Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences
School for Information Technology

Master of Statistics

Masterthesis

Can we use the Bispectral Index monitor in out-of-hospital cardiac arrest patients to differentiate a good neurological outcome from a poor neurological outcome?

Selidji Zomahoun Tchala

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics, specialization Biostatistics

SUPERVISOR :
dr. Stijn JASPERS

SUPERVISOR :
Mr. Ward EERTMANS

2017
2018

Faculty of Sciences
School for Information Technology
Master of Statistics

Masterthesis

Can we use the Bispectral Index monitor in out-of-hospital cardiac arrest patients to differentiate a good neurological outcome from a poor neurological outcome?

Selidji Zomahoun Tchala

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics, specialization Biostatistics

SUPERVISOR :

dr. Stijn JASPERS

SUPERVISOR :

Mr. Ward EERTMANS

Acknowledgement

I would like to express my deepest appreciation to Dr. Stijn JASPERS and Dr. Ward EERTMANS for their patience, tolerance, advice and guidance prior to realization of this thesis. A big thank you goes to all my professors and lecturers for their incredible service to us the students and to let them know that their sacrifice was not in vain.
To Jethem' Josiane my wife, a special thank you for her continuous support, her sacrifice. My deepest appreciation to my mother, my brother and friends Dr. Hervé T. V. ZOMAHOUN, Dr. Pascal WELE, Abou ISSIAKA, Louis BIAO, Carl, Sylvester, Koge, Timme and Eric.

Above all, I am most grateful to the Almighty GOD for his continuous blessings over my life.

Abstract

Neurological recovery can not be assessed accurately by physical examination before 72 hours[3]. Bispectral index monitoring (BIS), originally designed to determine the depth of anesthesia, has been found a user friendly and widespread EEG-monitoring, to test its usefulness as a prognostic tool for neurological outcome in the out-of-hospital cardiac arrest patients.

To address this topic, we used data collected on 77 patients over 36 hours. The given dataset has 30% missingness values and thus some missingness technique had been employed in order to handle the observed missing pattern. We used the complete deletion case for the assessment and one set of the five dataset obtained from Multiple imputation technique, to conduct the sensitivity analysis.

Five different prediction techniques had been used, to evaluate the highest accuracy, specificity, sensitivity and the time at which these metrics values are obtained. Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Neural Network (NN), Joint Modeling (JM) technique with Linear Mixed Effect model for the first sub-model and with Logistic Regression as second sub-model and then the second Joint Model with Neural Network as second sub-model. The Linear Mixed Effect model was fitted with BIS as continuous response variable and time as main covariate.

As a result, we obtained the best accuracy value for BIS alone at hour 9 and all metrics give 100% as value at hour 24 for the full model. While losing a little sensitivity and keeping the specificity at 100%, we gain in predicting an earlier neurological outcome status for hour 9 with 75% Sensitivity, 90.91% accuracy and 100% Specificity.

Thus, based on the JM-NN results, BIS values measured continuously can be used to predict poor neurologic outcome status for OHCA patients.

Keywords: Cardiac arrest, Bispectral index, Prediction, support Vector Machine, Neural Network, Joint model, complete case deletion, Single Imputation.

List of Tables

1 Correlation coefficients (P-value) . . . 15
2 Analysis of Maximum Likelihood Estimates of full model fitted over 17 hours . . . 21
3 Confusion matrix from JM-NN with BIS alone at hour 9 . . . 22
4 Table Summarizing Method's Evaluation . . . 23
5 Method evaluation table using imputed data . . . 26
6 Five first hours result for the SVM model with BIS alone . . . 31
7 Five first hours result for the SVM full model . . . 31
8 Five first hours result for the LDA model with BIS alone . . . 31
9 Five first hours result for the LDA full model . . . 32
10 Five first hours result for the NN model with BIS alone . . . 32
11 Five first hours result for the NN full model . . . 32
12 Joint Logistic regression on complete deletion case for model with BIS alone . . . 33
13 Joint Logistic regression on complete deletion case for the full model . . . 33
14 Joint Model with NN: Complete case deleted BIS alone . . . 33
15 Joint Model using Neural Network with all covariates . . . 34
16 Sensitivity analysis for SVM on model with BIS alone . . . 35
17 Sensitivity analysis for SVM on full model . . . 35
18 Sensitivity analysis using LDA for model of BIS alone . . . 35
19 Sensitivity analysis using LDA for the full model . . . 36
20 Sensitivity analysis using NN for the model of BIS alone . . . 36
21 Sensitivity analysis using NN for the full model . . . 36
22 Sensitivity analysis for joint logistic regression model on BIS alone . . . 37
23 Sensitivity analysis for full joint logistic regression model . . . 37
24 Joint model with NN : BIS alone . . . 37
25 Joint Model with NN: Full model results on sensitivity analysis . . . 38

List of Figures

1 Structure of the data set . . . 2
2 Working methodology steps . . . 4
3 Schema of prediction modeling steps . . . 5
4 Confusion matrix for diagnostic testing . . . 6
5 Mean profile plot . . . 14
6 Variance structure plot . . . 14
7 Missingness pattern of the original dataset . . . 15
8 Missingness pattern of the imputed dataset . . . 16
9 Convergence plot at 5000 iterations and 5 maxit . . . 17

Contents

1 Introduction **1**

2 Theoretical framework **2**
2.1 General Objective . . . 2
2.2 Specifics Objectives . . . 2
2.3 Hypotheses . . . 2

3 Data description **2**
3.1 Missingness in the data . . . 3
3.2 Missingness handling Mechanism: Multiple Imputation . . . 3

4 Methodology **4**
4.1 Exploratory Data Analysis . . . 4
4.2 Statistical data analysis . . . 4
4.3 Prediction models . . . 5
4.3.1 Support Vector Machine (SVM) . . . 7
4.3.2 Linear discriminant analysis . . . 8
4.3.3 Neural Network . . . 8
4.3.4 Joint modeling . . . 9
4.3.5 The first sub-model: Linear Mixed Model . . . 9
4.3.6 The second sub model: Generalized Linear Model . . . 10
4.4 Software . . . 12

5 Data analysis and interpretation of results **13**
5.1 Exploratory data analysis . . . 13
5.1.1 Descriptive statistics . . . 13
5.1.2 Individual Profiles . . . 14
5.1.3 Correlation analysis . . . 15
5.1.4 Data Pattern . . . 15
5.2 Predictive Models . . . 19
5.2.1 Support Vector Machine results . . . 19
5.2.2 Linear discriminant analysis results . . . 19
5.2.3 Neural Network results . . . 20
5.2.4 Joint Modeling with logistic regression . . . 20
5.2.5 Joint Modeling with Neural Network . . . 21
5.3 Evaluation of the best prediction method . . . 23
5.3.1 BIS alone . . . 23
5.3.2 BIS with all other covariates . . . 23
5.4 Sensitivity analysis . . . 25
5.4.1 Sensitivity analysis: SVM, LDA, NN and the two Joint models . . . 25
5.4.2 Sensitivity analysis for the joint models . . . 25
5.4.3 Sensitivity analysis: Method evaluation . . . 25

6 Discussion, Conclusion and recommendation **27**

Appendices **31**

A Complete case: Metrics measures results **31**

B Sensitivity analysis: Metrics results obtained **35**

1 Introduction

Cardiac arrest (CA) is an abrupt loss of heart functioning [1]. It is a major a major cause of disability and death. Emergency medical services respond to 375,000 cardiac arrest cases, each year in Europe[10]. Although some progress had been made, in care, out-of-hospital cardiac arrest (OHCA) remains a devastating event, with an overall mortality above 90% [23]. Understanding the chronobiology of OHCA is important to clarify the immediate precipitants of sudden cardiac death, develop preventive strategies, and to optimize resource planning for the prehospital and in hospital response to CA[2]. Nowadays, neurologic recovery cannot be assessed accurately by physical examination until at least 72 hours after return to normothermia due to the neurophysiologic effects of hypothermia, sedatives, and muscle relaxants. Therefore, a generally applicable prediction model supporting the early decision to continue with full supportive treatment or to remain conservative would be of major interest during the critical hours following resuscitation. Clinicians have noted the paucity of data to help identify patients early in the course of care that will not recover from cerebral anoxia[6]. Initial bispectral index (BIS) may help to identify patients who will awaken during therapeutic hypothermia after CA [21]. The BIS is a processed EEG signal that was initially developed to measure anesthetic depth but is used in some intensive care units (ICU) to monitor patients' brain activity and titrate sedatives. Several studies already showed that BIS scores and sedation requirements during therapeutic hypothermia may reflect neurologic function and could be used to clinically predict neurologic recovery. This study aimed to predict poor neurological outcome as early as possible using BIS values and other clinical and demographic characteristics.

The theoretical framework is presented in Section 2, the data description in section 3, the methodology in Section 4; thereafter the results and discussion are shown in Sections 5 and 6 respectively.

2 Theoretical framework

2.1 General Objective

The aim of the current study, is to understand the possibility of using BIS index monitor in OHCA to differentiate a good neurological Outcome from a poor.

2.2 Specifics Objectives

1. Find the best performance in predicting poor neurological outcome based on continuous measure BIS values;

2. Predict the appropriate time point for predicting poor neurological outcome using BIS values alone and in combination with demographical covariates.

2.3 Hypotheses

1. We could build a model with high discriminant performance in predicting poor neurological outcome based on continuous measure BIS values;

2. We could predict the appropriate time point for poor neurological outcome using BIS values alone and in combination with demographical covariates.

3 Data description

The dataset given is composed of 77 subjects in whom BIS, SR and EMG values were measured over the first 36 hours after admission to the ICU. The total number of observations in the dataset is 2,772 for 13 variables. PatientID, neurological outcome, Bispectral Index Score (BIS), Gender, Age, Initial Rhythm, Initial Lactate, Initial pH, SR, EMG NSE 24 and NSE 48. There are two type of variables, variables measures one time (PatientID, Outcome, Gender, Age, Initial Rhythm, Initial Lactate, Initial pH, NSE 24 and NSE 48) and those measured continuously over 36 hours (BIS, SR and EMG). Hence, the transposition of the data set gives 77 subjects and 118 variables. After removing the missing values, we end up with 50 observations.

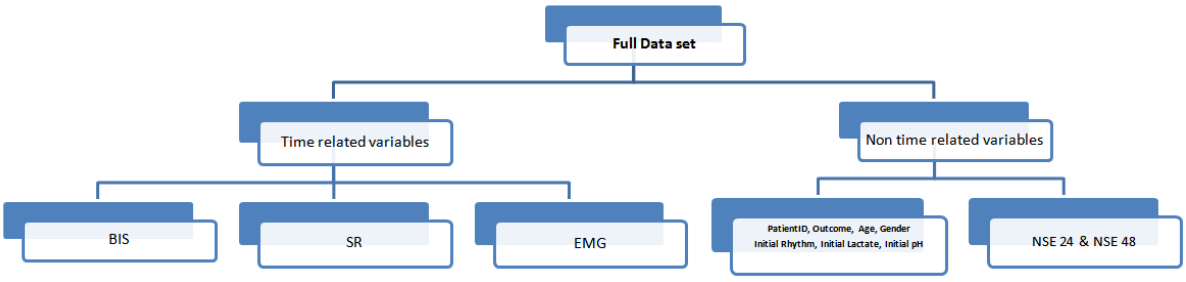

Figure 1: Structure of the data set

3.1 Missingness in the data

An important difficulty when streaming statistical data collected from the field occurs when, not all planned measurements are observed. This results into a reduction of the study subjects leading to a decreasing sample size, hence, increasing variability which in turn leads to loss in efficiency in the validity of statistical inference. Rubin (1976) and Little and Rubin (1987) classified these mechanisms into three possible categories, namely data missing completely at random, at random, or not at random. The missing pattern had been explored, to determine the completer, dropouts as well as those with intermittent missing patterns. In attempting to handle missingness patterns in this project many techniques such as weighted estimating equations, direct likelihood approach could be employed since Verbeke and Molenberghs (2005)[19] argued that likelihood based inference for instance is valid, whenever missingness is in covariates, the mechanism is Missing at random (MAR) and provided that the parameters describing the non-response mechanism are distinct from measurement model parameters (ignorable). According to Rubin (2004), Multiple imputation is a useful mechanism to handle missing data that occur in more than one variables. Multiple imputation is more credible when missing data mechanism is Missing At Random (MAR). A comprehensive knowledge about these mechanisms can be gotten from excellent books by (Little, 1992; Little and Rubin, 2014; Schafer, 1997; Rubin, 1976). Hence, we chose to conduct a multiple imputation technique ([22] [24]), after applying the Complete Case (CC) analysis technique, that is deleting all rows that have missing values.

3.2 Missingness handling Mechanism: Multiple Imputation

Multiple imputation (MI) was formally introduced by Rubin (1987) [22]. This technique replaces each missing value with two or more acceptable values representing a distribution of possibilities. Each missing value is replaced by a set of m plausible values (m being the number of imputations, $m \geq 1$). The missing values are filled in m times to generate m complete data sets. These are generated from a plausible model which is based on a plausible set of parameters drawn from a sampling distribution of the parameter estimates. The m values are ordered in a sense that the first components of the vectors when substituted for the missing values result in one data set, the second components also result in a second data set, and so forth. These imputed values are stored in an auxiliary matrix with one row for each missing value and m columns. These m complete data sets are analyzed by using standard procedures. Results from the analyses are combined for the inference. This process results in valid statistical inferences that properly reflect the uncertainty due to missingness, that is, valid confidence intervals for parameters (Rubin, 1987). For this data set with MAR missing data mechanism, the R package MICE (Buuren et al., 2015) was applied. MICE, is a parametric method which assumes that the missing data are MAR that can handle mixed data type [25]. In practice, after 5000 iterations, ran using the function **"maxit"** in R to ensure convergence within a reasonable time, five (05) samples were generated from which we used the fifth in this study for the sensitivity analysis.

4 Methodology

4.1 Exploratory Data Analysis

The exploratory data analysis (EDA) methods was used to illustrate the structures in the data set [13]. The EDA was conducted by plotting individual and mean profiles together with the variance function and correlation structure in order to get preliminary insight and a better understanding of how the bispectral index score evolved over time and also to identify unrevealed structures in the data.

4.2 Statistical data analysis

The applied methodology cover two types of data set, the dataset obtained using complete deletion case and the one obtained from the multiple imputation technique that will serve for sensitivity analysis. As presented by the figure below, we made use of classification methods such as LDA, SVM, NN Joint Models applied using BIS as only covariate and the full model composed of BIS and demographic covariates.

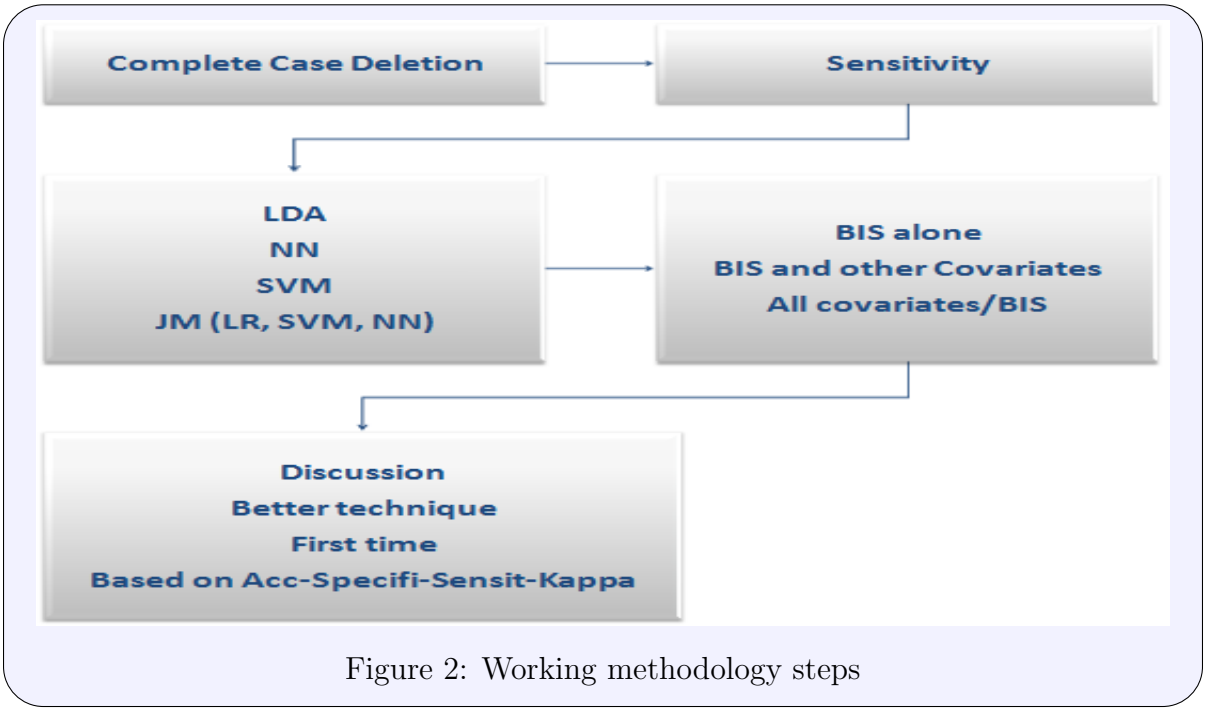

Figure 2: Working methodology steps

4.3 Prediction models

In order to classify patients into survivor group (patients with good neurological outcome) and into non survivor group (those with poor neurological outcome), using their demographics and clinical characteristics, the data set was first split into two parts. The training set, was 80% of the observations and 20% for the test set. The test set was kept aside, in order to assess the accuracy of the prediction made by the model. This model was built under three different classification methods plus two joint models, namely:

1. Support Vector Machine (SVM);
2. Linear Discriminant Analysis (LDA);
3. Neural Network (NN);
4. Joint Model with Logistic Regression (JM LR);
5. Joint Model with Neural Network (JM NN).

In the following lines, we will briefly describe each of these procedures and how they are applied to our data.
This principle lead to many efficient and effective classifiers. It is applied to most learning machines such as Linear Discriminant Analysis (LDA), Neural Networks (NN) and most instance-based methods.
In practice, the procedure had been applied in two (02) steps, the training step and the test step.

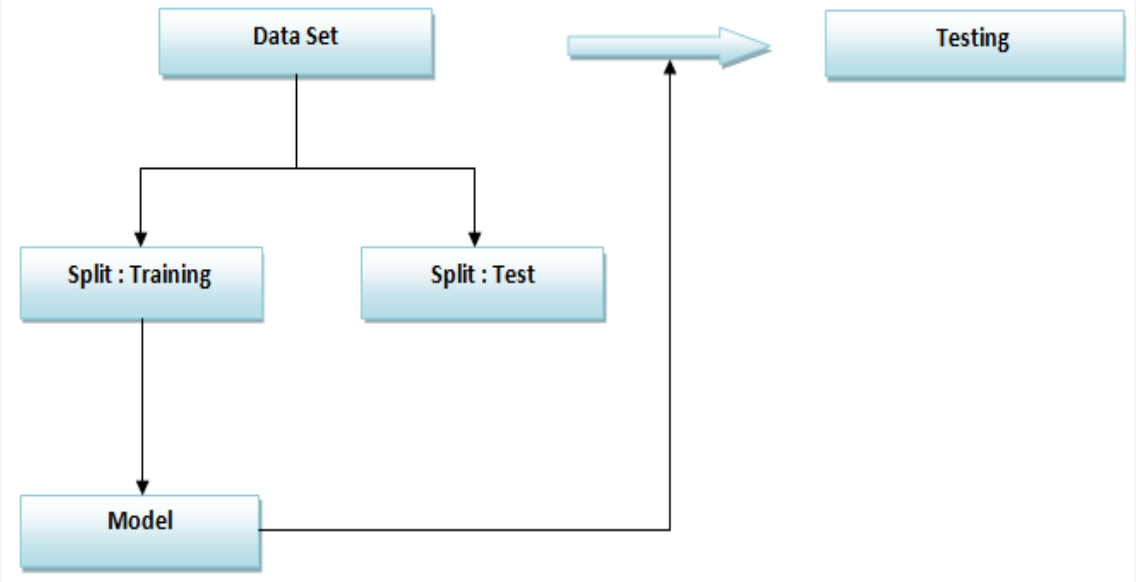

Figure 3: Schema of prediction modeling steps

Step one: Training
The model was fitted on the training data set, using forward model building technique and taking into account the clinical relevance of different covariates. Using a forward selection procedure, we started by a model with one covariate (BIS) then added Age, Gender and other three clinically relevant variables (Initial Rhythm, Initial Lactate and Initial pH) of the patient.

Each model was improved, using tune function that allow for the choice of the best parameters in order to get the most accurate prediction from the fitted model.

Step two: Test
We conducted an accuracy test on the fitted model, using the test dataset, and interpreted the results.
As result, this technique allows us to generate a confusion matrix.

		Actual Class y		
		Positive	Negative	
$h_\theta(x)$ **Test outcome**	Test outcome positive	**True positive** (TP)	**False positive** (FP, Type I error)	Precision = $\frac{\#\text{TP}}{\#\text{TP}+\#\text{FP}}$
	Test outcome negative	**False negative** (FN, Type II error)	**True negative** (TN)	Negative predictive value = $\frac{\#\text{TN}}{\#\text{FN}+\#\text{TN}}$
		Sensitivity = $\frac{\#\text{TP}}{\#\text{TP}+\#\text{FN}}$	Specificity = $\frac{\#\text{TN}}{\#\text{FP}+\#\text{TN}}$	Accuracy = $\frac{\#\text{TP}+\#\text{TN}}{\#\text{TOTAL}}$

Figure 4: Confusion matrix for diagnostic testing

The confusion matrix is a 2 by 2 contingency table that contains in columns the actual class of the response variable (Outcome in our case) and in rows, the prediction result also known as test outcome. Each of them have two (02) levels positive and negative (no or yes) in our case, survivors (0) and non survivors (1).
Indeed, the first column and first row contains the True Positive – the model predicted "+" (non- survivor) and the true class of the patient is "+" (non-survivor). The second column and second row contains the True Negative - the model predicted "-" (survivor) and the true class of the patient is "-" (survivor). The second column, first row contains the False Positive - the model predicted "+" (non-survivor) and the true class of the patient is "-" (survivor) (Type I error). The first column and second row contains the

False Negative - the model predicted "-" (survivor) and the true class is "+" (non-survivor) (Type II error).

4.3.1 Support Vector Machine (SVM)

Support Vector Machines (SVMs) are a fairly new technique for general (nonlinear) classification, regression and outliers detection with an intuitive model representation according to Chang & Lin (2001), developed by Cortes & Vapnik (1995) for binary classification[18]. They have been applied to many different problems, and have been very successful in areas such as face recognition[5], text-categorization ([20], [12]), time-series prediction[14], and hand-written digit recognition [17]. In many of these areas, SVMs have shown to out-perform well-established methods such as Neural Networks and Radial Basis Functions [11].
The SVM[9] in two-class classification problem could be stated as follows:

1. Given a data set D of N pairs: $(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$, with $x_i \in R^p$ and $y_i \in 0, 1$. Each pair is composed of a training example x_i of length M, with elements $x_i = (x_1, x_2, ..., x_M)$.

2. Define a hyperplane by $x : f(x) = x^T\beta + \beta_0$ where β is a unit vector: $\|\beta\| = 1$. A classification rule induced by f(x) is $G(x) = sign[x^T\beta + \beta_0]$

3. $f(x)$ gives the signed distance from a point x to the hyperplane.

4. $f(x) = x^T\beta + \beta_0 = 0$. The classes are separable, we can find a function.

5. $f(x) = x^T\beta + \beta_0$ with $y_i f(x_i) > 0 \ \forall i$, there exist an hyperplane that creates the biggest margin between the training points for class 1 and 0.

6. The goal is to find a classifier with decision function, f(x), such that $f(x_i) = y_i, \forall (x_i, y_i) \in D$.

7. The performance of such a classifier is measured in terms of the classification error defined in equation:

$$error(f(x), y) = \begin{cases} 0, & \text{if } f(x) = y \\ 1, & \text{if otherwise} \end{cases}$$

Consider a learning machine with a set of adjustable parameters α. Given the above binary classification task, the machine seeks to find α such that it learns the mapping $x \mapsto y$. This will result in a possible mapping $x \mapsto f(x, \alpha)$ that defines the machine.
The performance of the machine is measured by the empirical risk error:

$$R_{emp}(\alpha) = \frac{1}{N} \sum_{i=1}^{N} error(f(x_i, \alpha), y_i)$$

where N is the size of the training set and α the set of adjustable parameters. This risk minimization principle is called Empirical Risk Minimization (EMP) [11].

4.3.2 Linear discriminant analysis

Linear discriminant analysis (LDA) is a fundamental tool for classification that classify new observation into the class with closest centroid. LDA method is used in statistics, pattern recognition and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification.
LDA is related to analysis of variance (ANOVA) and regression analysis, which also attempts to express one dependent variable as a linear combination of other variables [16] [7]. However, ANOVA makes use of categorical independent variables and a continuous dependent variable, whereas discriminant analysis has continuous independent variables and a categorical dependent variable [26]. Logistic regression and probit regression appear to be more similar to LDA as compared to ANOVA. Indeed, they explain a categorical variable by the values of continuous independent variables. These methods are preferable in applications where it is not reasonable to assume that the independent variables are normally distributed, which is a fundamental assumption of the LDA method. The LDA function[9] is as follow:

$$\delta_k(x) = x^T\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k + log(\pi_k)$$

The unknown parameters of the Gaussian distributions will be estimated using the training data:

- $\hat{\pi_k} = N_k/N$, where N_k is the number of class-k observations;
- $\hat{\mu_k} = \sum_{g_i=k} x_i/N_k$
- $\hat{\Sigma} = \sum_{k=1}^{K}\sum_{g_i=k}(x_i - \hat{\mu_k})(x_i - \hat{\mu_k})^T/(N-K)$
- $G(x) = argmax_k\delta_k(x)$.

4.3.3 Neural Network

Neural Network (NN) is a network of neurons that is composed of the nonlinear functions of two or more a nonlinear, parameterized function of its input variables. Neural networks (NN) have attracted much attention for their ability of solving a numerous difficult problems in various areas, such as dynamic modeling, pattern recognition and system control involving uncertainty parameters. NN is a version of the additive model (Hastie and Tibshirani, 2016). In practice, there are two approaches, the Backward Propagation (BP) algorithm and the multilayer feed forward NN. Normally the optimization problem is usually handled using the Backward Propagation (BP) algorithm in which the error evaluated at output layer is propagated back through the hidden layers (Chen et al., 2003). However, the common NN structure is the multilayer feed forward NN, which is proven to be a universal approximator of non-linearity. Here, the outputs of nodes in one layer are inputs to the next layer. At each node, NN modifies the inputs, using the logistic function. $\Upsilon(Z) = \frac{1}{1+e-Z}$ where Z is the weighted linear combination. The

functional form[9] of such model is as follow: $Z_j = \beta_j + \sum_{i=1}^{n} w_{i,j} X_i$. The parameters $\beta_1, ...\beta_j$ and $w_{1,2}, ..., w_{i,j}$, are learned from the training set.

4.3.4 Joint modeling

In addition to these prediction techniques, we also used a joint modeling technique consisting of two sub-models to cross check the previously obtained results. The use of this latter is again justified by the fact that, the BIS (despite its covariate nature) was measured over time on the same subjects (assuming a longitudinal scheme for the continuous data BIS). As such, measurement time (Hour) could have an unknown influence on the predictive ability of BIS. In such case, more complicated methods could be used. For instance, a summary measure that collapses the longitudinal information over time (e.g. overall average, change score, slope, intercept, maximum, minimum, achievement of a threshold, measurement on a particular day, etc.) could be used as a covariate in a logistic regression. However, it may be difficult to decide which summary measure to use, and as there are many candidate measures, there is perhaps a problem of multiple comparisons. Furthermore, if the longitudinal measurements are non-linear in time, it may be difficult to come up with a single summary measure. Generalized Estimating Equations (GEE) can be used to estimate the effect of a time-varying covariate on a time-varying multi-dimensional response; in our application however, the response has a single dimension, so GEE is not appropriate. Hence, the fitted joint model consist of:

- a first sub-model which is a linear mixed model with BIS as continuous response variable and time as covariate and patient ID as random component;
- a second sub-model which is a Generalized Linear Model (GLM) for the primary outcome (which in our application is a binary outcome predicting poor neurological outcome), using the random intercept and slope obtained from the first stage as covariate, in addition to all other demographic and clinical variables previously used in the classification models (see Laird and Ware 1982).

4.3.5 The first sub-model: Linear Mixed Model

Mixed modeling has become a major area of statistical research, including work on computation of maximum likelihood estimates, non-linear mixed effect models and missing data in mixed effects models. Mixed models are applied in many disciplines where multiple correlated measurements are made on each unit of interest or where measurements are made on clusters (here time point measurement) of related statistical units (here patient ID). They are prominently used in research involving human and animal subjects in fields ranging from medical research to business administration research (marketing) fields, and have also been used in industrial statistics.
As with all statistical problems, the method to use in data analysis depends on the type of data at hand. Perhaps the easiest data to work with due to the availability of extensive methods for analysis, especially in software implementation, is the continuous type of data. This stemmed from the fact that with continuous data, even if not normally distributed, transformations to attain normality are available and henceforth, the elegant

properties of the normal distribution can be used. A nice property also of normally distributed data is that, integrating the mixed model over the random effects produces a marginal model. As such, regression parameter estimates of the Linear Mixed Model have marginal interpretation and the random effects contribute in a simple way to the variance-covariance structure (Verbeke and Molenberghs, 2000).
Let $Y_{ij} = (Y_{i1}, Y_{i2}, ..., Y_{im_i})$ represent the continuous longitudinal measurement (BIS) for individual i, i= 1,2,....,n ($n = 77$), at measurement times (Hour) $X_{ij} = x_{i1}, ..., x_{im_i}$, $m_i = 36$. Individuals were measured approximately at the same time. The sub-model for the longitudinal data can be written as:

$$Y_{ij} = \alpha_0 + \alpha_1 X_{ij} + Z_i + \epsilon_{ij}$$

Where, α_0 is the overall mean BIS value at time ($Hour = 1$), α_1 is the effect of time on the BIS value. Z_i is a q x 1 vector of unknown random-effects and $\epsilon_{ij} = (\epsilon_{i1}, ..., \epsilon_{im_i})$ is a vector of measurement errors for individual i. We assume that:

$$Z_i \sim N(0, \Sigma) \quad and \quad \epsilon_{ij} \sim N(0, \sigma^2 I_{m_i})$$

and that Z_1,.....,Z_n, ϵ_1,....., ϵ_n are independent, where Σ is a q x q covariance matrix for the random effects and I_k indicates the k x k identity matrix. For example, letting the j^{th} row of X_i equal (1,x_{ij}) for $j = 1, ..., m_i$, so that $Z_i = (Z_{i1}, Z_{i2})^T$ corresponding to the random intercept and slope for subject i, and with:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

we obtain a random slopes and intercepts model introduced by Laird and Ware (1982) as sub model in a joint modeling.

4.3.6 The second sub model: Generalized Linear Model

The sub model for the primary endpoint (neurological outcome), R_i, is assumed to be a Generalized Linear Model (GLM) as described previously.
The second applied GLM, fitted a logistic regression model, using the random intercept and slope obtained from the first sub-model, added as covariates to all other relevant independents variables.
Logistic regression models are the most prominent, from the numerous distribution functions proposed for use in the analysis of a dichotomous outcome variable from continuous and/or discrete explanatory variables measured at a single point in time (Hosmer Jr, Lemeshow, and Sturdivant, 2013). The rational for this choice stand in two points, a mathematical point of view, as it is an extremely flexible and easily used function, and an easy clinically meaningful interpretation. More so, logistic regression easily accommodates the binary nature of the response. On the other hand, it uses the method of maximum likelihood for parameter estimation which has better statistical properties (James et al., 2013). For notation simplicity, $\pi(x) = E(R|x)$ which is the conditional mean of the Neurological outcome given patients characteristics for a logistic function.

Precisely, it expressed the probability that individual i died (non survivor) given to his BIS value and other characteristics. The distribution of R_i is a binary indicator in our application, assumed to belong to the exponential family of distributions. The logistic regression makes use of logistic function that takes only two values (0 and 1). It is a classification model of two categories response. As such, and likewise all classification models, it predicts the neurological outcome using one or more independent variables. The independent variables can be either categorical or numerical. Replacing the dependent variable of the logistic function with a linear combination of dependent variables we intend to use for regression, we arrive at the formula for logistic regression as in the following equation:

$$logit(p_i) = \beta_0 + X\beta + Z\gamma$$

Where, $x_1, ..., x_p$ are the p patients characteristics in the model, $Z\gamma$, the Random intercept and Random slope estimates obtained from the Linear Mixed Effect Model and $\beta_0, \beta_1, ..., \beta_p$ are unknown parameters which are estimated based on the training data. We can interpret $\pi(x)$ as the probability $p(R = 1|x)$. That is, the probability of a given neurological outcome, given the patients BIS and other characteristics.

This logistic regression fits a maximum likelihood model to the data. That is a model that supplies probabilities for each patient and the product of all the predicted probabilities is least surprising (close as possible to their true probabilities of neurological outcome) (Friedman, Hastie, and Tibshirani, 2001; Kutner et al., 2005). The maximum likelihood model fitted to the data is of form:

$$\pi(x) = \frac{1}{1 + e^{-}(\beta^T x + z\gamma)}$$

Where, $\beta^T x = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$, L($\beta$) is the likelihood of neurological outcome, and $p(X_i) = p(R_i = 1|X_i)$ is the probability of a patient neurological outcome given patient characteristic.
Since the neurological outcome predicted probabilities value range from 0 to 1 and in the validation data is binary, The predicted probabilities had to be dichotomized in order to get the prediction accuracy as in the formula:

$$\hat{R_i} = \begin{cases} 0, & \text{if predicted probability} \leq 0.5 \\ 1, & \text{if otherwise} \end{cases}$$

Where, $\hat{R_i}$ is the Neurological Outcome prediction of the i^{th} patient in the train data. If the predicted probability is greater than 0.5, the patient is classified in the class of those having a neurological outcome (1, non-survivor) otherwise the patient is classified as survivor (0, survivor). The class predictions were cross-tabulated with the gold standard in a confusion matrix from which the false positive rate, false negative rate, sensitivity, specificity, and predictive values are computed.

4.4 Software

The statistical software used for the statistical analysis of the current data are R 3.4.0, with different R-packages appropriated to each procedure and SAS 9.4.

5 Data analysis and interpretation of results

5.1 Exploratory data analysis

As a first step of the analysis, the data were explored in different ways in order to get details that may help to make decision in the subsequent steps of the analysis.

5.1.1 Descriptive statistics

The table below present the Summary statistics at baseline per group (survivor and Non-survivor) for continuous covariates Age, Initial Lactate and Initial pH.

Table 1: Summary statistics at baseline per group

Variable	Mean	Median	Mode	Std Dev	Minimum	Maximum
Outcome = 1 (Non-survivors)						
Age	67	67	61	12.97	38	89
Initial_lactate	6.55	5.85	2.5	3.9	1.35	18.5
Initial_pH	7.2	7.2	7.17	0.11	6.91	7.43
Outcome = 0 (Survivors)						
Age	60.58	61	53	12.74	28	91
Initial_lactate	5.53	4.55	5	4.29	0.9	20
Initial_pH	7.24	7.26	7.26	0.16	6.74	7.62

The study include 77 patients. The outcome is composed of 38 survivors and 39 non survivors.
The mean survivors group age is 60.58 against 67. At a first stage, the analysis, was performed using a complete deletion case of missingness pattern handling. Hence, patients with missing values were discarded from the analysis that were performed on the evolution of BIS values in the remaining 50 patients (24 survivors and 26 non-survivors) during the first 36 hours of their stay.

5.1.2 Individual Profiles

The score of bispectral index was recorded every hour. All over the study period, the recorded scores were higher and relatively stable in survivors group as compared to non survivors'.

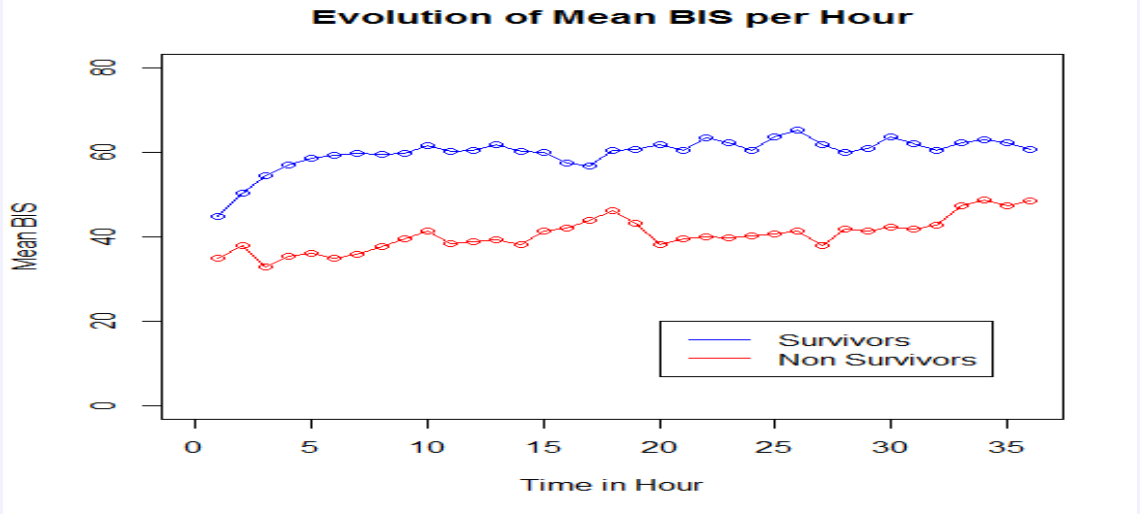

Figure 5: Mean profile plot

In another hand, the variability of BIS values was lower in survivors patients as compared those with poor neurological outcome (figures 4 and 5).

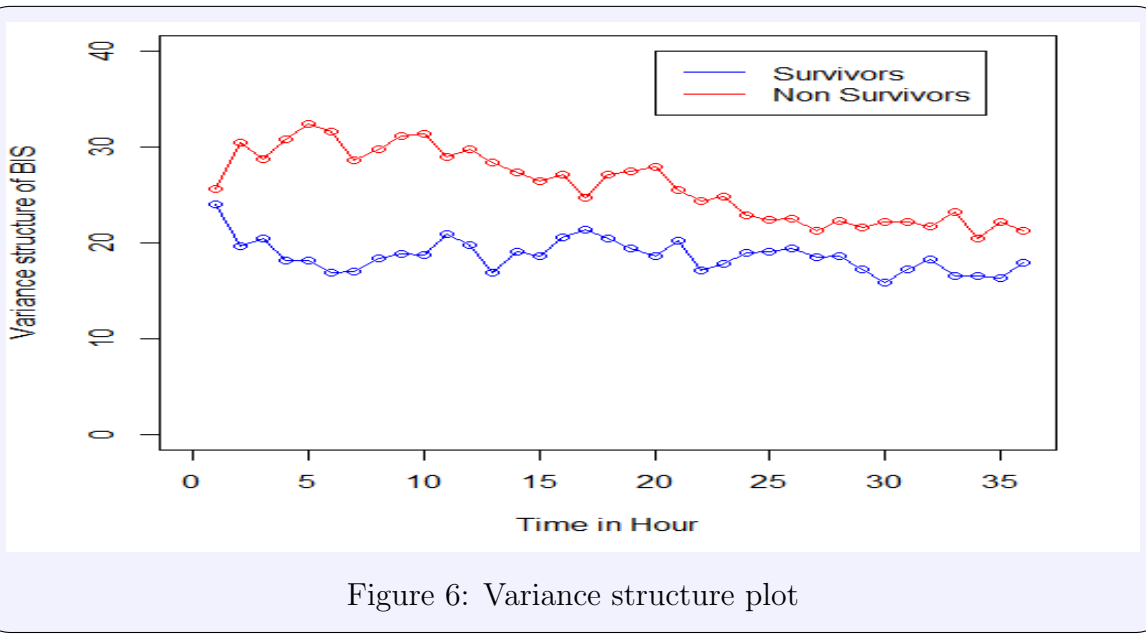

Figure 6: Variance structure plot

5.1.3 Correlation analysis

The data analysis show a high correlation between BIS and EMG (0.59) and SR (-0.91) as presented in the correlation matrix at baseline. The relevance of BIS as compared to over covariates allows us to discard EMG and SR from a model that contain already, BIS.

Table 1: Correlation coefficients (P-value)

	Outcome	**BIS.1**	**EMG.1**	**SR.1**	**NSE24.1**	**NSE48.1**
Outcome	1.00	-0.20(0.11)	0.12(0.35)	0.37(0.00)	0.22(0.05)	0.12(0.29)
BIS.1		1.00	0.59(0.00)	-0.91(0.00)	-0.21(0.09)	-0.12(0.33)
EMG.1			1.00	-0.39(0.00)	0.02(0.86)	0.06(0.64)
SR.1				1.00	0.21(0.08)	0.15(0.21)
NSE24.1					1.00	0.70(0.00)
NSE48.1						1.00

5.1.4 Data Pattern

The first plot, is a histogram which clearly depicts the influence of missingness values in each variable of the dataset.

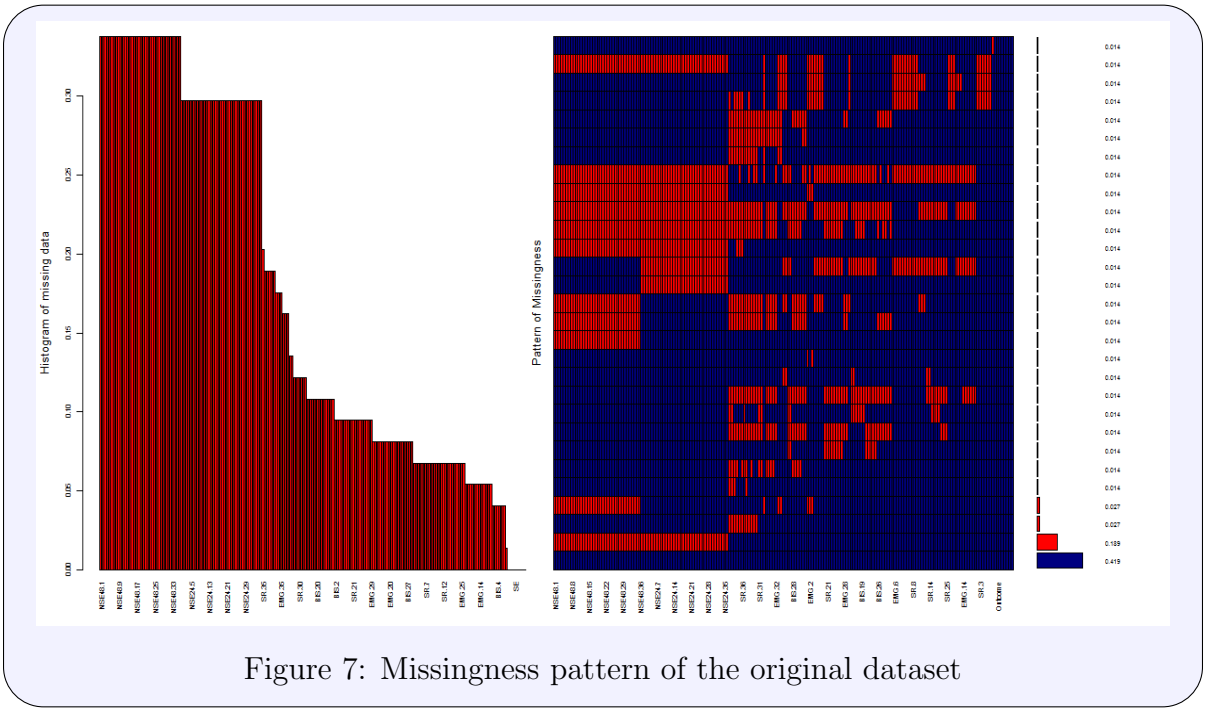

Figure 7: Missingness pattern of the original dataset

The second figure shows the missingness pattern in the dataset. Each cell represents an observation. The red color indicates the missing observation per variable. Hence, the first observation which is the first row from below, do not have any red color, indicating that there were no missing value. However, the overall pattern shows the presence of a lot of missing values. Indeed, there are 41.9% values in the data set with no missing value. There are 18.9% missing values in NSE24 and NSE48, 2.7% missing values in EMG, and in BIS. A proportion of around 1.4% of missing values was found in the remaining variables. Therefore, there is a need of missing handling technique application in the current study.

This figure present the missing pattern in the dataset obtained from the imputation process ran at 5000 iterations.

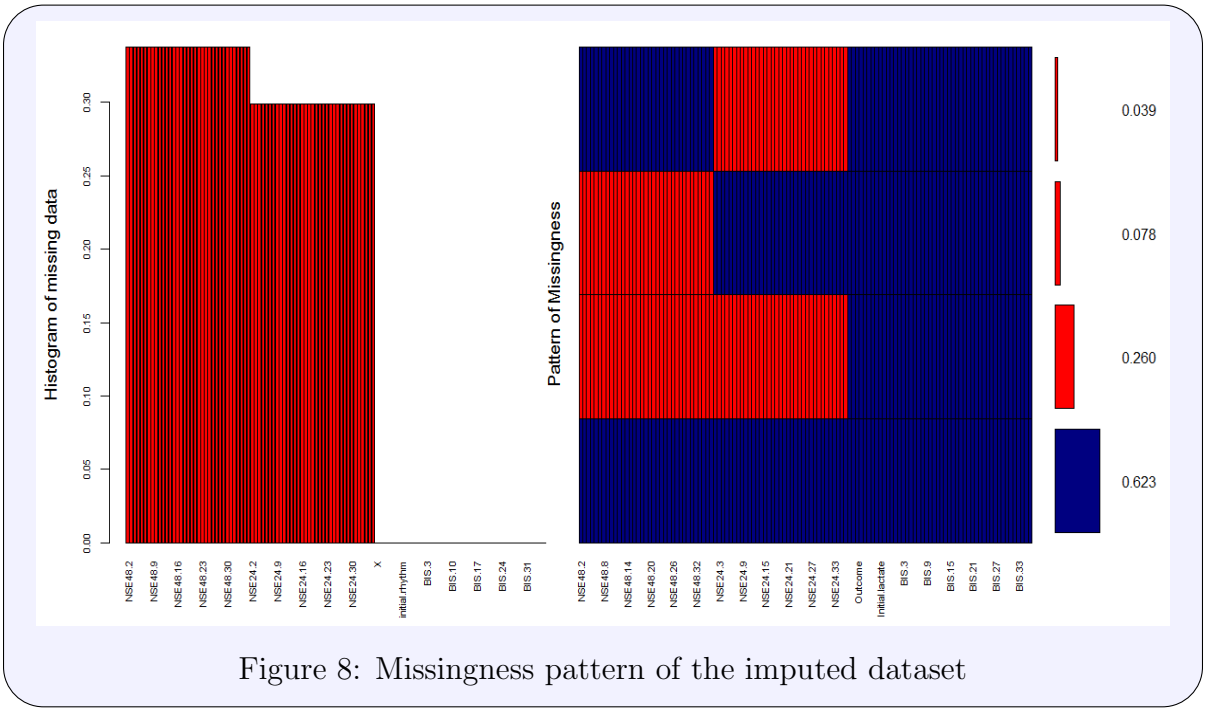

Figure 8: Missingness pattern of the imputed dataset

After 5000 iterations with 5 maxit, we obtain a dataset containing some missing values for the two variables NSE24 and NSE48. A little comparison of the figure 7 and 8 show that the imputation process allow us to have more completed dataset. This will be confirmed by the figure 9 that presents the convergence plot of the dataset obtained from the imputation process. This plot shows quite an acceptable convergence of the data, leading for less deviation from the observed values, thus less bias.

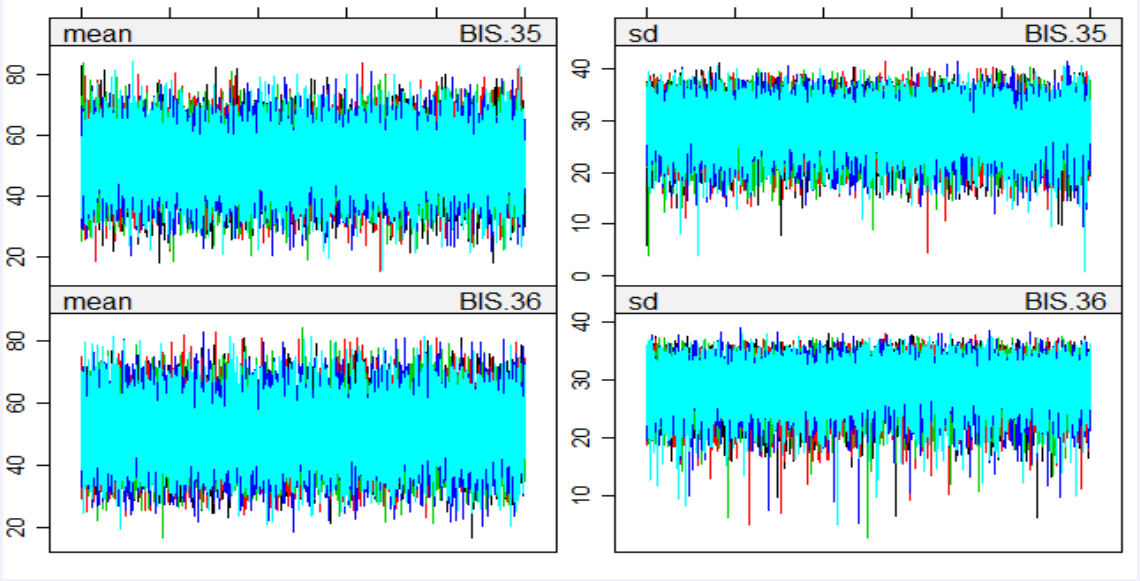

Figure 9: Convergence plot at 5000 iterations and 5 maxit

5.2 Predictive Models

This first part of the current report, was conducted using a complete case analysis missingness handling technique. A second sensitivity analysis was done using the fifth dataset obtained from the multiple imputation process. Discrimination measures used are accuracy, specificity and sensitivity obtained from the test set prediction. The most relevant right after Accuracy that give the global performance of the model, is Specificity since it allow us to predict rightly the poor neurologic outcome of OHCA patients. The R function confusion- Matrix() in the caret library was used to compute these values ([4], [8], [15]).

5.2.1 Support Vector Machine results

Using SVM procedure, for a model fitted with BIS alone as covariate, the first time response appears to be hour 27 with accuracy of 90.91%, a specificity of 85.7% and a sensitivity of 100%. But at hour 3, we recorded a performance of 81.82% accuracy, for the same Specificity (85.7%) while sacrificing a little sensitivity (75%). One could also noticed that stopping treatment after only three hours could be too short for an so important decision and hence, check on the next best performance that raised at Hour 21 with the same metrics values as Hour 3 (table 6).
Using SVM procedure, for a model fitted with BIS and all other covariates, the first time response appears to be hours 6 and 11 with accuracy of 81.82%, a specificity of 85.7% and a sensitivity of 75%. Hour 27 comes as second, with the same values for accuracy, specificity 71.4% and sensitivity 100% (table 7).
However, other techniques had been conducted to appreciate the already obtained results.

5.2.2 Linear discriminant analysis results

Using LDA procedure, for a model fitted with BIS alone as covariate, the first time response appears to be hour 2 with accuracy of 63.64%, a specificity of 71.5% and a sensitivity of 50%. Hour 12 comes as second, with 54.55% as accuracy, specificity 85.7% and sensitivity 0%, followed by hour 17 which specificity is a bit lower but with a little gain in term of sensitivity as shown by table 8.
Using LDA procedure, for a model fitted with BIS and all other covariates, the first time response appears to be hour 27 with accuracy of 72.73%, a specificity of 71.4% and a sensitivity of 75% as one can see from the table 9.
The linear discriminant analysis technique conducted gives quite mitigated results. Thus, for a model including only BIS, the highest meaningful accuracy was obtained at hours 12 (54.55%) and 18 (54.6%) against a model including all other covariates at hour 27 (72.73%). However, it comes to confirm the prediction of hour 27 obtained from the SVM method.
Hence, including other covariates in the model help improving the prediction accuracy.

5.2.3 Neural Network results

As a third classification procedure conducted in this study, the Neural Network allows the appreciation of the predictability faculty of BIS for neurological status of patients.
Using Neural Network procedure, for a model fitted with BIS alone as covariate, the first time response appears to be hour 27 with accuracy of 90.91%, a specificity of 85.7% and a sensitivity of 100%. Hour 6 comes next in terms of performance, with an accuracy of 81.82%, a very good specificity of 100% but a sensitivity reduced by half (50%) as compared to the one of the best time predicted by the model (hour 27) as presented in table 10. One should noticed that hour 27 is still confirmed here also, but disappeared in the full model, which results are presented in table 11, where hour 13 appears as the best time predicted by that model followed by hours 3, 6 and 21 already predicted by the SVM models among the five first time response.
Using Neural Network procedure, for a model fitted with BIS and all other covariates, the first time response appears to be hour 13 with accuracy of 90.91%, a specificity of 100% and a sensitivity of 75.0%.

5.2.4 Joint Modeling with logistic regression

The Joint model fitted on BIS alone, make use of the random effect component intercept and slope estimates obtained from the linear mixed effect model as BIS values in the logistic regression model fitted on BIS alone. The fitted model helps to predict the neurological status of the patients from the test set with 63.6% accuracy, 85.7% specificity and 25% sensitivity at time 17 as first time (table 12).
From the joint model fitted using the logistic regression model as second sub model on BIS and all covariates, it had appeared that, hour 17 is still the best with an accuracy of 81.8%, a specificity of 100% and a sensitivity of 75%. Then, follows hour 13 already predicted by the Neural Network model but with less percentage of accuracy and specificity but a better sensitivity (75%) as presented in table 13.
Thus, adding other covariates improves the accuracy of the prediction method.
Hence, combined with the Generalized Linear Model (GLM), we obtained more accurate results with high significance for the mean BIS (intercept of the random component) for hours 1 to 17. Indeed, BIS appears significant with a P-value of 2.7% less than 5% significance level. All others covariates in the model appear non significant except Age that while forcing, in order to explain the source of the gain in accuracy we could indicated it as significance at 10% level. This result is summarized in table 2.

Table 2: Analysis of Maximum Likelihood Estimates of full model fitted over 17 hours

Parameter	**Estimate**	**Standard Error**	**Wald Chi-Square**	$Pr > ChiSq$
Intercept	56.0771	39.6568	1.9996	0.1573
Rand. intercept	-0.0498	0.0225	4.9039	0.0268
Rand. Slope	-0.4878	0.3601	1.8344	0.1756
Gender	-0.549	1.376	0.1592	0.6899
Age	0.0557	0.0315	3.1283	0.0769
initial_rhythm	-0.0077	0.0298	0.0667	0.7961
Initial_lactate	-0.0306	0.148	0.0429	0.8359
Initial_pH	-8.1641	5.4194	2.2694	0.1319

Thus, Age appears as significant covariates at 10%, indicating that a patient's poor neurological outcome is age related. From the negative sign of the parameter estimate, we can say that old patients have higher probability to decease.
The regression methods, logistic regression, did well as compared to LDA. This is most likely due to their ability to handle nonlinear class boundaries.

5.2.5 Joint Modeling with Neural Network

The joint modeling technique allows us to appreciate given to time spent before waking up or death of the patient, the effect of the bispectral index score (BIS) on his neurological status.
The final model retained for the analysis is a joint model build with a linear mixed effect model as first sub-model to handle time varying covariate (BIS) and then the best classifier identify from the model evaluation process as second sub-model. While fitting such joint model, the result appears quite convincing. Hence, trained on the training set, this model correctly classifies 34 patients over 39 with an accuracy of 87.18% while the test set prediction gives an accuracy of 72.73% showing that the model is stable since the difference between these two accuracy is 14.45%, less than 20%. The prediction capability of the model had improved with tune function.
Indeed, with the parameter tuning of 'nnet', the 10-fold cross validation sampling method gives size 4, decay- 0.0015 as best parameters and a best performance of 0.3220708 as error.
Under all those consideration, we found that the fitted model is more accurate than all previous technique used. Thus for BIS alone, the accuracy obtained is 90.91%, with a specificity of 100% and a sensitivity of 75% at time 9 (Hour at which the patient died or wake up). The classification by the model prediction defaulted by 1 misclassified patients over 11 (table3).

With the parameter tuning of 'nnet', the 10-fold cross validation sampling method gives size 5, decay- 0.0015 as best parameters and a best performance of 0.2843488 as error, while fitting BIS (Intercept and slope values obtained from the linear mixed model with all other demographic and clinical characteristics covariates, the model get the most best performance that could be expected in terms of prediction accuracy that rise to 100%, specificity, (100%) and sensitivity (100%) at hour 24. The model rightly predicted correctly all patients into their observed categories. Then predicted hour 12 as second best performance as one can see from table 15.

Table 3: Confusion matrix from JM-NN with BIS alone at hour 9

		Neurologic Outcome		**Total**
		Non Survivor	**Survivor**	
Predicted	**Non Survivor**	3	0	3
	Survivor	1	7	8
Total		4	7	**11**

5.3 Evaluation of the best prediction method

5.3.1 BIS alone

While assessing prediction methods performance on the model fitted with BIS alone, we noticed that Joint model with NN as sub-model recorded the first high level performance followed by NN and SVM that recorded a test set accuracy value of 0.9091. The joint model fitted with logistic regression as sub model comes next (0.636) and finally, the Linear Discriminant Analysis recorded 0.5455 (table 4).

5.3.2 BIS with all other covariates

The prediction methods performance on the model fitted with BIS and all other covariates show clearly a better performance for JM-NN, followed by NN then SVM in third position, JM-LR in fourth rank and finally Linear Discriminant analysis in last. Thus, driving all other methods, JM-NN produced a prediction with 100% accuracy followed by NN (90.91%), SVM (81.82%), then joint model of Logistic regression recorded 81.8% and lastly, LDA (72.73%). In terms of Specificity values, this order is quite similar with the JM-LR coming right after NN followed by SVM and ended by LDA as summarized in table 4. The late metrics values gave the incentive of choosing NN as second sub-model instead of the Logistic regression in order tho make use of the good performance recorded from NN on our data, while accounting for time influence on BIS values.

Table 4: Table Summarizing Method's Evaluation

Classifier	Models	Sensitivity	Specificity	Accuracy	Kappa	Time	Best classifier
SVM	**BIS**	1	0.857	0.9091		27	2ex
	Full mod.	0.75	0.857	0.8182	0.6071	11	3
LDA	**BIS**	0	0.857	0.5455		12	5
	Full mod.	0.75	0.714	0.7273	0.4407	27	5
NN	**BIS**	1	0.857	0.9091		27	2
	Full mod.	0.75	1	0.9091	0.7925	13	2
JM LR	**BIS**	0.25	0.857	0.636		17	4
	Full mod.	0.5	1	0.818		17	4
JM NN	**BIS**	0.75	1	0.9091	0.7925	9	1
	Full mod.	1	1	1	1	24	1

From all above, we noticed that hour 27 appears more frequently and could be taken as the best compromise from all three classification techniques (SVM, LDA and NN), for a model of only BIS and a model including all other covariates. However it does not show off at all in the two joint models fitted.
From table 4, while assessing the best prediction model, it appears relevant to check on the performance of a JM-NN, using a Linear Mixed Effect Model as first sub-model and

Neural Network as second sub-model to predict poor neurological outcome of the patients involved in this study.
Applying JM-NN technique appears moreover relevant since SVM and NN predicted hour 27 as the first time while using BIS alone as covariate, but gives different others hours (11 for SVM and 13 for NN) in their full models while JM-LR predicted hour 17 consistently from the two models BIS alone and full model one could expect such consistency from JM-NN.

5.4 Sensitivity analysis

In order to assess the sensitivity of the different results obtained we conducted the same process using the fifth dataset obtained from the MICE multiple imputation. This data set allow the usage of all 77 patients, randomly split with the same ratio (80%: 20%) in training set of size 62 and test set of size 15 patients.

5.4.1 Sensitivity analysis: SVM, LDA, NN and the two Joint models

The model fitted on BIS alone using SVM method gives hour 7 as best test set accuracy (80%) and a specificity of 85.7%, followed by hour 6 with same accuracy but specificity decreases to 71.4%. These results are presented in table 16 in appendices.
In table 17, the results of the SVM technique obtained from the full model are presented. This place the best response hour at 23 with an accuracy of 73.33%. It also show hour 7 already predicted by the model fitted using BIS alone. but do not confirm none of the time observed from the complete case.
A model fitted with BIS alone as covariate while applying linear discriminant analysis technique gives hour 17 as the best response time with an accuracy of 80%, followed by hours 16 and 12 with less accuracy value but same specificity of 71.4%. This method confirm actually hours 12 and 17 with a little improvement of accuracy and specificity values. hence, for the same specificity value, one could sacrifice 0.2 point of base in accuracy value, to gain five hours prediction while choosing to stop OHCA care at hour 12 (table 18).
Using the imputed data, we noticed that the full model with Neural Network classifier method gives hour 35 as best response time, which could be seen as a very late response time, while BIS alone in table 20, gives an very early response time of hour 7 at the same accuracy level, but perform less in specificity which drop at 85.7%. The full model confirm hours 6 and 21 predicted previously in the complete case part. This comes out with less accuracy value measured on the test set but better specificity value (100%) for hour 21 these can be seen in table 21 in appendices.

5.4.2 Sensitivity analysis for the joint models

The table 22 presents the sensitivity analysis result of JM-LR fitted on BIS alone. This gives hour 3 as the best response time with an accuracy of 60%, confirmed by the full model. None of the models fitted with this method confirm hour 17 previously predicted in complete case. Across methods, hours 7 and 21 are still showing off but this high changes could inspired a need of more data.
From table 24 which presents the sensitivity analysis result of JM-NN fitted on BIS alone hour 21 appears as the best response time with an accuracy of 73.33%, followed by hour 24.

5.4.3 Sensitivity analysis: Method evaluation

From the table 5 below it can be seen that Joint Model with does no longer appear as the best classifier. The best classier from the sensitivity analysis predictions is Neural

Network followed by SVM, then LDA followed by JM-NN and JM-LR comes last. The order seen in complete case stayed for the three classifiers (Neural Network is still the best method followed by Support Vector Machine and Linear Discriminant Analysis). However using Neural Network as second sub-model here does not improve the performances as obtained from the complete case. Hence as our data set is becoming more informative, the effect of time get weak, by having less influence on the prediction performances. Hour 7 in this part of the work appears more frequently while make an across method consideration, same as hour 27 obtained previously in the complete case.

Table 5: Method evaluation table using imputed data

Classifier	Models	Sensitivity	Specificity	Accuracy	Kappa	Time	Best classifier
SVM	**BIS**	0.75	0.857	0.8	0.6018	7	1ex
	Full mod.	0.5	1	0.7333	0.4828	23	2
LDA	**BIS**	0.875	0.714	0.8	0.5946	17	3
	Full mod.	0.75	0.714	0.7333	0.4643	7	3
NN	**BIS**	0.75	0.857	0.8	0.6018	7	1
	Full mod.	0.625	1	0.8	0.6087	35	1
JM LR	**BIS**	0.625	0.29	0.533		7	5
	Full mod.	0.375	0.571	0.467		8	5
JM NN	**BIS**	0.75	0.714	0.7333	0.4643	21	4
	Full mod.	0.625	0.714	0.6667	0.3363	3	4

6 Discussion, Conclusion and recommendation

From all above, we noticed a lot of changes in the results obtained from complete case and sensitivity analysis, leading to a need of more data. A lot of changes in the results produced by the different classification methods used (SVM, LDA, NN, JM-LR and JM-NN) and by models fitted with BIS alone and the full model.

Although, some similarities appear usually, across methods, across models.

The three classifiers NN, SVM and LDA predicted Hour 27 as best time with high level of 90.91% accuracy and 85.7% specificity. While the prediction results obtained from the imputed dataset mostly predicted hour 7 as the best time, with less accuracy value (80%) but the same specificity value of 85.7%. When saying the earlier the better, one could just conclude on hour 7 to stop care for OHCA patients however a caution most be held.
Across complete case and imputed dataset consideration, the different performances recorded putted Neural Network and the Joint model with the Neural Network as the best techniques to predict using BIS the neurological status of Out-hospital cardiac arrest patients.
In this regard, we can conclude this discussion, based on the JM-NN results, since hour 24 appears with all metrics values at 100% which place it as the best time response. However, the prediction of hour 9 with less than 20% loss of accuracy while keeping specificity high at 100%, could be also seen as a very good result. Moreover, hour 9 had been already predicted by the model fitted using BIS alone as best time as presented in table 4. This lead to an earlier prediction of poor neurologic outcome, based on BIS values measured continuously on the same patients.

The prediction result for hour 9 while using BIS alone shows a good compromising between metrics values. The rate of specificity is stable at 100% in the reduced and full model (tables 14 and 15).

Hence, we can use BIS values measured continuously over time, to predict poor neurologic outcome for OHCA patients. As recommendation, which also account for limitation for the present work, one could take into account for next study a larger sample size, adding additional clinical relevant covariates such as BIS of 0, Status Epilepticus (SE), NSE 24 and NSE 48.

References

[1] Box, George EP (1979). “All models are wrong, but some are useful”. In: Robustness in Statistics, p. 202.

[2] Brooks S. C. & al., (2009). ”Out-of-Hospital Cardiac Arrest Frequency and Survival: Evidence for Temporal Variability”, [Online]. 81(2): 175. doi:10.1016/j.resuscitation.2009.10.021., 16. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2815250/pdf/nihms161643.pdf [Accessed 15 August 2017].

[3] Burjek, N. E., Wagner, C. E., Hollenbeck, R. D., Wang, L., Yu, C., McPherson, J. A., & Billings, F. T. (2014). ”Early bispectral index and sedation requirements during therapeutic hypothermia predict neurologic recovery following cardiac arrest.”, Critical Care Medicine, 42(5), 1204–1212. http://doi.org/10.1097/CCM.0000000000000126.

[4] DeLong, Elizabeth R, David M DeLong, and Daniel L Clarke-Pearson (1988). “Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach”. In: Biometrics, pp. 837–845.

[5] E. Osuna, R. Freund, and F. Girosi., (1997). ”Training support vector machines: An application to face detection.” In Proceedings of CVPR’97, Puerto Rico.

[6] Engdahl J., Holmberg M. & all, (2002), ”The epidemiology of out-of-hospital ‘sudden’ cardiac arrest”.

[7] Fisher, R. A. (1936). ”The Use of Multiple Measurements in Taxonomic Problems”. Annals of Eugenics. 7 (2): 179–188.

[8] Hanley, James A and Barbara J McNeil (1983). “A method of comparing the areas under receiver operating characteristic curves derived from the same cases.” In: Radiology 148.3, pp. 839–843.

[9] Hastie T., Tibshirani R., and Friedman J., (2016). “The elements of statistical learning”, Springer.

[10] Heart.org. (2018). ”About Cardiac Arrest”. [online] Available at: http://www.heart.org/HEARTORG/Conditions/More/CardiacArrest/About-Cardiac Arrest UCM 307905 Article.jsp.WlxOrlnkTIV [Accessed 15 Jan. 2018].

[11] JESPER S., (2001). ”Support Vector Machines for Phoneme Classification. Master of Science School of Artificial Intelligence Division of Informatics University of Edinburgh”.

[12] Joachims., T., (1998), ”Text categorization with support vector machines: learning with many relevant features.” In Proc. 10th European Conference on Machine Learning ECML-98, pages 137–142.

[13] Kaski, Samuel and Teuvo Kohonen (1996). “Exploratory data analysis by the self-organizing map: Structures of welfare and poverty in the world”. In: Neural networks in financial engineering. Proceedings of the third international conference on neural networks in the capital markets.

[14] Müller, K.-R., Smola A., R¨atsch G., Sch¨olkopf B., Kohlmorgen J., and Vapnik. V., (1999), ”Predicting time series with support vector machines.” In B. Sch¨olkopf, C. J. C. Burges, and A. J. Smola, editors, Advances in Kernel Methods — Support Vector Learning, pages 243–254. MIT Press.

[15] Kuhn, Max (2015). “A Short Introduction to the caret Package”. In: R Project website. cran. r-project. org/web/packages/ caret/ vignettes/ caret. pdf. Published August 6.

[16] McLachlan, G. J. (2004). ”Discriminant Analysis and Statistical Pattern Recognition. Wiley Interscience”

[17] M. Schmidt and H. Gish., ”Speaker identification via support vector classifiers.” In ICASSP, volume 1, 1996.

[18] MEYER D., (2017). ”Support vector machines: the interface to libsvm in package e1071.” FH Technikum Wien, Austria.

[19] Molenberghs, G., and Verbeke, G. (2005). ”Models for Discrete Longitudinal Data.” New York: Springer-Verlag.

[20] R. Cooley., (1999). ”Classification of news stories using support vector machines.” In Proc. 16th International Joint Conference on Artificial Intelligence Text Mining Workshop.

[21] Riker RR1, Stone PC Jr, May T, McCrum B, Fraser GL, Seder D., (2012). ”Resuscitation. Initial bispectral index may identify patients who will awaken during therapeutic hypothermia after cardiac arrest: a retrospective pilot study.”, [Online]. Volume 84, Issue 6, Pages 794–797, 4.

[22] Rubin, D.B. (1987). ”Multiple Imputation for Nonresponse in Surveys.” New York: Wiley & Sons.

[23] Seder, D. B., Fraser, G. L., Robbins, T., Libby, L., & Riker, R. R., (2009). ”The bispectral index and suppression ratio are very early predictors of neurological outcome during therapeutic hypothermia after cardiac arrest. Intensive Care Medicine”, 36(2), 281-288. doi:10.1007/s00134-009-1691-1.

[24] Sander Greenland, William D. Finkle, (1995). ”A Critical Look at Methods for Handling Missing Covariates in Epidemiologic Regression Analyses,” American Journal of Epidemiology, Volume 142, Issue 12, Pages 1255–1264

[25] Shah, Anoop D et al. (2014). “Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study”. In: American Journal of Epidemiology 179.6, pp. 764–774.

[26] Wetcher-Hendricks D., (2011). ”Analyzing Quantitative Data: An Introduction for Social Researchers”, p.288.

Appendices

A Complete case: Metrics measures results

Table 6: Five first hours result for the SVM model with BIS alone

Order	SVM BIS	Sensitivity	Specificity	Accuracy	Hour
1	**BIS.27**	1.000	0.857	0.9091	**27**
2	**BIS.30**	1.000	0.857	0.9091	**30**
3	**BIS.3**	0.750	0.857	0.8182	**3**
4	**BIS.21**	0.750	0.857	0.8182	**21**
5	**BIS.22**	0.750	0.857	0.8182	**22**

Table 7: Five first hours result for the SVM full model

Order	SVM	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.6**	0.750	0.857	0.8182	0.6071	**6**
2	**BIS.11**	0.750	0.857	0.8182	0.6071	**11**
3	**BIS.27**	1.000	0.714	0.8182	0.6452	**27**
4	**BIS.28**	1.000	0.714	0.8182	0.6452	**28**
5	**BIS.29**	1.000	0.714	0.8182	0.6452	**29**

Table 8: Five first hours result for the LDA model with BIS alone

Order	LDA BIS	Sensitivity	Specificity	Accuracy	Hour
1	**BIS.2**	0.500	0.714	0.6364	**2**
2	**BIS.12**	0.000	0.857	0.5455	**12**
3	**BIS.1**	0.000	0.857	0.5455	**1**
4	**BIS.18**	0.250	0.714	0.5455	**18**
5	**BIS.17**	0.250	0.714	0.5455	**17**

Table 9: Five first hours result for the LDA full model

Order	LDA	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.27**	0.750	0.714	0.7273	0.4407	**27**
2	**BIS.26**	0.750	0.571	0.6364	0.2903	**26**
3	**BIS.29**	0.750	0.571	0.6364	0.2903	**29**
4	**BIS.30**	0.750	0.571	0.6364	0.2903	**30**
5	**BIS.35**	0.750	0.571	0.6364	0.2903	**35**

Table 10: Five first hours result for the NN model with BIS alone

Order	NN: BIS alone	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.27**	1.000	0.857	0.9091	0.8136	**27**
2	**BIS.30**	1.000	0.857	0.9091	0.8136	**30**
3	**BIS.6**	0.500	1.000	0.8182	0.5600	**6**
4	**BIS.3**	0.750	0.857	0.8182	0.6071	**3**
5	**BIS.29**	0.750	0.857	0.8182	0.6071	**29**

Table 11: Five first hours result for the NN full model

Order	NN	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.13**	0.750	1.000	0.9091	0.7925	**13**
2	**BIS.3**	0.500	1.000	0.8182	0.5600	**3**
3	**BIS.6**	0.500	1.000	0.8182	0.5600	**6**
4	**BIS.7**	0.750	0.857	0.8182	0.6071	**7**
5	**BIS.21**	0.750	0.857	0.8182	0.6071	**21**

Table 12: Joint Logistic regression on complete deletion case for model with BIS alone

Order	LR alone	Sensitivity	Specificity	Accuracy	Hour
1	**BIS.17**	0.250	0.857	0.6360	**17**
2	**BIS.18**	0.250	0.857	0.6360	**18**
3	**BIS.19**	0.250	0.857	0.6360	**19**
4	**BIS.20**	0.250	0.857	0.6360	**20**
5	**BIS.21**	0.250	0.857	0.6360	**21**

Table 13: Joint Logistic regression on complete deletion case for the full model

Order	LR Full	Sensitivity	Specificity	Accuracy	Hour
1	**BIS.36**	1.000	1.000	1.0000	**36**
2	**BIS.17**	0.500	1.000	0.8180	**17**
3	**BIS.13**	0.750	0.714	0.7270	**13**
4	**BIS.28**	0.250	0.714	0.5450	**28**
5	**BIS.34**	0.250	0.714	0.5450	**34**

Table 14: Joint Model with NN: Complete case deleted BIS alone

Order	JNN BIS Alone	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.9**	0.750	1.000	0.9091	0.7925	**9**
2	**BIS.33**	0.750	1.000	0.9091	0.7925	**33**
3	**BIS.34**	0.750	1.000	0.9091	0.7925	**34**
4	**BIS.12**	0.750	0.857	0.8182	0.6071	**12**
5	**BIS.30**	0.750	0.857	0.8182	0.6071	**30**

Table 15: Joint Model using Neural Network with all covariates

Order	JNN Full	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.24**	1.000	1.000	1.0000	1.0000	**24**
2	**BIS.12**	0.750	1.000	0.9091	0.7925	**12**
3	**BIS.13**	0.750	1.000	0.9091	0.7925	**13**
4	**BIS.9**	0.500	1.000	0.8182	0.5600	**9**
5	**BIS.10**	0.500	1.000	0.8182	0.5600	**10**

B Sensitivity analysis: Metrics results obtained

Table 16: Sensitivity analysis for SVM on model with BIS alone

Order	SVM : BIS	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.7**	0.750	0.857	0.8000	0.6018	**7**
2	**BIS.35**	0.750	0.857	0.8000	0.6018	**35**
3	**BIS.6**	0.875	0.714	0.8000	0.5946	**6**
4	**BIS.20**	0.500	1.000	0.7333	0.4828	**20**
5	**BIS.33**	0.750	0.714	0.7333	0.4643	**33**

Table 17: Sensitivity analysis for SVM on full model

Order	SVM	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.23**	0.500	1.000	0.7333	0.4828	**23**
2	**BIS.7**	0.750	0.714	0.7333	0.4673	**7**
3	**BIS.5**	0.500	0.857	0.6667	0.3478	**5**
4	**BIS.33**	0.500	0.857	0.6667	0.3478	**33**
5	**BIS.3**	0.625	0.714	0.6667	0.3363	**3**

Table 18: Sensitivity analysis using LDA for model of BIS alone

Order	LDA : BIS	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.17**	0.875	0.714	0.8000	0.5946	**17**
2	**BIS.16**	0.750	0.714	0.7333	0.4643	**16**
3	**BIS.12**	0.500	0.714	0.6000	0.2105	**12**
4	**BIS.8**	0.625	0.571	0.6000	0.1964	**8**
5	**BIS.11**	0.375	0.714	0.5333	0.0870	**11**

Table 19: Sensitivity analysis using LDA for the full model

Order	LDA	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.7**	0.750	0.714	0.7333	0.4643	**7**
2	**BIS.9**	0.500	0.857	0.6667	0.3478	**9**
3	**BIS.25**	0.500	0.857	0.6667	0.3478	**25**
4	**BIS.32**	0.500	0.857	0.6667	0.3478	**32**
5	**BIS.36**	0.500	0.857	0.6667	0.3478	**36**

Table 20: Sensitivity analysis using NN for the model of BIS alone

Order	NN: BIS	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.7**	0.750	0.857	0.8000	0.6018	**7**
2	**BIS.31**	0.750	0.857	0.8000	0.6018	**31**
3	**BIS.35**	0.750	0.857	0.8000	0.6018	**35**
4	**BIS.20**	0.500	1.000	0.7333	0.4828	**20**
5	**BIS.33**	0.750	0.714	0.7333	0.4643	**33**

Table 21: Sensitivity analysis using NN for the full model

Order	NN	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.35**	0.625	1.000	0.8000	0.6087	**35**
2	**BIS.21**	0.500	1.000	0.7333	0.4828	**21**
3	**BIS.36**	0.500	1.000	0.7333	0.4828	**36**
4	**BIS.6**	0.625	0.857	0.7333	0.4737	**6**
5	**BIS.31**	0.625	0.857	0.7333	0.4737	**31**

Table 22: Sensitivity analysis for joint logistic regression model on BIS alone

Order	LR alone	Sensitivity	Specificity	Accuracy	Hour
1	**BIS.3**	0.625	0.571	0.6000	**3**
2	**BIS.4**	0.625	0.571	0.6000	**4**
3	**BIS.7**	0.625	0.429	0.5330	**7**
4	**BIS.15**	0.500	0.429	0.4670	**15**
5	**BIS.6**	0.375	0.429	0.4000	**6**

Table 23: Sensitivity analysis for full joint logistic regression model

Order	LR Full	Sensitivity	Specificity	Accuracy	Hour
1	**BIS.3**	0.750	0.714	0.7330	**3**
2	**BIS.8**	0.375	0.571	0.4670	**8**
3	**BIS.5**	0.500	0.429	0.4670	**5**
4	**BIS.7**	0.500	0.429	0.4670	**7**
5	**BIS.21**	0.500	0.429	0.4670	**21**

Table 24: Joint model with NN : BIS alone

Order	JNN BIS	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.21**	0.750	0.714	0.7333	0.4643	**21**
2	**BIS.24**	0.750	0.714	0.7333	0.4643	**24**
3	**BIS.22**	0.625	0.714	0.6667	0.3363	**22**
4	**BIS.34**	0.750	0.571	0.6667	0.3243	**34**
5	**BIS.35**	0.750	0.571	0.6667	0.3243	**35**

Table 25: Joint Model with NN: Full model results on sensitivity analysis

Order	JNN Full	Sensitivity	Specificity	Accuracy	Kappa	Hour
1	**BIS.3**	0.625	0.714	0.6667	0.3363	**3**
2	**BIS.16**	0.625	0.714	0.6667	0.3363	**16**
3	**BIS.17**	0.625	0.714	0.6667	0.3363	**17**
4	**BIS.20**	0.625	0.714	0.6667	0.3363	**20**
5	**BIS.21**	0.625	0.714	0.6667	0.3363	**21**