2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Prediction of severe adverse event from vital signs for post-operative patients

Ying Gu[1], Søren M. Rasmussen[1], Jesper Mølgaard[2,5], Camilla Haahr-Raunkjær[2,5], Christian S. Meyhoff[3,4,5]*, Eske K. Aasvang[2,5]*, and Helge B. D. Sørensen[1]*, IEEE Senior Member

Abstract— **Monitoring post-operative patients is important for preventing severe adverse events (SAE), which increases morbidity and mortality. Conventional bedside monitoring system has demonstrated the difficulty in long term monitoring of those patients because majority of them are ambulatory. With development of wearable system and advanced data analytics, those patients would benefit greatly from continuous and predictive monitoring. In this study, we aim to predict SAE based on monitoring of vital signs. Heart rate, respiration rate, and blood oxygen saturation were continuously acquired by wearable devices and blood pressure was measured intermittently from 453 post-operative patients. SAEs from various complications were extracted from patients' database. The trends of vital signs were first extracted with moving average. Then four descriptive statistics were calculated from trend of each modality as features. Finally, a machine learning approach based on support vector machine was employed for prediction of SAE. It has shown the averaged accuracy of 89%, sensitivity of 80%, specificity of 93% and the area under receiver operating characteristic curve (AUROC) of 93%. These findings are promising and demonstrate the feasibility of predicting SAE from vital signs acquired with wearable devices and measured intermittently.**

I. INTRODUCTION

Major abdominal surgery is associated with risk of severe adverse events (SAE) [1], which might lead to preventable and avoidable deaths in the hospital. It has been reported that clinical deterioration was preceded with SAEs and often reflected by the abnormal changes of vital signs [2]. Therefore, monitoring of vital signs and timely prediction of deterioration from those monitoring would make difference in morbidity and mortality for post-operative patients.

Monitoring of vital signs is a standard procedure for the hospitalized patients. It is often performed intermittently by bedside monitoring system, which show some difficulties for monitoring of mobilized patients. Mobilization after the surgery has proven to be a vital step in recovery and rehabilitation [3]. This calls for the use of wearable monitoring system. Currently, the early warning scores (EWS), which is an aggregate weighted scoring system based on values of vital signs, is the most commonly used scoring system to assist in evaluating patients' risk of complications and provide corresponding treatments. Despite being widely used, EWS has some limitations due to its simple model. Time-based correlation among vital signs is not the focus of EWS. It only represents present status and provides no information about future possible development of vital signs. In addition, EWS is often calculated on intermittent observations of vital signs, which might be inadequate. Patients may deteriorate significantly between observations. With technical development in electronic miniaturization, wearable technology, wireless communication, computing power and data analytics, continuous monitoring of vital signs combined with advanced data analysis would overcome the limitations and challenges faced by current monitoring system for hospitalized patients. The researches have shown that random forest performed better than EWS for prediction of clinical deterioration [4], [5]. Machine learning based approaches for predictive monitoring have been adopted greatly. The performance varied substantially based on various clinical settings, machine learning method used and different observation and prediction windows, see [6] for a review.

A lot of research has been devoted to predicting ICU readmission and mortality by using bedside monitoring system or medical records [7], [8]. Some studies have focused on predicting one type of SAEs such as sepsis onset [9] and cardiac arrest [10]. Other study tried to predict the SAEs resulting from cardiac arrest, intensive care unit transfer and death [4]. Research by Clifton et al. [11] tried to detect abnormality by statistical model trained from normal states of patients. In this paper, we extracted SAEs resulting from neurologic, respiratory, circulatory, infectious and other complications from patients' database. The patients' vital signs were monitored. The objective was to prove the feasibility of prediction of SAE based on continuous monitoring of heart rate (HR), respiration rate (RR), and blood oxygen saturation (SpO_2) by wearable devices and intermittent measurement of blood pressure (BP). To the best of our knowledge, this is the first study to predict SAEs representing various severe complications.

* shared last authorship.
Corresponding author: ying Gu (email: yingu@dtu.dk)
[1]Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
[2]Department of Anaesthesiology, Centre for Cancer and Organ Dysfunction, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
[3]Department of Anaesthesia and Intensive Care, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark
[4]Copenhagen Center for Translational Research, Copenhagen University Hospital, Bispebjerg and Frederiksberg, Copenhagen, Denmark
[5]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

II. MATERIALS AND METHODS

A. Patients

The study took place at Rigshospitalet and Bispebjerg Hospital in Copenhagen, Denmark from February 2018 to August 2020. It is a sub-project of Wireless Assessment of Respiratory and circulatory Distress' (WARD) project. The study and experimental procedures on patients were approved by the Danish Data Protection Agency (2012-58-0004) and registered at http://ClinicalTrials.gov (project: NCT03491137). 500 post-operative patients participated the study. 8 patients were excluded because they were not part of the study and 39 were excluded due to having less than 10 hours vital sign recording (10 hours data were used for observation window, see detail in section *C*). Finally, 453 post-operative patients (278 males, 175 females) were included for further analysis. Mean of age was 71 years old (range: 60–93). Mean of monitoring hours was 79 hours (range: 0.73-168.8). Patients in the study had a wide range of clinical SAEs ranging from neurologic, respiratory, circulatory, infectious and other complications. Information about SAEs were registered by medical doctors. Patient data containing all clinically relevant information were organized and stored in a local database. All patients gave their written informed consent for the study.

B. Clinical vital signs monitoring

The vital signs HR, RR and SpO_2 were acquired continuously by the wearable sensors and BP was measured intermittently. The acquisition of vital signs was managed by Isansys patient status engine (PSE) (Isansys Lifecare Ltd). The Isansys Lifetouch was attached to patients' chest for acquiring single lead ECG with sampling frequency of 1000Hz, from which HR in beats per minute and RR in breaths per minute were derived. Pulse Oximeter (Nonin Model 3150 WristOx2) was attached to the finger for the acquisition of the photoplethysmogram (PPG) with sampling frequency of 75 Hz, from which SpO_2 as a percentage was derived. The wearable sensors' data and derived values were first transmitted via Bluetooth to gateway of PSE, which was located near the bed of the patient, and then to a hospital server for storing data in patients' database via WiFi every minute. Systolic blood pressure (sysBP) in mmHg was measured intermittently by using Meditech BlueBP-05. These sysBP measurements were entered into gateway by medical staff and then automatically transmitted to patients' database. HR, RR, SpO_2 and sysBP were synchronized through their timestamps.

C. Severe adverse event prediction

In essence, predicting SAE is a classification problem. It aims to classify SAE versus no SAE in a few hours (prediction window) based on last recordings (observation window). In this study, prediction window was chosen to be two hours and observation window was chosen to be ten hours as shown in Fig. 2(a). The prediction of SAE was based on the features extracted from trends of four time series HR, RR, SpO_2 and sysBP and on classification carried out with support vector machine (SVM). Fig. 1 depicts the steps of the procedure.

1) Extraction of SAE class and control class: SAE class was identified based on SAEs' timestamps. To account for class imbalance, SAE class was oversampled. SAE class

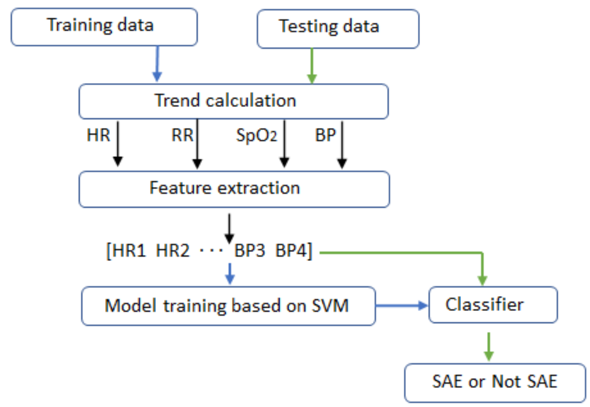

Fig. 1. Overview of SAE prediction

samples were extracted as eight hours' time series of vital signs with overlapping from two hours before to twelve hours before SAE timestamp. Four samples were extracted for each SAE as illustrated in Fig. 2(a). Control class samples were extracted from patients who did not have SAEs during vital sign monitoring at hospital and the monitoring duration was at least eight hours. Fig. 2(b) illustrates the extraction of control samples. The samples were extracted during the whole monitoring period to cover all possible patients' statues.

2) Feature extraction: Extracting discriminative features is important for the prediction of SAE. The clinical deterioration is often preceded with SAE and is reflected in vital signs. In this study, first the trends of HR, RR, SpO_2 and sysBP were extracted by using moving average with sliding window of 60 minutes. The trends were supposed to represent the deterioration. Then four descriptive statistics (maximum, minimum, mean, and standard deviation) were calculated from the trend of each modality as features. The features from each modality were concatenated into one feature vector. The length of the feature vector was sixteen.

3) Classification based on support vector machine: The SVM is a supervised machine learning algorithm for solving

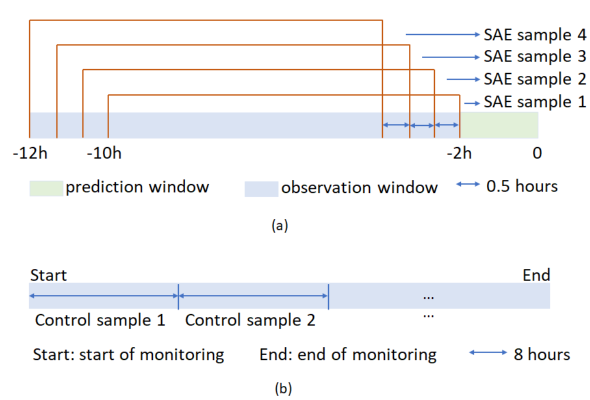

Fig. 2. (a) Illustration of extraction of SAE samples (0 represent SAE timestamp, h: hour) (b) Illustration of extraction of control samples

classification and regression problems. It has shown good generalization property in many applications [12], [13]. The basic idea is to construct an optimal hyperplane for linearly separable patterns. The optimal hyperplane is the one that has maximal margin between two classes. For the non-linearly separable patterns, which most real world problems involve, one solution is to transform original data into a higher or indefinite dimensional space and then find a separating hyperplane in the transformed space by using kernel function. Given a training set (x_i, y_i), $i = 1, \ldots, N$ where $x_i \in R^n$ and $y_i = \{\pm 1\}$, x_i is a data point and y_i indicates the class which the point x_i belongs to. The output of the classifier is defined as

$$y(x_i) = sign\left[w^T \varphi(x_i) + b\right] \quad (1)$$

where the function φ maps x_i into a higher dimensional space. w is the weight vector and b is the bias of the hyperplane. The standard SVM requires the solution of the following optimization problem [14]:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + c \sum_{i=1}^{N} \xi_i \quad (2)$$

subject to

$$\begin{cases} y_i\left(w^T \varphi(x_i) + b\right) \geq 1 - \xi_i, & i = 1, \ldots, N \\ \xi_i \geq 0, & i = 1, \ldots, N \end{cases} \quad (3)$$

where ξ_i is a slack variable and c is a penalty parameter. They are used when the training samples cannot be separated without error. Under the circumstances, training samples can be on the wrong side of the hyperplane with a small distance ξ_i. In practice, there is a trade-off between a low training error and a large margin. This trade-off is controlled by the penalty parameter c. A Gaussian kernel k was chosen for non-linear SVM classifier in this study:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) = \varphi(x_i)^T \varphi(x_j) \quad (4)$$

where σ is the width of Gaussian kernel. Tuning of σ is important for optimizing classifier performance.

The classification performance was estimated with 3-fold cross-validation procedure. The misclassification cost $(n\_SAE + n\_control)/n\_SAE$ was given to SAE data samples, whereas $(n\_SAE + n\_control)/n\_control$ to control data samples. Here, n_SAE and $n\_control$ represent the number of data samples belonging to SAE class and control class, respectively. The dataset was randomly partitioned into three subsets. One subset (a testing set) was used to validate the classifier trained on the remaining two subsets (a training set). This process repeated until each subset was validated once. During training, the training set was further divided into subsets for optimizing Gaussian kernel parameter σ and boxconstraints (inner cross-validation). The set of boxconstraints and σ were searched among positive values, with a log-scaled in the range $[10^{-3}, 10^3]$. The optimal boxconstraints and σ were then applied to build classifier for the testing set. The performance of classifier was evaluated in terms of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and the area under receiver operating characteristic curve (AUROC).

III. Results

The performance of the classifier with 3-fold cross validation was summarized in the Table I. The accuracy, sensitivity, specificity, PPV, NPV and AUROC are relatively close among three tests. The classifier achieved an averaged accuracy of 89%, sensitivity of 80%, specificity of 93%, PPV of 82%, NPV of 92% and AUROC of 93%. Additionally, Fig. 3 presented the receiver operating characteristic curves (ROCs) for three tests. The averaged AUROC of 93% indicated the good discriminative power of the classifier.

TABLE I

The performance of classifier from 3-fold cross validation (Acc.: accuracy; Sen.: sensitivity; Spe.: specificity)

	ACC.	SEN.	SPE.	PPV	NPV	AUROC
Test 1	88.91%	80.00%	92.41%	80.56%	92.16%	92.47%
Test 2	89.28%	82.75%	91.85%	80.00%	93.11%	91.96%
Test 3	90.25%	78.62%	94.84%	85.71%	91.84%	94.12%
Average	89.48%	80.46%	93.03%	82.08%	92.37%	92.86%

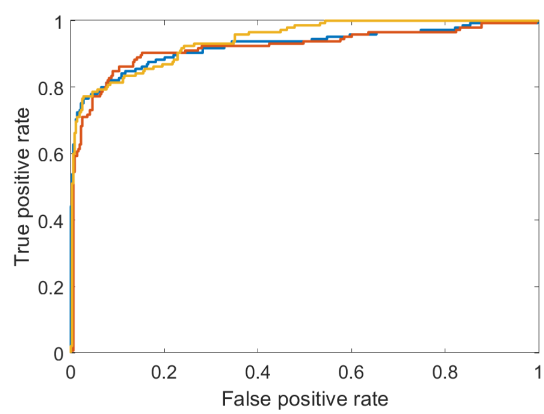

Fig. 3. ROCs from 3-fold cross validation

IV. Discussion

Currently monitoring of post-operative patients relies on intermittent bedside monitor and simple model of EWS in the hospital. Wearable system would facilitate continuous and predictive monitoring and therefore improve the management of patients. The objective of this study is to develop an approach for early prediction of SAEs based on both continuous and intermittent vital signs monitoring and advanced machine learning techniques.

In this study, HR, RR and SpO_2 were acquired continuously with wearable devices. Because reliably continuous measurement of BP was not found during the period of patients' data collection for the study, it was then measured intermittently by the available device. With those acquired vital signs, we have successfully developed an algorithm based on SVM, which could predict SAE in two hours based on last ten hours' recording with AUROC of 93% shown in the Table I and Fig. 3.

For machine learning based approach, the features for training a model are important and should be representative for the difference between classes. In order to reduce the effects of random and transient noises, the trends were first extracted from raw recordings of vital signs by using moving averaging with sliding 60 minutes window. Then the maximum, the minimum, the mean and the standard deviation were calculated from the trends as features. Those descriptive statistics are simple to be calculated and understood and has been used by previous research [15]. 3-fold cross validation was adopted to evaluate the performance of the classifier. The three tests had quite similar result among accuracy, sensitivity, specificity, PPV, NPV and AUROC, which reflected that the developed approach for predication is quite robust. Those six measures indicated that the approach had a powerful discrimination.

Various machine learning based prediction of clinical deterioration has been reported in the literature [9], [11], [16], [17], [18]. The research by Clifton et al [11] applied novelty detection approach for detection of abnormality. The clinicians first identified patients being sufficiently abnormal manually, then using the rest normal patients' data, being larger compared to patients being abnormal, to train a classifier to detect abnormality. They achieved an accuracy of 94%, sensitivity of 96% and specificity of 93%. By calculating variability of vital signs as features, the study using SVM predicted onset of sepsis within the next 4 hours based on recordings from the last 8 hours with an AUROC of 88% [9]. Chen and Qi reported prediction performance of heart failure with AUROC of 84% [16]. In our study, we directly extracted samples of SAEs resulting from neurologic, respiratory, circulatory, infectious and other complications from patients' database. Those SAEs' samples were regarded as SAE class. At the same time, the control class' samples were extracted from patients who did not have SAEs during monitoring period. A classifier for prediction of SAE was trained from those two classes. We have achieved an AUROC of 93% for predicting SAE in 2 hours based on last 10 hours' observation. As main contribution, the paper proves that SAEs resulting from various complications can be predicted by HR, RR and SpO_2 acquired by wearable devices and BP by intermittent measurement. Using descriptive statistics extracted from trends as features and SVM based machine learning technique will reduce computational complexity and therefore require less resources, which is crucial for its implementation in wearable systems. In the future, cuffless-based and continuous measurement of BP will be investigated. The developed approach will be adapted and integrated for clinical validation.

V. CONCLUSIONS

We have developed a machine learning based approach to predict SAEs for post-operative patients. The study has shown that SAEs can be predicated with high AUROC of 93% by four common vital signs, three of which from wearable sensors and one from intermittent measurement. The promising results present an important step towards continuous and predictive monitoring for post-operative patients.

REFERENCES

[1] J. S. Thompson, B. T. Baxter, J. G. Allison, F. E. Johnson, K. K. Lee, and W. Y. Park, "Temporal Patterns of Postoperative Complications," Arch Surg, vol. 138, no. 6, pp. 596–603, 2003.

[2] D. R. Goldhill, "Preventing surgical deaths: critical care and intensive care outreach services in the postoperative period," Br J Anaesth, vol. 95, no. 1, pp. 88–94, 2005.

[3] O. Ljungqvist, M. Scott, and K.C. Fearon, "Enhanced Recovery After Surgery: A Review," JAMA Surg, vol. 152, no. 3, pp. 292-298, 2017.

[4] M. M. Churpek, T. C. Yuen, C. Winslow, D. O. Meltzer, M. W. Kattan, and D. P. Edelson, "Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards," Crit Care Med, vol. 44, no. 2, pp. 368-374, 2016.

[5] M. Green, H. Lander, A. Snyder, P. Hudson, M. Churpek, and D. Edelson, "Comparison of the Between the Flags calling criteria to the MEWS, NEWS and the electronic Cardiac Arrest Risk Triage (eCART) score for the identification of deteriorating ward patients," Resuscitation, vol. 123, pp. 86-91, 2018.

[6] S. Muralitharan, E. Nelson, S. Di, M. McGillion, P. J. Devereaux, N. G. Barr, and J. Petch, "Machine Learning-Based Early Warning Systems for Clinical Deterioration: Systematic Scoping Review," J Med Internet Res, vol. 23, no. 2, e25187, 2021.

[7] J. M. Kwon, Y. Lee, Y. Lee, S. Lee, H. Park, and J. Park, "Validation of deep-learning-based triage and acuity score using a large national dataset," PLoS One, vol. 13, no. 10, e0205836, 2018.

[8] Y. D. Chi, S. S. Villar, J. W. Brand, M. V. Patteril, D. J. Morrice, J. Clayton, and J. H. Mackay, "Logistic early warning scores to predict death, cardiac arrest or unplanned intensive care unit re-admission after cardiac surgery," Anaesthesia, vol. 75, no. 2, pp. 162-170, 2020.

[9] E. Bloch, T. Rotem, J. Cohen, P. Singer, and Y. Aperstein, "Machine learning models for analysis of vital signs dynamics: a case for sepsis onset prediction," J Healthc Eng, vol. 2019, 5930379, 2019.

[10] J. C. Ho and Y. Park, "Learning from different perspectives: Robust cardiac arrest prediction via temporal transfer learning," Annu Int Conf IEEE Eng Med Biol Soc, 2017, pp. 1672–1675.

[11] L. Clifton, D. A. Clifton, M. A. F. Pimentel, P. J. Watkinson, and L. Tarassenko, "Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors," IEEE J Biomed Health Inform, vol. 18, no. 3, pp. 722-730, 2014.

[12] K. Vandecasteele, T. De Cooman, J. Dan, E. Cleeren, S. Van Huffel, B. Hunyadi, and W. Van Paesschen, "Visual seizure annotation and automated seizure detection using behind-the-ear electroencephalographic channels," Epilepsia, vol. 61, no. 4, pp. 766-775, 2020.

[13] A. Nguyen, S. Ansari, M. Hooshmand, K. Lin, H. Ghanbari, J. Gryak, and K. Najarian, "Comparative Study on Heart Rate Variability Analysis for Atrial Fibrillation Detection in Short Single-Lead ECG Recordings," Annu Int Conf IEEE Eng Med Biol Soc, 2018, pp. 526-529.

[14] C. Cortes and V. Vapnik, "Support-Vector Networks," Mach Learn, vol. 20, pp. 273–297, 1995.

[15] B. Hunyadi, M. Signoretto, W. Van Paesschen, J. A. K. Suykens, S. Van Huffel, and M. De Vos, "Incorporating structural information from the multichannel EEG improves patient-specific seizure detection," Clin Neurophysiol, vol. 123, no. 12, pp. 2352–2361, 2012.

[16] Y. Chen and B. Qi, "Representation learning in intraoperative vital signs for heart failure risk prediction," BMC Med Inform Decis Mak, vol. 19, no. 1, 260, 2019.

[17] R. M. Olsen, E. K. Aasvang, C. S. Meyhoff, and H. B. Dissing Sorensen, "Towards an automated multimodal clinical decision support system at the post anesthesia care unit," Comput Biol Med, vol. 101, pp. 15-21, 2018.

[18] A. Youssef Ali Amer, F. Wouters, J. Vranken, D. de Korte-de Boer, V. Smit-Fun, P. Duflot, M. H. Beaupain, P. Vandervoort, S. Luca, J. M. Aerts, and B. Vanrumste, "Vital Signs Prediction and Early Warning Score Calculation Based on Continuous Monitoring of Hospitalised Patients Using Wearable Technology," Sensors, vol. 20, no. 22, 6593, 2020.