Personalized Sedation Level Monitoring in ICU Patients using Heart Rate Variability

Sunil B Nagaraj [1], Siddharth Biswal [1], Valdery M Junior [1], Patrick L Purdon [2], M Brandon Westover [1]

[1]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, U.S.A
[2]Department of Anesthesia, Massachusetts General Hospital, Harvard Medical School, Boston, U.S.A

Abstract

Aim: To develop an automated system to monitor sedation levels in intensive care unit patients using heart rate variability (HRV).

Methods: We developed an automatic sedation level prediction system using HRV as input to a support vector machine learning algorithm. Our data consisted of electrocardiogram recordings from a heterogeneous group of 50 mechanically ventilated adults receiving sedatives in an ICU setting. The target variable was the Richmond agitation-sedation scale score, grouped into four levels: "comatose" (-5), "deeply sedated" (-4 to -3), "lightly sedated" (-2 to 0), and "agitated" (+1 to +4). As input we used 14 features derived from the normalized-RR (NN) interval. We used leave-one-subject-out cross-validated accuracy to measure system performance.

Results: A patient-independent version of the proposed system discriminated between the 4 sedation levels with an overall accuracy of 52%. A patient-specific version, where the training data was supplemented with the patient's labeled HRV epochs from the preceding 24 hours, improved classification accuracy to 60%.

Conclusions: Our preliminary results suggest that the HRV varies systematically with sedation levels and has potential to supplement current clinical sedation level assessment methods. With additional variables such as disease pathology, and pharmacological data, the proposed system could lead to a fully automated system for depth of sedation monitoring

1.   Introduction

Accurate administration and optimization of sedatives dosing is important in intensive care units (ICU) [1]. Patients are often sedated in the ICU to increase their level of comfortness, tolerate several procedures and facilitate mechanical ventilation and analgesia [2,3]. Therefore, it is critical to maintain patient's optimal sedation level since both over- and under sedation could lead to adverse patient outcomes including prolonged mechanical ventilation and ICU stay, increased risk of pneumonia and delirium [4,5]. Currently, subjective methods such as the Riker sedation agitation scale (SAS), Richmond agitation sedation scale (RASS), Ramsay sedation scale are used to score the level of sedation in ICUs. These scoring systems rely mainly on patient's response to external and noxious stimulation, which are subjective and are less discriminative during over-sedated state [6].

Several electroencephalogram (EEG) based methods have been developed to assess patient's level of consciousness during general anesthesia based on spectral and entropy measures [7,8]. However, several studies have demonstrated their ineffectiveness in monitoring patients sedation level in the ICU environment [9,10]. Therefore, a more robust and reliable objective method is crucial to assess the depth of sedation in the ICU.

Heart rate variability (HRV) is a popular noninvasive method that can provide information about the functioning of the autonomous nervous system using electrocardiogram (ECG). We recently investigated the potential of HRV for automatically predicting level of sedation using machine learning method [11]. We proposed a novel system using HRV and support vector machine (SVM) to classify level of sedation. In this paper, we extend our preliminary work to develop a novel automatic patient-specific or personalized sedation level assessment system using SVM and HRV. Due to limited prior knowledge about best set of HRV features for predicting sedation levels, we obtained 14 different time, frequency, nonlinear and complexity measures from the HRV signal. The proposed patient-specific system was rigorously tested and validated on a large ICU database and we believe that the proposed system will provide a reference for developing HRV based sedation level assessment monitors.

2.   Methodology

2.1.   Dataset

ECG recordings from 50 patients (35 males; 15 females) admitted to several ICUs at Massachusetts General

Hospital (MGH), Boston, USA were used in this work. The demographic characteristics of patients are presented in Table 1. This work was performed under an IRB approved protocol. ECG data from GE bedside patient monitors was recorded using BedMaster (Excel Medical Electronics, Jupiter FL, USA) software with a sampling frequency of 240 Hz.

Table 1: Summary of patient demographics used in this study. The values here are reported as minimum, maximum, mean and standard deviation.

Variable	min, max, mean ± SD
Age	20, 86, 58.7 ± 15.2
Weight (kg)	67, 126, 97.4 ± 6.2
No. of days in ICU	1.5, 36, 13.1 ± 8.2
No. of drugs	1,18, 5.5 ± 3.2

2.2. Sedation Assessment

The Richmond agitation-sedation scale (RASS) was used to score patient sedation levels. 10 possible RASS scores were grouped into four categories a priori on clinical grounds for the current work [12]. These RASS groupings were selected as the smallest set of sedation states that seem important to discriminate between within the RASS scale. For convenience in this study, we refer to RASS groupings as A = Agitated = RASS >0; B = "Lightly sedated", RASS 0, -1, -2; C = "Deeply sedated", RASS -3, -4; and D = "Comatose", RASS -5.

2.3. Sedation Level Classification System

The architecture of the proposed HRV based automatic sedation level classification system is shown in Figure 1.
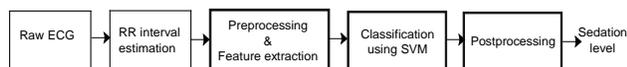
Figure 1: Architecture of the proposed HRV based automatic sedation level classification system

2.3.1 Preprocessing and feature extraction

The raw ECG signal was segmented into short duration epochs (*dur.*) and the Pan-Tompkins method was used to obtain the RR intervals (RRI) from each ECG epoch [13]. Presence of artifacts in the RRI sub-signal is an important issue which could strongly influence the classifier performance. Abnormal ectopic beats were removed using a threshold based method proposed by Clifford et al. [14]. After pre-processing, we regarded the cleaned and interpolated RRI as normalized RRI (or NN interval).

Based on our previous work, 14 features were extracted from the NN interval [11], shown in Table 2. All features were normalized using the box-cox transformation to have uniform mean and standard deviation before feeding it to the SVM classifier for classification.

2.3.2 Classification

The support vector machine (SVM) with a Gaussian kernel was used in this study for classification. Sequential feature selection (SFS) method was used for automatic feature selection. The output of the SVM classifier was converted into posterior probabilities bounded within [0,1] via Platt scaling using a sigmoid function as:

$$P(y = 1 \mid \mathbf{x}) = \frac{1}{1 + exp(A\mathbf{d} + B)}$$

where $\mathbf{d}$ is the distance to the separating hyperplane, A and B are the parameters of the sigmoid function which are estimated over the training dataset.

Table 2: List of HRV features used in this work for the classification of sedation levels.

Domain	Features
Time	Standard deviation of NN (SDNN), Root mean square of SDNN (RMSDD), mean heart rate (MHR), standard deviation of heart rate (SDHR).
Frequency	PVLF-Power in very low frequency spectrum (0.003-0.04 Hz), PLF-Power in low frequency spectrum (0.04-0.15Hz), PHF-Power in high frequency spectrum (0.15-0.4 Hz), PLF/PHF, PLF/PTOT x 100, PTOT is the total power spectrum, PHF/PTOT x 100.
Nonlinear &Complexity	Poincaré plot measures (SD1, SD2), Spectral entropy, Kolmogorov complexity.

2.4 Performance assessment

The performance of the proposed system was evaluated using leave-one-subject-out (LOSO) cross-validation as it results in an unbiased estimation of the true generalization error. In each iteration of the LOSO, 69 patient's data were used for training and the remaining patient's data was used for testing. A 10-fold cross validation on the training data

was used for model selection. This process was repeated until each patient's data was used for testing. Initially, a training set was obtained by leaving out a single subject from the database. This resulted in a split of 69:1 between training and testing set. For the patient-independent system, no data from the testing set was used for training the classifier or tuning other parameters. However, for a patient-specific classification, the first 24 hours data (day *1*) from the testing set was used in the training process and the remaining data (day 2 to day *M*, where *M* is the total no. of days in the ICU) was used for testing. In this manner, we capture the variation in the heart beat morphology specific to the patient being tested.

The resulting training dataset had unequal number of epochs in different RASS groups which could severely bias the performance of the classifier. Due to this, we used *under-sampling* method in which random epochs from each group in training set were selected corresponding to the length of minimum class to create a balanced set. For the patient-specific system it was ensured that all data from day*1* recording of the test patient was included in the balanced training set.

3.   Results

The grouping of RASS scores resulted in a total of 4259 epochs (A = 346, B = 1608, C = 1214 and D = 180). For the purpose of comparison, we estimated the chance level accuracy by calculating the average accuracy attainable by guessing sedation levels A, B, C, or D with probabilities $[p_1, p_2, p_3, p_4]$=[346, 1608, 1214, 180]/3348 = [0.1033,0.4803, 0.3626, 0.0538], respectively. This calculation yields an estimated chance level accuracy of $p_1^2 + p_2^2 + p_3^2 + p_4^2 = 0.37$, or 37%.

The training process of the 10-fold cross validation provided 50 different values of *dur.* and the optimal (maximum likelihood) value of 15 minutes (45/50) provided maximum accuracy during the training process. The overall estimated accuracy of the patient-specific and the patient-independent system were 60% and 52%, respectively, substantially better than a chance level accuracy of 37%.

Table 3: Confusion matrix of the proposed sedation classification system and the actual sedation score.

| Actual Sedation group	System output			
	A	B	C	D
A	165	125	75	8
B	143	1105	331	45
C	128	190	826	35
D	11	30	67	64
Accuracy (%)	44.2	68.0	70.0	37.2
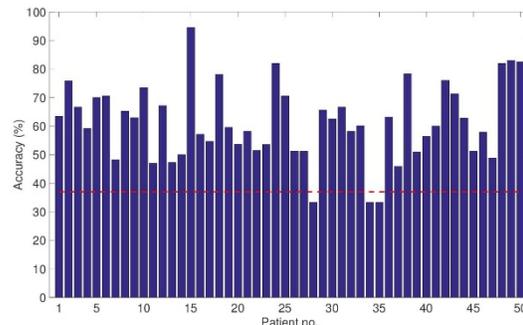
Figure 3. Performance (accuracy, %) of the sedation level detection system for each patient.

Figure 3 shows the accuracy obtained using patient-specific system for individual patients. Except for 3 patients, the accuracy obtained was better than random chance for all other patients. For these three recordings, most misclassification occurred between RASS scores bordering the two groups (-4 and -5, 1 and 0). Since the precision with which nurses assign patients into nearby RASS categories is limited, the problem of category label noise is an important issue to resolve.

4.   Discussion

This paper extends the preliminary work demonstrating the potential of HRV features to predict patient's sedation level [11] into an automated patient specific sedation level classification system. Several HRV features involving time, frequency and complexity domain were used in this work. A novel artifact reduction pre-processing technique was also proposed to improve the quality of the HRV signal for classification. The system, basically, has 3 steps after preprocessing - extracting several HRV features using filtered HRV signal, selecting the best subset of features for each patient by SFS, and identifying optimal window length and model parameters for best classification by means of multi-class SVM. The results from cross validated classification on a large database (50 patients) show that the proposed system can predict the sedation levels with an accuracy of 60%, 23% better than the theoretical chance-level performance (60% vs 37%). The proposed system in this work is a personalized sedation level classification system in which the classification system is trained and tested for individual patient separately. Using feature selection, different optimal subset of features was obtained for each patient in the LOSO cross validation. However, mean heart rate (MHR), high frequency spectral power (PHF), Poincaré plot measures (SD1, SD2) and Kolmogorov complexity features were always selected across all patients. This means that the SVM model trained for each patient are

different from each other and therefore the presented system can be used as a personalized sedation level monitor. The proposed system using HRV has several advantages over traditional methods for assessing depth of sedation. The HRV (1) is unaffected by inter-observer variability, (2) can be used for continuous monitoring, and (3) does not require multiple electrodes for recording ECG when compared to EEG. The proposed sedation level classification system represents a positive step towards developing a multimodal fully automated system for prediction of sedation levels in ICU patients.

References

[1] Kress JP, Pohlman AS, O'Connor MF, Hall JB. Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. N. Engl. J. Med. 2000;342:1471–1477.

[2] Devlin JW. The pharmacology of oversedation in mechanically ventilated adults. Curr. Opin. Crit. Care. 2008;14:403–407.

[3] Mascia MF, Koch M, Medicis JJ. Pharmacoeconomic impact of rational use guidelines on the provision of analgesia, sedation, and neuromuscular blockade in critical care. Crit. Care Med. 2000;28:2300–2306.

[4] Gehlbach BK, Kress JP. Sedation in the intensive care unit. Curr. Opin. Crit. Care. 2002;8:290–298.

[5] Ouimet S, Kavanagh BP, Gottfried SB, Skrobik Y. Incidence, risk factors and consequences of ICU delirium. Intensive Care Med. 2007;33:66–73.

[6] Jonghe BD, Cook D, Appere-De-Vecchi C, Guyatt G, Meade M, Outin H. Using and understanding sedation scoring systems: a systematic review. Intensive Care Med. 2000;26:275–85.

[7] Jospin M, Caminal P, Jensen EW, Litvan H, Vallverdú M, Struys MM, et al. Detrended fluctuation analysis of EEG as a measure of depth of anesthesia. Biomed. Eng. IEEE Trans. On. 2007;54:840–846.

[8] Li X, Li D, Liang Z, Voss LJ, Sleigh JW. Analysis of depth of anesthesia with Hilbert–Huang spectral entropy. Clin. Neurophysiol. 2008;119:2465–2475.

[9] Arbour R. Impact of bispectral index monitoring on sedation and outcomes in critically ill adults: A case series. Crit. Care Nurs. Clin. North Am. 2006;18:227–241.

[10] Walsh TS, Ramsay P, Lapinlampi TP, Särkelä MO, Viertiö-Oja HE, Meriläinen PT. An assessment of the validity of spectral entropy as a measure of sedation statein mechanically ventilated critically ill patients. Intensive Care Med. 2008;34:308–315.

[11] Nagaraj SB, McClain LM, Zhou DW, Biswal S, Rosenthal ES, Purdon PL, et al. Automatic Classification of Sedation Levels in ICU Patients Using Heart Rate Variability. Crit. Care Med. 2016;

[12 Sessler CN, Gosnell MS, Grap MJ, Brophy GM, O'Neal PV, Keane KA, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. Am. J. Respir. Crit. Care Med. 2002;166:1338–44.

[13] Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm. IEEE Trans. Biomed. Eng. 1985;BME-32:230–6.

[14] Clifford GD, McSharry PE, Tarassenko L. Characterizing artefact in the normal human 24-hour RR time series to aid identification and artificial replication of circadian variations in human beat to beat heart rate using a simple threshold. Comput. Cardiol. 2002;31: 129–132.

Address for correspondence:

Dr. Sunil Belur Nagaraj, PhD
MGH Neurology Department
WACC 827
Boston, MA 02114, USA.
Email: sbelurnagaraj@mgh.harvard.edu
Phone: 617-415-6191 Fax: 617-724-6513