAP Dynamics
Dynamics Only	0.67 (0.58, 0.74)	0.61 (0.58, 0.74)	0.70 (0.60, 0.84)
Dynamics + SAPS I	0.76 (0.69, 0.83)	0.67 (0.59, 0.75)	0.78 (0.70, 0.85)
Dynamics + APACHE IV	0.78 (0.71, 0.85)	0.78 (0.70, 0.85)	0.82 (0.75, 0.88)

Table 2. Classification performance in detecting sepsis in ICU patients using ten most common HR and MAP dynamics from the first 24 hours in the ICU. AUCs (and 95% confidence intervals) are from 10-fold cross validation. The performance of SAPS I and APACHE IV alone were 0.70 [0.63, 0.78] and 0.79 [0.72, 0.86] respectively.

display the correlation coefficients and p-values ($p< 0.05$) between the proportion of time patients spent in different MAP dynamic modes and patient age, daily lab measurements, and other physiological variables from the first 24-hours in the ICU (Figure 1 a, b). Overall, high-risk MAP dynamic modes (3 and 5) are correlated with high heart rate, high respiratory rate, and low systolic blood pressure, which may be signs of hemodynamic instability. Modes 3 and 5 are also correlated with high lactate, creatinine, BUN, and low urine output, suggesting potential end-organ and tissue hypo-perfusion. High-risk mode 5 is significantly correlated with a low pH, HCO3 (bicarbonate), in addition to high lactate; it has a weak correlation with high temperature, though the correlation is not significant. Together, these signs point to a state of potential metabolic acidosis or sepsis.

Low-risk MAP modes (1, 9, 7 and 2) are, in general, correlated with the same physiological variables as the high-risk modes, though the correlations are in the opposite direction. Notably, low-risk dynamic modes 1 and 2 are correlated with lower heart rate, and higher systolic blood pressure. All four low-risk modes are correlated with low creatinine, and modes 9 and 7 are additionally correlated with high urine output, typical of normal renal function. All four low-risk modes are correlated with low lactate, and modes 1 and 7 with high pH, an indication of likely absence of metabolic acidosis.

In the bivariate HR/MAP case (Figure 1 c, d), note that dynamic modes 2 and 9 are significantly correlated with high respiratory rate, temperature, and heart rate, which

point to symptoms related to systemic inflammatory response syndrome (SIRS) [7]. Mode 2 is further correlated with high creatinine and low urine output, which suggest end-organ damage characteristic of patients with severe sepsis [7, 8]. Table 1 summarizes the association analysis between the top ten most common bivariate HR/BP dynamic modes and sepsis. We built a separate univariate/multivariate logistic regression model for each of the top ten most common dynamic modes. In the multivariate case, the dynamic mode proportion was the primary predictive variable, and APACHE IV was added as a covariate. Four of the modes (see Figure 3) had a significant association with sepsis ($p < 0.05$): mode 2 and 9 ("high-risk modes") had odds ratios greater than 1.0 and therefore were associated with an increased chance of sepsis; modes 7 and 8 ("low-risk modes") had odds ratios less than 1.0 and thus were associated with a lower sepsis risk. The results presented in Table 1 indicate that even after adjustment for APACHE IV, two modes (2 and 7) were significantly associated with sepsis.

Table 2 displays the 10-fold cross-validated classification performance in detecting sepsis. We report the area under the receiver operating characteristic curve (AUC) with 95% confidence intervals. Among the 453 patients, 62 (13.7%) had sepsis (as defined by Martin sepsis criteria [8]). For sepsis detection, the bivariate dynamics of HR/MAP achieved an AUC of 0.70 [0.60, 0.84] which is comparable to the performance of SAPS I alone with an AUC of 0.70 [0.63, 0.78]. Combining the bivariate dynamics of HR/MAP with SAPS I significantly (p=0.02) improved the performance of SAPS I from an AUC of 0.70 [0.63, 0.78] to 0.78 [0.70, 0.85]. State-of-the-art APACHE IV achieved an AUC of 0.79 [0.72, 0.86]. Combining HR/MAP dynamics with APACHE IV improved the APACHE IV performance slightly from 0.79 [0.72, 0.86] to 0.82 [0.75, 0.88], but the performance gain was not statistically significant (p=0.18).

4. Discussions and Conclusions

In this work, we presented a framework for extracting and interrogating the clinical significance of vital sign dynamics shared across a patient cohort. Using HR/BP time series of over 450 adult ICU patients, we demonstrated that the fluctuation patterns in HR and BP were significantly correlated with several laboratory measurements routinely monitored in the ICU, and can potentially be used to track the patho-physiological states of the patients. We demonstrated that the bivariate dynamics of HR and MAP achieved similar performance in detecting sepsis in comparison to that of the SAPS I scores, which used age and the most extreme values of 13 variables. Combining the bivariate dynamics of HR/MAP time series and SAPS I provided a more accurate assessment of patients' risks in sepsis than using SAPS I alone, suggesting that the dynamics of the interaction between HR and BP may contain additional predictive value beyond that contained in the SAPS I scores alone. Our ongoing work involves integrating multimodal measurements of vital sings with clinical data, including lab tests, medications, and clinical interventions (administration of fluids, pressors, and titration of medications) to design a risk score capable of continuous monitoring of ICU patients and informing therapeutic interventions.

Acknowledgments

The authors thank Dr. Louis Mayaud for identifying the sepsis cohort in MIMIC II. This work was supported by the National Institutes of Health (NIH) grant R01-EB001659 and R01GM104987 from the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and the James S. McDonnell Foundation Postdoctoral grant.

References

[1] Lehman LH, Adams RP, Mayaud L, Moody GB, Malhotra A, Mark RG, Nemati S. A physiological time series dynamics-based approach to patient monitoring and outcome prediction. IEEE Journal of Biomedical and Health Informatics 2014;Accepted.

[2] Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman LH, Moody G, Heldt T, Kyaw TH, Moody B, Mark RG. Multiparameter intelligent monitoring in intensive care (MIMIC II): a public-access intensive care unit database. Critical Care Medicine May 2011;39(5):952–960.

[3] Murphy KP. Switching Kalman filter. Compaq Cambridge Research Laboratory Tech Rep 98 10 1998;Cambridge, MA.

[4] Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978;6(2):461–464.

[5] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 1995;57(1):289–300.

[6] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;837–845.

[7] Bone R, Balk R, Cerra F, Dellinger R, Fein A, Knaus W, Schein R, Sibbald W. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. CHEST 1992;101:1644–1655.

[8] Martin GS, Mannino DM, Eaton S, Moss M. The epidemiology of sepsis in the United States from 1979 through 2000. The New England Journal of Medicine 2003;1546–1554.

Address for correspondence:

Li-wei Lehman (Email: lilehman@mit.edu)