Continuous Prediction of Sedation Levels Based on Signal Inputs: Evaluation of Different Modeling Approaches Including Machine Learning

Presenting Author: Pedro L Gambús

First Author: Matthew McDermott

Co-Authors: Sergio Vide, Mireia Alenyà, Iñaki F Trocóniz, Xavier Borrat, Wei-Hu Weng, Peter Szolovits

Background: Monitoring systems are an essential part in diagnosing and therapeutic management of patients, more specially to control and individualize management of patients under anesthesia. Collecting signals that are direct reflect of organs or systems activity allows the clinician to observe in real time how the patient is doing or responding to a treatment. Adequately collected and analysed data from other populations of patients could help in providing more useful information regarding how a patient might respond to a therapy. The present work wants to demonstrate that mathematical models based on prior information from large sets of data can predict level of sedation or probability of side effects in a period of two minutes ahead of time potentially improving monitoring systems with predictive ability.

Methods: Under IRB Approval a new analysis of a database of 380 patients undergoing deep sedation (TCI propofol and remifentanil) for gastrointestinal endoscopy in Hospital Clinic of Barcelona was performed. Hemodynamic variables, raw and processed EEG, transcutaneous pCO2, pulse oximetry as well as drug input were collected every second. Ramsay Sedation Scores (RSS), response to noxious stimulation such as endoscopy tube introduction were also recorded.

Individuals were divided into ten random folds for cross-validated model evaluation under three high-utility tasks: predicting RSS, GAG response, and spontaneous decompensation. Three classes of model were examined: logistic regression models, random forest models, and deep learning models. Models used 21 continuously measured input features, spanning hemodynamic variables, processed EEG signals, pCO2 and SpO2 variables, and drug input, all with missingness. Models optimized a cross-entropy loss and output a probabilistic prediction of the target.

Results: On all tasks, the optimal model outperformed a majority class prediction baseline by a statistically significant ($p \leq 0.05$) margin. For RSS prediction and GAG prediction, L1-regularized logistic regression performed best, yielding accuracies of 59.2 ± 1.0% (chance 36.6 ± 4.9%) and 78.7 ± 2.2% (chance 74.5 ± 4.1%), respectively. On spontaneous decompensation, a random forest model performed best, yielding accuracy 62.5 ± 4.1% (chance 50%). Dataset size is yet too small to allow deep learning approaches to succeed on these tasks, but sufficient for simpler non-linear models (random forests) to offer significant advantages on some tasks.

Conclusions: Machine learning models based solely on the continuous, in-procedure covariates can yield local predictive value in predicting three clinically useful measures: RSS, GAG after endoscopy tube insertion, and spontaneous decompensation. This suggests that increased use of in-procedure machine learning techniques could yield clinical benefits.

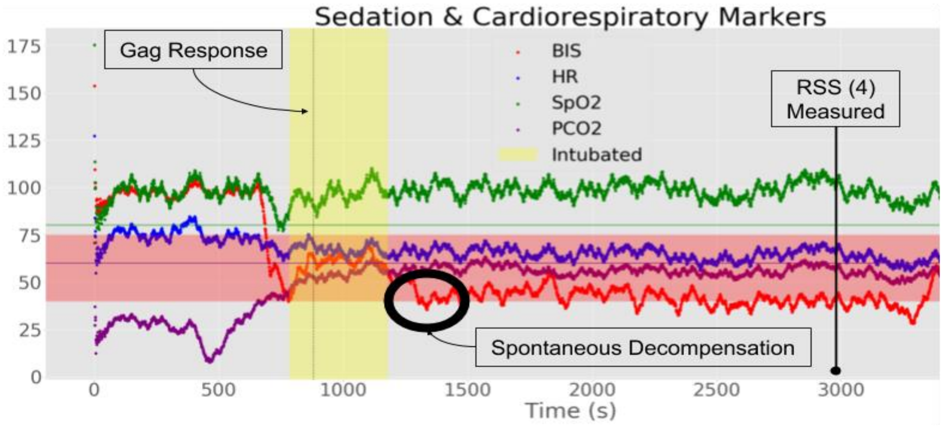

Figure 1. An example of our input data, and prediction target sites.