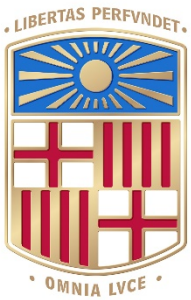

UNIVERSITAT DE BARCELONA

Final Degree Project

Biomedical Engineering Degree

Predicting optimal anesthesia level from propofol and remifentanil concentration: analysis of covariate factors for individualization

Barcelona, 14 of June of 2021

Author: Marta Rey Prieto

Tutor: Dr. Pedro Gambús Cerrillo

External Tutor: Sebastián Jaramillo Selman

ABSTRACT

General anesthesia involves some targeting effects which aim to prevent the patient from suffering against the therapeutic aggression. These effects are hypnosis, analgesia, amnesia and immobility and to achieve them a combination of drugs is delivered into the patient, from which propofol and remifentanil are highlighted.

In the operating room, monitoring systems are used to assess the depth of anesthesia in real time. This monitoring includes basic systems such as arterial blood pressure, oxygenation or electrocardiogram and electroencephalogram derived measures, which are more complex; from this last group, BIS index is a good indicator. Being able to predict the anesthetic depth from a set of input variables could be valuable during the surgery, as it would help the anesthesiologists to prevent adverse effects, and it would help the post-operative recovery.

Knowing this, the aim of this project is to predict the probability to be in the optimal level of anesthesia, which is related to the BIS index. This probability is obtained from the input concentration of propofol and remifentanil, a hypnotic and an analgesic drug respectively, and from the demographic variables such as age, height or gender. To do so, a Logistic Regression model will be built with data from patients undergoing general anesthesia in Cirurgia Major Ambulatòria (CMA) in Hospital Clínic.

ACKNOWLEDGEMENTS

Firstly, I would like to express my appreciation to my tutor and supervisor, Dr. Gambús, for smiling back at me and for his guidance during the whole project. It has been a pleasant experience being offered to work in his group acquiring data. My appreciations also to all the staff of the operating room, surgeons and nurses; they have always been really kind to me and I have learned a lot from them.

I would also like to extend my thanks to Dr. Jaramillo, for having helped me since the beginning of the model building and for his constant advices.

Moreover, I would like to help Joan Altés a Master student working in the same research group, SPEC-M, for giving me advices about the model. And I would also like to thanks to Pep, Irene and Joana, a Biomedicine and Biomedical Engineering students who were doing their TFG and practices in the same operating room; they made my stay there more pleasant.

Finally, I wish to thank my family and friends who have encouraged me to finish the project in the desired schedule.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. General anesthesia

General anesthesia is a drug-induced, reversible condition that includes behavioral and physiological endpoints, such as unconsciousness, amnesia, analgesia and immobility. These effects are induced by anesthetic drugs, highly powerful and with intense side effects such as respiratory depression, cardiovascular instability, altered control of internal temperature and others. The anesthesiologists must adopt measures to maintain physiologic functions inside normality range [1,2]. For instance, to avoid absence of respiration the anesthesiologist must introduce an endotracheal tube to connect the patient to a ventilation machine providing positive pressure ventilation [3].

Drugs used for general anesthesia produces distinct patterns on the electroencephalogram (EEG), from which we can highlight a progressive increase in low-frequency and a high-amplitude activity as the level of anesthesia intensifies [1]. The main characteristics of general anesthesia can be understood by comparing it with sleep and coma. Rapid-eye-movement (REM) sleep and non-REM sleep are the sleep cycles that human present; the EEG shows active high-frequency, low-amplitude and high-amplitude, low frequency rhythms, respectively. On the other hand, coma is a state of profound unresponsiveness, resulting from a severe brain injury. EEG changes induced by general anesthesia are different from coma and from physiologic sleep. Some authors have defined general anesthesia as a reversible drug-induced coma state [1].

The following sections analyze the state of general anesthesia and its components, the drugs required to achieve general anesthesia and the different periods in which anesthesia is divided.

1.1.1. Targeting effects

In the operating room, there a lot of surgical procedures conducted, and the anesthesiologists must generate a state of protection of the patient against the surgical aggression. To achieve that, a combination of some pharmacological effects is required.

Hypnosis

The hypnotic effect gradually starts with some degree of sedation that progresses to unconsciousness [5]. It can be achieved with different kind of drugs, a fact that supports the idea that related to unconsciousness different systems are controlling the same endpoint. Physiologically, it is referred as mind activity characterized by focused attention, dissociation and plastic imagination. During induction of unconsciousness, the patient feels general well-being, a sense of deep relaxation, eyelid heaviness and regular breathing that progressively loses its physiologic pattern, in order to concentrate on the internal environment and disconnect from external stimuli [6]. Several drugs produce hypnosis, from which γ–aminobutyric acid type A (GABA-A) receptor agonists can be highlighted. Propofol is classified into this group.

After induction of anesthesia, the hypnotic effect must be evaluated to avoid it being excessive or insufficient. Once unconsciousness is achieved, some reflexes can no longer be used to assess it, especially when neuromuscular blocking agents are used. So, the most used way to assess the

loss of consciousness, and so the depth of anesthesia and the hypnotic effect, is by analyzing the EEG. [5]

Analgesia

Analgesia is referred as the loss or modulation of pain perception. Analgesia is an important effect to target during the surgical procedure since the intensity of noxious stimulation is very high and unbearable under normal conditions. The pain stimuli activate the μ, k and δ nociceptive receptors, [7] from where the pain pathways translating the noxious stimulation to the cortex are activated. One way to avoid it is by blocking these receptors by using opioids, such as remifentanil, fentanyl or alfentanil.

The analgesia level cannot be directly measured, a problem that is more significant because of unconsciousness. There are commercially available monitors of antinociception, most of them are indirect systems based on the quantification of the variations in tonus of the sympathetic and parasympathetic activity. These monitors include physiological signals, such as heart rate, blood pressure, skin conductance, plethysmograph waveform, pupil diameter and EEG integrated-signals. [5]

Amnesia

Amnesia is the memory suppression during surgery; this is, the impossibility of retain what has occurred since anesthetic state has started and for the whole surgical procedure.

Immobility

Immobility is described as absence of movement during the procedure, related or not to noxious stimulation [9]. Besides adequate analgesia, it can be achieved by the administration of neuromuscular blocking agents, also known as muscle relaxants, which block the connection between nerve and muscles at the endplate level. During the induction of anesthesia, they are mainly used to open the vocal chords and allow the introduction of the endotracheal tube. It might be required also during the course of surgery to facilitate surgical maneuvers [10].

Neuromuscular blockade can be obtained by using depolarizing drugs, such as succinylcholine, and non-depolarizing agents, such as rocuronium. Both of them have the aim to avoid the union of Acetylcholine (Ach) molecules to their postsynaptic receptors to generate the muscle contraction. While the firsts are acetylcholine agonists, the others compete with the Ach to occupy the receptors. [11]

The degree of neuromuscular block can be measured with mechanomyography (MMG) and acceleromyography (AMG) preferably, because they measure the actual movement in muscles, but also with electromyography (EMG). [5]

1.1.2. Anesthetic drugs

Anesthetic agents can be classified in 5 main classes [4]: intravenous (IV) anesthetics, inhalational anesthetics, IV sedatives, synthetic opioids and neuromuscular blocking drugs. Anesthetics can be administered by inhalation but, an alternative method is injecting them exclusively through veins:

total intravenous anesthesia (TIVA). This method is seen as a good alternative to inhaled drugs, as patients who have TIVA has less nausea and vomiting, as well as a better recovery conditions after anesthesia. [5] However, further study is required to determine whether inhaled or intravenous drugs lead to better postoperative cognitive response.

TIVA drugs can be administered either by manual techniques (intravenous bolus or constant continuous infusion) or by target-controlled infusion (TCI). TCI systems use predicted concentrations in plasma (Cp) or to the site of drug effect, also called biophase os effect-site (Ce), as the target controlled by the anesthesiologist. TCI systems use pharmacokinetic-pharmacodynamic (PKPD) models to adjust, every ten seconds, drug infusion to achieve and maintain the target concentration selected by the anesthesiologist. TCI systems are used to deliver the hypnotic and the analgesic agents. The PKPD models are based on population studies correlating drug infusion to blood concentrations and effect by means of effect-site concentrations. [12]. Covariates such as weight, age or gender are used to individualize better the performance of the model and adjust the infusion rate to the characteristics of the patient [13].

The most common TIVA drugs are propofol, remifentanyl, fentanyl, alfentanil and sufentanil; ketamine is sometimes used but it has powerful effects on the brain such as hallucinations as well as a long duration of effects, delaying recovery. [14]

Any combination of hypnotic and opioids can be used in TIVA but, the synergy between propofol and remifentanil for different clinically relevant endpoints is highly effective, and it achieves an adequate depth of anesthesia, considered as hypnosis and analgesia, and a rapid recovery [15]. While specific drug classes target specific clinical end-points, when they are combined, opioids also have sedative effects and propofol also has analgesic properties. [5]

The following table includes some examples of the drugs that can be used and its classification [16]. In this section, the most common anesthetic drugs used in Cirugia Major Ambulatòria (CMA) of Hospital Clínic will be reviewed.

Hypnotics	Opioids	Neuromuscular blockers
Pentothal	Morphine	Succinylcholine (depolarizing)
Etomidate	Fentanyl	Atracurium (non-depolarizing)
Propofol	Remifentanil	Cis Atracurium (non-depolarizing)
Ketamine	Alfentanil	Mivacurium (non-depolarizing)
		Rocuronium (non-depolarizing)
		Vecuronium (non-depolarizing)

***Table 1**: Classification of anesthetic agents*

Propofol

Propofol is a safe, effective, hypnotic, and amnesic anesthetic agent. [17] Its rapid administration produces vasodilatation and decrease of peripheric resistances, but these effects can be avoided if the perfusion is slow or continuous [18]. This drug can lead to respiratory depression, hypotension, decrease of oxygen consumption or decreased cardiac blood flow. At anesthetic doses the level of respiratory depression requires mechanical ventilation. It has an unspecific effect among lipid cell membranes and it acts as an allosteric modulator of $GABA_A$ receptors. It enhances de Cl^- conductivity and the action of the GABA neurotransmitter, which inhibits the synaptic transmission through a hyperpolarization mechanism, opening the Cl^- channels. As it is highly lipophilic, it can reach quickly the central nervous system (CNS), leading to the hypnotic effect. [17]

It is known for its rapid onset of action, its rapid recovery time and its little post-anesthetic sedation and sequalae. Nonetheless, due to its lipidic component and related also to the diameter of the vein in which it is injected it can produce a stinging feeling at injection, which in some cases could be avoided with the administration of lidocaine. Despite this, it is an excellent drug for the ambulatory surgical setting.

Remifentanil

Remifentanil is a synthetic opioid whose main goal is to provide the analgesia [19]. It is a μ- agonist, with low affinity to κ, σ, δ receptors and produces bradycardia and hypotension. Furthermore, it has a synergistic effect with other hypnotic agents delivered intravenously, and enhances the loss of consciousness. As it is an opioid, it can produce a respiratory depression, when used for general anesthesia, mechanical ventilation will be required and, as in the case of propofol, patency of the airway must be achieved by either the placement of a laryngeal mask airway (LMA) or through laryngoscopy and

Its main advantage is its short half-life and rapid clearance, due to the metabolism by unspecific plasmatic and tissue esterases. [20]

Rocuronium

Rocuronium is a non-depolarizing neuromuscular blocker, used to produce muscle relaxation and immobility during surgery. Its use facilitates airway management, and the ability to ventilate and oxygenate the patient. As with many other drugs there are very few cases reported of allergic reactions including anaphylaxis [21]. Its main advantage among other blockers is its rapid acting and its reversibility by anticholinesterase drugs or the specific reversal agent sugammadex [22]

It is not metabolized into active metabolites and it has a low lipid solubility, so it cannot pass lipid membrane barriers, such as the blood-brain barrier as the previous drugs. Hypothermia, hypovolemia, concomitant volatile agents and hepatic and renal diseases prolong the effects of the agent. [22]

1.1.3. Anesthesia stages

When anesthesia appeared, the only way of assessing the patient's depth of anesthesia was by physical examination. It was not until 1937 when a monitoring approach appeared; it was developed

by Dr. Arthur Guedel and it consisted on 4 stages [23]: analgesia or disorientation, excitement or delirium, surgical anesthesia and overdose [24]. The first stage takes place at the preoperative anesthesiology holding area, with patients sedated but conversational and conscious, and the delirium phase includes disinhibition, uncontrolled movements, loss of eyelash reflex, hypertension and tachycardia but with the presence of airway reflexes. Then, the surgical anesthesia is the targeted anesthetic level in general anesthesia, where ceased eye movements and respiratory depression are highlighted. Finally, overdose occurs, proportional to the amount of surgical stimulation, which leads to severe brain or medullary depression; it goes from respiratory cessation to potential death.

Although this classification is still used, due to the arrival of newer anesthetic medications, delivery techniques and monitoring technologies, the stages described by Guedel are not used. General anesthesia is considered as a continuum that can be adapted to the specific responses of the patient by modulating anesthetic drug administration and concomitant drugs depending on the expected changes due to the surgical process or to the specific responses of the individual patient.

For better description, general anesthesia can be divided in 3 other periods: [1]

Induction period

This period begins with the injection of hypnotic drugs, such as propofol, with which sedation is induced. The patient goes from a calm state to unconsciousness although sometimes a slight paradoxical excitation appears. Loss of verbal response to commands is evident. At the same time, opioid administration starts to obtain adequate analgesia to block noxious stimulation during airway management maneouvers. Absence of spontaneous breathing and intense decrease in values of processed EEG parameters indicate that the brain is losing its normal function patterns due to general anesthesia. When more hypnotic and opioid agents being injected, the respiratory pattern becomes irregular and breathing must be supported with manual ventilation and face mask. When adequate hypnotic level and enough analgesia has been achieved, airway should be warranted to allow respiratory function and oxygen administration. This means that either a laryngeal mask or endotracheal tube must be placed. The latter will require a neuromuscular blocking agent.

Maintenance period

During this phase, a combination of hypnotic agents (inhalational or intravenous) and opioids, sometimes with added neuromuscular blocking agents, maintains general anesthesia. Hemodynamic support is often required by means of sympathomimetic drugs. Ventilatory and thermoregulatory support are also required. Besides basic and advanced monitoring technologies, there are many indicators which evaluate the state of anesthesia, such as the EEG to control the hypnotic component and others to control response to nociceptive stimulation (changes in pupil size, heart rate variability, sympathetic activity indicators and also EEG changes…). The changes in these indicators are used by the anesthesiologist to adapt drug administration to individual requirements.

Emergence period

Emergence and recovery from general anesthesia starts when anesthetic delivery stop. The patient gradually recovers protective reflexes such as coughing and ability to breath followed by opening the eyes and sluggish response to commands. It is a passive period that depends on the patient metabolizing and eliminating anesthetic drugs remaining in the body, the physiology of the patient and the duration of the surgery [25]. During this period, a normal pattern of spontaneous ventilation is recovered with adequate respiratory rate and tidal volume, and salivation, tearing and responses to painful stimulation begin. Recovery of breathing means removal of laryngeal mask or extubation. Once stable physiologic patterns are observed, the patient is taken from the operating room to the post anesthesia care unit until everything comes back to normal physiological ranges.

1.2. Anesthesia monitoring

The clinical observation and evaluation by the anesthesiologists are essential to ensure the safety of the patient during the anesthesia. Thus, anesthesia monitoring is a key aspect to be achieved. Due to its importance, monitoring systems are one of the areas that have evolved more during the recent years.

The main objective of monitoring is to observe and to register the temporal evolution of the basic physiologic variables during the different stages of anesthesia and surgery. The equipment must be evaluated prior to the intervention [26]. It is highly recommended to check the oxygenation, ventilation, cardiovascular function, temperature, neuromuscular blocking level and depth of anesthesia [27]. However, the patient must receive a direct vigilance by the anesthesiologist.

Anesthesia monitoring technology can be divided into two groups: basic and advanced monitoring.

1.2.1. Basic monitoring systems

Basic monitoring systems are those which are essential to control vital signs.

1.2.1.1. Pulseoximeter

During the anesthesia, it is required a quantitative monitoring of the oxygen saturation in arterial blood; to do so, a pulseoximeter is used. It is a non-invasive method, based on the placement of a sensor on the fingertip. Light of two different wavelengths passes through the finger of the patient from the emitter to a photodetector and the change of absorbance of each of the wavelengths is measured. Apart from providing the blood oxygen saturation, it can provide indirect data from the cardiovascular state of the patient as well as measuring different variants of hemoglobin levels like carboxyhemoglobin or methemoglobin.

1.2.1.2. Electrocardiogram

The electrocardiogram (ECG) analysis is a way to analyze the cardiovascular function of the patient. It displays the electrical signal coming from the heart and the output parameters are the heart rate, the heart rhythm, conduction, repolarization and asystole. These parameters can give us some insight about the heart rhythm, ischemia and an indirect estimation of the autonomic response to stimulation. The number of R pikes per time gives an estimate of heart rate and the ST segments provides information about the absence of O_2 arriving to the myocardium and the induction of ischemia. It is advisable to register the 5 derivations [32].

1.2.1.3. Arterial blood pressure

Arterial blood pressure is another method of monitoring the cardiovascular function. It is an indicator of the pressure exerted by blood flow on arterial walls, and the measurements are obtained periodically. If the blood pression is out of an established range, it is alarming and can be physiologically important [28]. This can be measured either by a non-invasive form with an inflatable blood pressure cuff or invasively with catheters placed inside a peripheral artery such as radial [29].

1.2.1.4. Capnography

Capnography gives information about the ventilatory function. It is the monitoring of the concentration of partial pressure of carbon dioxide (CO_2) in the respiratory gases.

During anesthesia, it is useful to confirm, by the CO_2 detection, the endotracheal intubation and to ensure a good communication between the anesthetic circuitry and the respiratory tract, to evaluate the ventilation and pulmonary perfusion and to detect cardiac output alterations and pulmonary embolism. [27]

1.2.2. Advanced monitoring systems

There are other advanced systems that allow clinicians to explore in deeper detail the function of organs that are not vital or are vital in a different sense. They are mostly used for surgeries where it is advanced that a complex management of the patient can occur such as cardiovascular procedures, neurosurgery or in the the critical care setting for postoperative management.

1.2.2.1. Electroencephalogram

Nowadays, the EEG is the most used advanced monitoring system. It measures the effect of anesthetics in individual patients and allows to monitor the effects induced by the doses of anesthetic drugs that are being administered. The EEG allows to observe the changes in brainwaves due to the anesthetic effects. The EEG signal is 100 times smaller than the electroencephalogram (ECG) signal, so the artifacts on the EEG can contaminate the tracing of the waves.

This technique registers the electric activity generated by the layer of pyramidal neurons in the cerebral cortex. In order to record this activity, generated by the movement of ions, some electrodes are put on the scalp surface of the patient, so it is a non-invasive method. These electrodes should optimally have an impedance to the electric signal, which ranges from 0.3-5 kΩ [35] with the objective of refusing the artifacts and processing only the correct information. Once the scalp has been conveniently cleaned, the electrodes are placed and the EEG signal is visualized in a monitor.

During anesthesia, each EEG waveform is associated with a different anesthetic state. Figure 1 shows the raw, unprocessed, EEG tracing of a patient under the effects of the hypnotic propofol, transitioning from awake to unconsciousness, passing through paradoxical excitation and sedative state. We can observe how at the beginning, in a basal state, low amplitude and high frequency pattern dominates, but as the drug is administered, it changes to high amplitude and low frequency waves. During deep anesthesia, a “burst-suppression” pattern predominates. At this point, high-frequency and large amplitude waves (bursts) are present, alternating with period of absence of electric activity appearing, isoelectric EEG, as flat lines (suppression) [31].

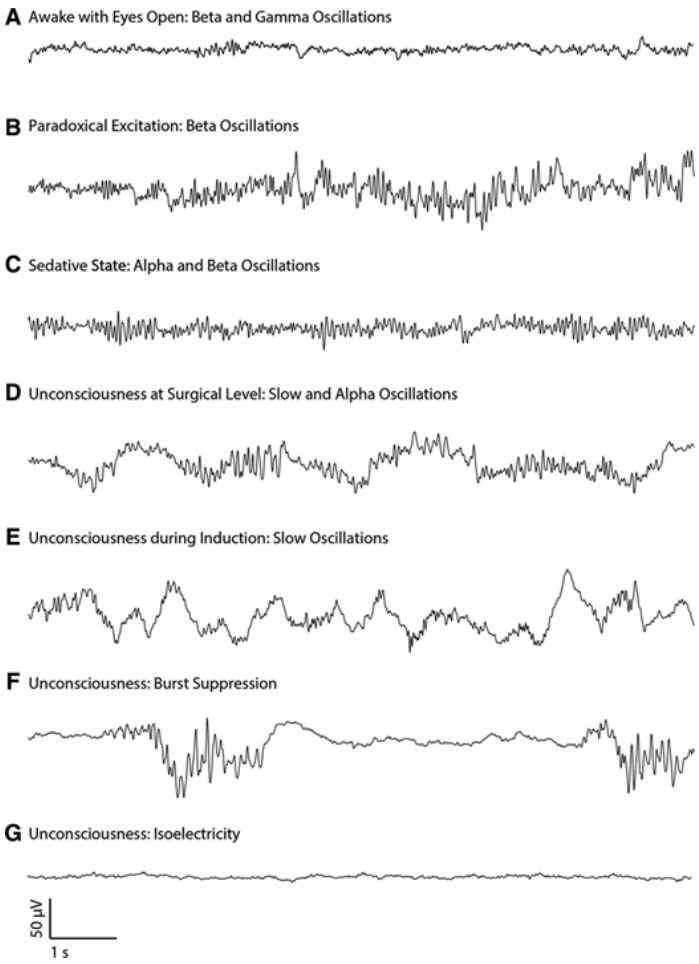

Figure 1: *Effect of propofol in the EEG unprocessed signal. From an awake to an unconscious state.*

However, interpreting this time domain analysis of the EEG in the operation room is challenging. So, a spectral analysis is computed by the spectrum, which decomposed a given segment of electroencephalogram data in different segments to derive its frequency components through Fast Fourier Transform (FFT) methods [32]. From here, brainwaves are classified by their characterizing frequencies in four main types, depicted in Table 2 Changes in power in these bands can be identified with changes in the brain's anesthetic state.

Wave	Frequency (Hz)	Clinical meaning
β	13-25	Conscious, awake, alert state
α	9-12	Relaxed, calm, lucid state
θ	5-8	Deeply relaxed
δ	1-4	Deep sleep

Table 2: *Brainwaves derived from the spectral analysis [33]*

As mentioned previously, using raw EEG to measure the depth of anesthesia is difficult because of the effort and training required to interpret such a complex signal. In order to extract from the

complex EEG signal the information only relevant to anesthetic drug effects, different monitors have been commercialized based on the application of different methods of signal analysis. The Bispectral Index (BIS) is one of the most used parameters and the one with more clinical work supporting its use. Although its algorithm is not public, the value of the BIS changes over time according to the proportions of high vs low frequency component of the EEG signal. It is a dimensionless number ranging from 99, meaning normal awake activity, to 0, which refers to absence of electrical activity in the cortex. [34]

In order to calculate the BIS value, a sensor with four electrodes is placed in the forehead of the patient, according to a predefined position, and the value displayed in the monitor corresponds to a degree of hypnotic effect. The different BIS ranges with their meaning are shown in the table below. [35]

BIS range	Degree of sedation
100-90	The patient is awake and shows response to verbal stimulation.
80-70	The patient responds to loud commands or mild shakings. Below this level, mechanical ventilation is required.
70-60	There is still a response but due to intense tactile stimulation.
60-40	General anesthesia. The patient is not able to respond to verbal stimulus.
<40	Hypnotic state.
<20	Burst suppression. At this stage, the patient shows a limited respiratory drive.
0	Totally suppressed EEG.

Table 3: *BIS ranges and its degree of sedation.*

1.3. Predictive modelling

Predicting analysis consists on extracting information from existing data and using it to predict trends and behavior patterns, to apply them to predict new events [36]. This is achieved by finding relationships between the variables of the initial data and, by an iterative process, deriving an algorithm to create a model with the ability of predicting future events. However, how well the predictive model is generated depends on the way it has been computed. The pre-processing and the amount of data is a key factor to consider just before the modelling stage.

This can be directly related with Machine Learning (ML); it is the field of study through which computers learn without being explicitly programmed [37]. Machine Learning is one of the branches of Artificial Intelligence (AI). Depending on how the algorithm has been created, we distinguish two main classes of Machine Learning approaches: [37,38]

- Supervised Learning: This approach uses labeled datasets; these datasets train the algorithm. As the output label is known, the model can measure the accuracy and learn over time, to achieve the expected output. Two problems are derived from this category:

- Classification problems: The algorithm assign new events into a specific category. They are used to predict discrete values.
- Regression problems: The algorithm is used to understand the relationship between dependent and independent variables. They are used to predict continues values.

- Unsupervised Learning: These kinds of algorithms learn by their own, from observing patterns in the input data. The unlabeled input data is divided into different clusters.

1.3.1. Logistic regression

Logistic regression is a technique borrowed by ML from the field of statistics, and based on supervised learning. It is used for binary classification problems, which aim to estimate the relationship between a dependent variable and a set of independent variables [39]. The dependent variable is binomial and it takes values of 1 (event, with probability *p*) and 0 (no event, with probability *1-p*). This method calculates the probability that the realization of the output variable belongs to the appropriate category, based on observation of input variables. Moreover, the model equation is an exponential function, which can be transformed to a linear function if a logarithmic transformation is applied.

The main objective of a logistic regression is to model the influence of some variables in the occurrence of an event and to study which factors contribute to it and which ones are negligible [39].

This model uses a sigmoid function as a cost function, which is depicted in Figure 2. The domain of the sigmoid function is limited between 0 and 1, so it is suitable for classification methods; the predicted output lies within this range.

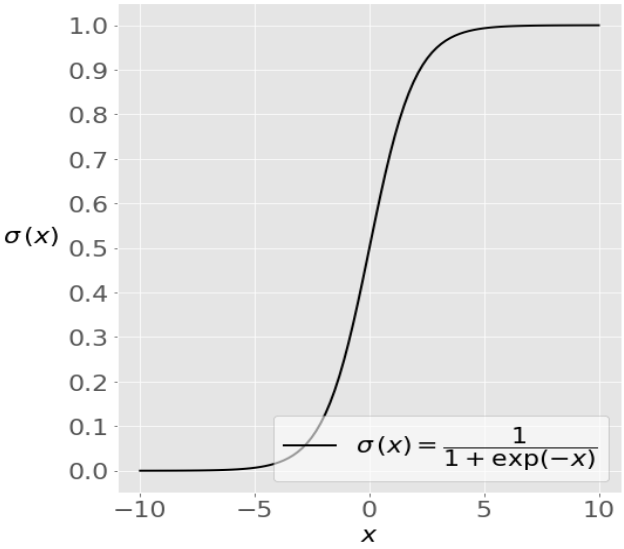

***Figure 2**: Sigmoid function.*

Logistic regression models the probability of an outcome based on individual characteristics. As the probability is a ratio, the logarithm of this probability will be modeled and the equation will be [40]:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m \beta_m$$

Equation 1: *Multivariate logistic regression model equation.*

where p is the probability that an event occurs, β_i are the estimators of the regression coefficients (predicted weights or coefficients), x_i are the independent variables or risk factors and β_0 is the intercept, which is the expected mean value of $f(x)$ when all x_i are 0.

In the case of a simple logistic regression, the equation above is reduced to:

$$f(x) = \log\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta_1 x_1$$

Equation 2*: Simple logistic regression model equation*.

If the cost function is applied to Equation 2, the logistic regression function, representing the probability for an event to occur, is obtained: [41]

$$p(x) = \log\left(\frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_m \beta_m)}}\right) \quad p(x) = \log\left(\frac{1}{1+e^{-(\beta_0+\beta_1 x_1)}}\right)$$

Equation 3, 4: *Probabilities of occurrence in multivariate and simple logistic regressions models, respectively.*

The model is trained to get the unknown parameters β which predict better the output. To do so, the maximum likelihood estimation is used; this method maximizes the log-likelihood function (LLF) for all observations:

$$\text{LLF} = \sum (y_i \log(p(x_i)) + (1-y_i)\log(1-p(x_i)))$$

Equation 5: *Log-likelihood function to obtain the best coefficients for the model.*

Once the probability $p(x_i)$ for a given input x_i is defined, the classification is proceeded. The input will be classified as 0 or 1 depending on the probability threshold; by default, 0.5 is taken.

1.3.2. Model evaluation

When building a model, apart from training and testing it with different groups of data, its performance has to be evaluated. This evaluation is done with the test set because if it was done with the training set, the model would be overfit. This overfitting would be consequence of a correct prediction of all the labels, as the model would remember the data with which it had been trained.

In classification problems, there are some well-known metrics to evaluate it [42]:

- Classification Accuracy

The accuracy consists on the ratio between the correct predictions and the total predictions. The higher the accuracy, the better the performance.

- Confusion matrix

The confusion matrix is a NxN table, where N is the number of classes, that contains more detailed information about the correct and the incorrect types of predictions. This is a more accurate metrics than the accuracy. The diagonal of the confusion matrix contains the true predictions, while the numbers out of the diagonal represent the classifier mislabels. So, it is interesting to have high numbers on the diagonal, because it indicates a correct performance of the model. Figure 3 shows how a confusion matrix is built.

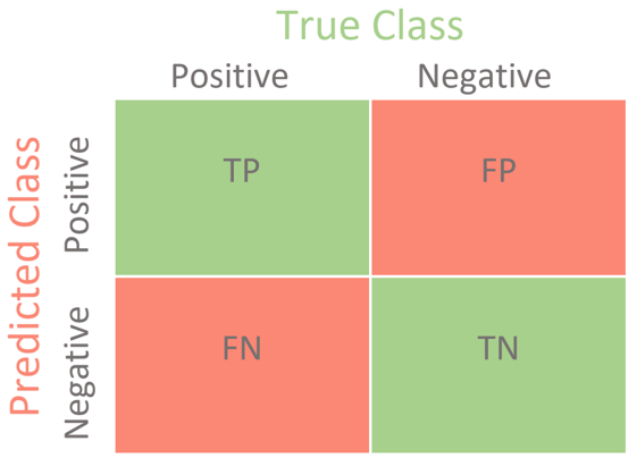

Figure 3: *Confusion matrix scheme for a 2-classes classification problem.*

As shown in Figure 3, the confusion matrix returns 4 values:

- True Positive (TP): The model predicts a positive label, and the real value is positive.
- True Negative (TN): The model predicts a negative label, and the real value is negative.
- False Positive (FP): The model predicts a positive label, and the real value is negative. (Type 1 error)
- False Negative (FN): The model predicts a negative label, and the real value is positive. (Type II error)

From these parameters, apart from the accuracy introduced before, some other evaluation metrics can be obtained.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{Precision} = \frac{TP}{TP+FP} \quad \text{Sensitivity} = \frac{TP}{TP+FN}$$

Equation 6,7,8*: Accuracy, precision and recall expressions*

$$\text{Specificity} = \frac{TN}{TN+FP} \quad F1-\text{score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}}$$

Equation 9, 10: *Specificity and F1-score expressions*

- Area under the receiver operating characteristic (ROC) curve (AUC-ROC)

The AUC-ROC is one of the most widely used metrics for performance evaluation. It summarizes the overall accuracy of the model. It is a graph which shows the performance of the model throughout all the classifiers thresholds. This curve is created by representing two parameters: true positive rate (TPR or sensitivity) versus false negative rate (FPR or 1-specificity). The different points of the curves correspond to the different cutpoints used to determine whether the test results are positive [43]. Figure 4 shows how the ROC curve is plotted.

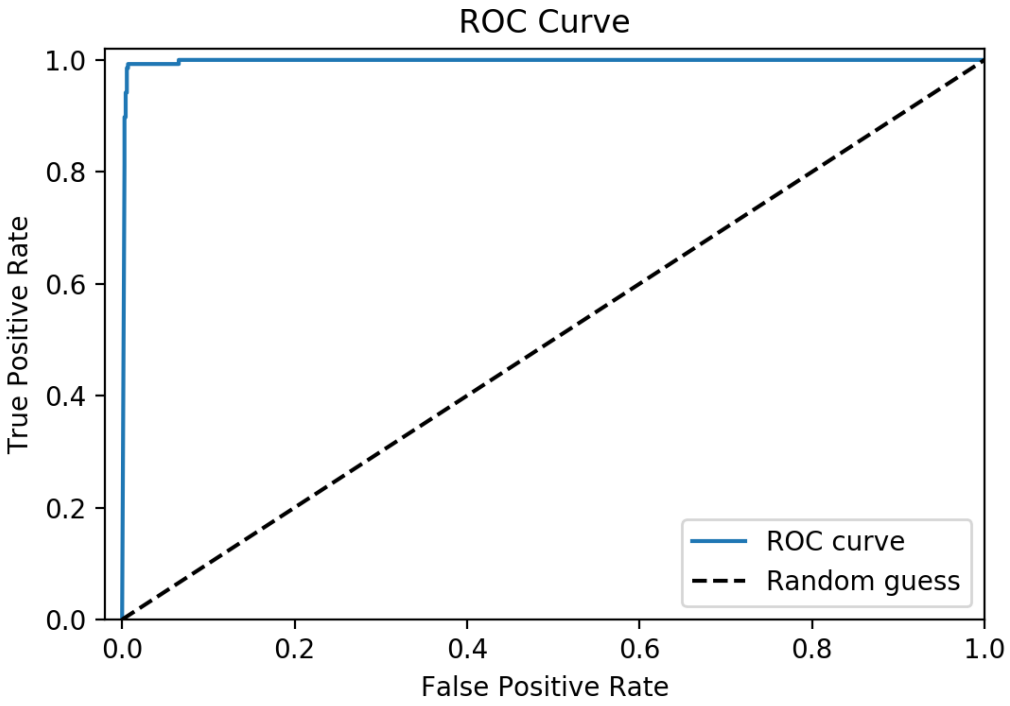

***Figure 4**: ROC curve plot*

Although all of these metrics are used to measure the performance of the classifier, their use to evaluate the performance will depend on the objectives of the project. All of them evaluate the classifier for a particular range while the ROC curve includes all the range of decision thresholds from a diagnostic test [43].

The AUC-ROC can take values between 0 and 1; 0 indicates perfectly inaccurate test and 1 reflects a perfectly accurate test. Between them, an AUC of 0.7-0.8 is acceptable, 0.8-0.9 is excellent and more than 0.9 is outstanding [43].

2. STATE OF ART

General anesthesia is defined as the combination of loss of consciousness and inability to perceive painful stimulation due to the administration of anesthetic drugs. [44] Evaluating depth on anesthesia, either the hypnotic effect, referred to the loss of consciousness, and the analgesic effect, referred to the loss of pain perception, requires monitoring. This section will introduce the main parameters used to monitor both states.

If we first focus on the hypnotic effect, nowadays anesthesiologists use different approaches; an example could be clinical signs during the induction stage, such as the absence of response to verbal command or loss of palpebral reflex.

These clinical signs are not continuous. That is why the EEG is also used to measure the depth of the hypnotic effect. Although understanding the raw EEG is challenging, there are some monitors which process the data coming from it and give as an output a parameter which can be clinically interpretable by the anesthesiologist. The most important EEG monitoring methods are: the EEG spectrum frequency, the bispectral analysis (see section 1.1.2), EEG entropy and the evoked potentials (visual, sensorial or auditory). [44] While the first three monitors are passive, based in spontaneous signals coming from the EEG, the last ones are active and measure the response to a stimulation state.

In order to proceed with the spectrum frequency analysis, an EEG-Fourier Transform has to be done; as the hypnotic drugs' effect increase, the EEG activity decreases. Among all of the evoked potentials, the auditory evoked potentials (AEP) are the most affected by the hypnotic drugs. They are calculated by applying repeatedly an auditory stimulus to the patient and analyzing a period of 100 ms after the stimulation, just when the anesthetic drugs can have more incidence on them. [45]

As it has been introduced at the beginning of the section, the analgesic state is another key stage to analyze during the monitorization of the depth of anesthesia. The first physiological response to nociceptive stimulus is an activation of the autonomic neural and hormonal pathways [46]. Although there is not an objective measure of the analgesic state, some parameters are nowadays used to generate the nociceptive monitoring monitors. Vital signs are used, such as the arterial blood pressure, heart rate, sweat or tears (known as PRST) or the absence of movement [44]. Moreover, the heart rate (HR), specifically a reduction of the heart rate variability power in the high frequency band (HRV-HF), derived from the electrocardiogram [47], or a reduction in the photo-pletysmographic waveform amplitude (PA) are used. Moreover, electro galvanic skin properties are considered to evaluate responses to pain, which measure changes in skin conductance or number of skin conductance fluctuations. [46]

3. MARKET ANALYSIS

Once introduced the different parameters used to monitor the depth of anesthesia, this section will introduce the commercially available monitors to evaluate it. The way to proceed will be the same as in the previous section. First of all, the hypnotic monitors will be introduced. Then, the analgesia monitors will be reviewed.

3.1.1. Hypnosis monitors

Some monitors which aim to evaluate the hypnotic effect during anesthesia are presented below [45]. All of them are derived from the EEG signal.

- BIS monitor (Aspect Medical Systems): Its main component is the bispectral analysis, which evaluates the phase relations from a single channel of EEG signal. The output is the BIS index (explained in section 1.2.2). The BIS algorithm is unknown but it is constantly updated. This monitor, as all EEG measurement systems, is altered by artefacts, such as the electrocardiogram, eye movement and power grid interferences, which have to be removed using different filtering methods prior to the algorithm calculation.
- Narcotrend monitor (MonitorTechnik): This monitor classifies anesthesia into five stages (A-F), each one subdivided in three sub-stages. The parameters of the algorithm are extracted from the analysis of the frequency interval between 0.5 and 47 Hz and the output ranges between 0 and 100.
- Patient State Index (PSI, Masimo): It calculates the output PSI (patient state index) value from four EEG channels and features from the 0-50 Hz and the specific frequency bands are calculated. The PSI is also a 0-100 range dimensionless number.
- qCON monitor (Quantium Medical) [48]: In this case, qCON is obtained from a single EEG channel, and after eliminating artefacts and processing the signal using spectrum analysis. An Adaptative Neuro Fuzzy Interference System (ANFIS) trains a model to generate the output value. The output is the qCON index, which ranges between 0 and 99; high values of this index represent an awake state while low values indicate absence of brain activity. During anesthesia, a value between 40 and 60 is recommended.

3.1.2. Analgesia monitors

In the case of analgesia, the commercialized monitors detect one or two parameters: [49]

- Analgesia Nociception Index (ANI): It is a dimensionless index which ranges between 0-100, calculated from the area under the curve of the high-frequency spectrum of the HRV. Higher values of ANI are related with less stress and less pain, as they represent higher parasympathetic activity.
- Nociception level index (NOL): It is a multiparameter nociception monitor, as the algorithm comes from photoplethysmography, galvanic skin response, temperature and accelerometer. It is dimensionless and goes between 0 and 100. During general anesthesia, values between 10 and 25 are advisable.
- qNOX (Quantium Medical): This index takes values between 0 and 99 and, as qCON, it is modelled using ANFIS. Its aim is providing information about the probability of movement

when a noxious stimulus is applied. A qNOX value above 60 represents a high probability of response to a painful stimulus.

4. CONCEPT ENGINEERING

Monitoring the depth of anesthesia is an important factor in the operating rooms. This passes through being able to determine whether the anesthetic agents are being administered properly to maintain a good level of anesthesia. On contrary, underdosage or overdosage would lead to an intraoperative awareness with recall or to a prolonged recovery time with postoperative complications, respectively [45].

Moreover, as it has been reviewed previously, either the dose of propofol an remifentanil induces changes in the body which are translated in a level of depth of anesthesia. So, generating a model with the ability to predict whether new drug concentrations would lead to an acceptable depth of anesthesia would help anesthesiologists to predict adverse events.

First of all, among all the different monitors introduced in section 2.2, the BIS has been chosen as the output to train and fit the model. However, the goal of this project is not exactly monitoring the depth of anesthesia in real time but being able to predict this state, from a given parameters such as propofol or remifentanil doses in terms of BIS index.

When thinking about the possibilities to generate the model, different approaches came to my mind, all of them related with ML algorithms. Different possibilities such as Deep Learning (DL) techniques and Logistic Regression methods were evaluated. DL is a branch of ML which works with neuronal networks, with highly connected input and output nodes. Due to the over-parameterized black-box nature, it is difficult to interpret the prediction results. [50] Nonetheless, as it will be explained later on, as the available data can be processed to obtain binary input variables, Logistic Regression appeared to be a good option to generate the model.

Another important aspect to consider is the software used to implement the model. RStudio, MATLAB and Python were the different possibilities. They are programming languages which offer some libraries to build a logistic regression model and they are compatible with the given data, so any of them would be suitable. But, when choosing between one of them, the most important factor I considered was the previous experience. Although they have been introduced to me during the previous courses, Python is the option with which I am most familiarized and, from my point of view, the most intuitive one.

At the end, the proposed solution to proceed with the implementation of the model was using a Logistic Regression method with Python.

5. DETAIL ENGINEERING

Once the different possibilities have been considered and the logistic regression method has been chosen, it has to be implemented.

In order to understand the methodology followed, depicted in Figure 5, the objective of the project is introduced. As mentioned in sections above, the aim of this study is to calculate the probability to be in an optimal level of depth of anesthesia derived from an input concentration of drugs. This depth of anesthesia is obtained from the BIS index, and we will consider that the optimal range comprises values between 45 and 55. So, the logistic regression will be applied between the input drug concentration and the corresponding BIS. However, to make the modelling step easier and more interpretable, there is an intermediate step which includes transforming the drug concentration to a predicted BIS value, generated with the propofol's Schnider model, which will be explained later on. At the end, the data of the model will be the predicted BIS value, as independent variable, and the binomial real BIS index, obtained from the patients, as the dependent variable. Once this first model is generated, some covariables such as remifentanil concentrations or demographic variables (age, weight, height) are added in order to study their influence in the predictions.

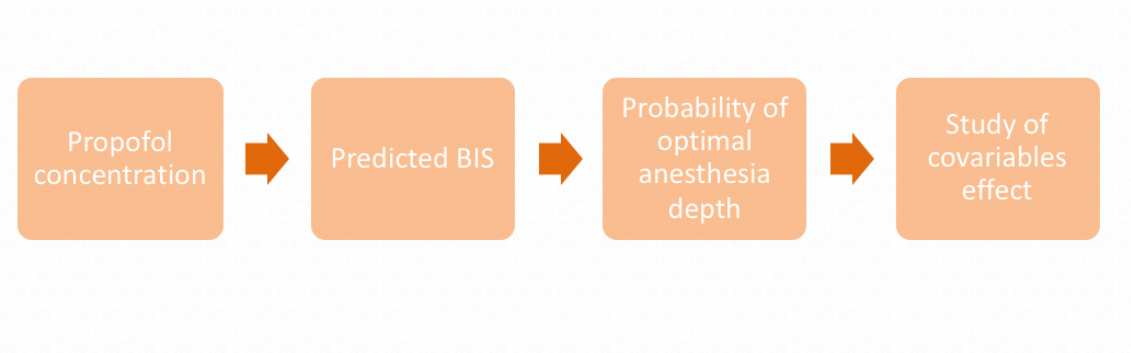

Figure 5: *Workflow followed during the development of the model building.*

All models are built from a set of available data; the first step was familiarizing with it. This consisted in a stay at CMA acquiring data from patients undergoing general anesthesia, and it lasted one month approximately. After being familiarized with the surgical environment and with the machines from the tower of control of anesthesia, it came the model building stage. But the initial data had to suffer some preprocessing steps before using it to train the model.

The next step was building the model. The data had to be separated into two groups: training and test. Once the model was trained with the first subgroup, and the proper parameters of the model were obtained, it had to be validated with the test group. Ideally, after validating the model with the test data, a prospective validation would have to be proceeded, with data collected after building the model. However, due to the lack of time this step is out of scope. Finally, interpretation of the results was carried out.

Taking this into account, the methodological part of my project consisted of three main parts. The first one was the data collection, then the preprocessing of the data was proceeded in order to generate the model and, finally, the model was built; these are reviewed in this section.

5.1. Data acquisition

In order to proceed with the modelling, the first thing needed is the data. The data belongs to the SPEC-M (Systems Pharmacology Effect Control & Modelling) Anesthesiology department of Hospital Clinic of Barcelona. Specifically, it is collected from patients who undergo surgery with general anesthesia in the operating room number 4, carried out by Dr. Gambús. These surgeries are all related with gynecologic and obstetrics procedures including laparoscopy, among others; thus, all the population of the study is formed by women.

The data has been recorded under the authorization of the Ethics and Clinical Research Commmittee (CEIC) of Hospital Clinic of Barcelona (Ref nº 2013/8356). My stay at CMA collecting data took place between February, 2021 and April, 2021. However, the database is highly extensive, as recordings from patients began in 2013; nowadays, it is fed with more than 1500 patients.

Moreover, one important fact to consider is that the anesthetic agent is administered by TIVA techniques (see section 1.1.2). As introduced above, the main drugs delivered to the patients in this surgical room are mostly propofol, remifentanil and rocuronium and the administering follows the same pattern in almost of all the patients. Firstly, propofol is administered by the infusion pump TCI-TIVA, with doses between 2-2.5 mg/kg and once the hypnotic effect is achieved, remifentanil is delivered to generate analgesia. Its initial dose is 1 mg/kg and it involves the placement of LMA or tracheal intubation, as explained at the beginning. This leads to the administering of rocuronium to block the musculature and immobilize the patient (usually a 30 mg dose). During surgery, other drugs can be administered to maintain the patient's constants stable: atropine (if bradycardia occurs) and ephedrine or urapidil (if hypotension).

The Anesthesiology control tower of SPEC-M research group in Hospital Clinic is depicted below. It consists of three monitoring devices (BIS VISTA, Conox and Drägger Infinity Gamma), the TCI-TIVA system and a computer where the data is stored. The data is acquired every 1 second and it is stored in a computer; each patient is saved in a different file. The different events occurred during surgery such as the loss of verbal or palpebral response, the corneal reflex, intubation and extubating during mechanical ventilation or infusion of ephedrine or atropine are recorded with the remote control (see Figure 6), also connected to the computer.

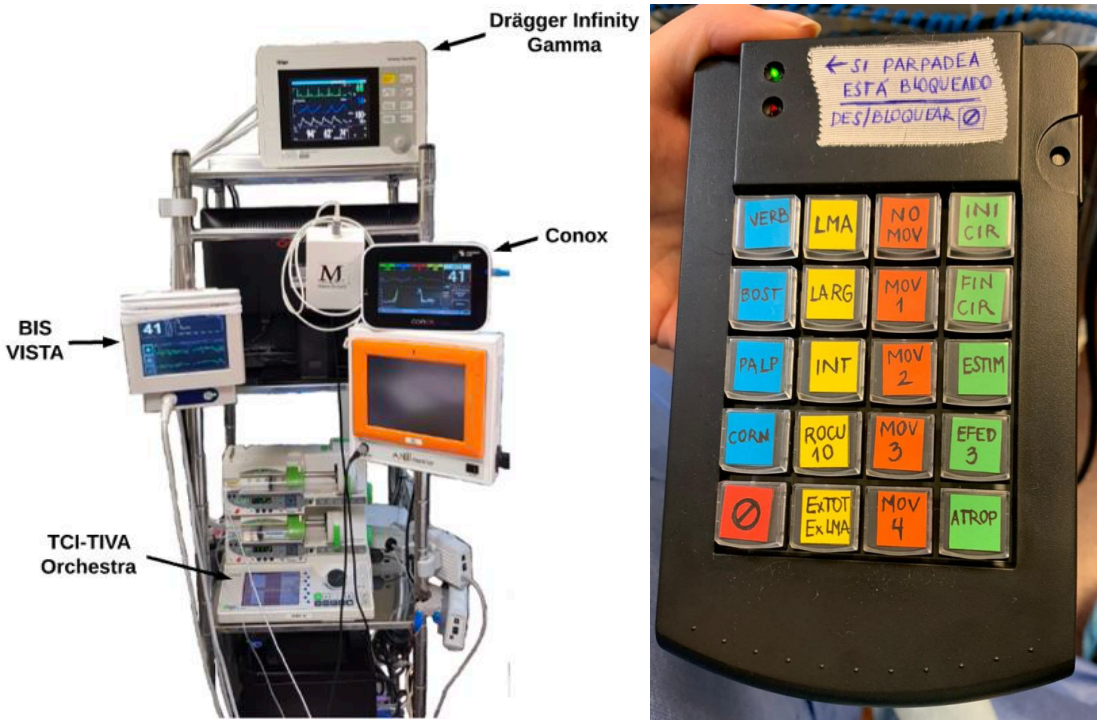

Figure 6: *On the left, anesthesiology control tower, with the monitoring devices and the TCI infusion pump. On the right, the remote control used to collect data.*

Data from BIS VISTA, Drager Infinity Gamma and TCI-TIVA is computed with Rugloop II® software. In addition, this software allows to write down the different events that take place during surgery, obtained from pressing the remote-control buttons. Then, Conox® collects data separately and both softwares are then synchronized to obtain the recordings in the same file.

The different monitoring machines comprising the control tower, with the data obtained from each one, are descripted in the table below [51-53]:

Monitor (commercial company)	Description	Parameters
BIS VISTA™ (Aspect Medical Systems)	It monitors the hypnotic state of the brain, from the acquisition and processing of EEG signals. Four electrodes are placed on the patient's forehead to acquire them.	SystTimeBIS BIS (Bispectral index) EMGBIS (Electromiogram index) BBSBIS (Burst Supression index) SQI09 (Signal Quality Indicator)
Infinity® Gamma (Dräger)	It is a compact vital signs (hemodynamic) monitor. It provides information about ECG, respiration, pulseoximetry (SpO_2), temperature and blood pressure.	HR (Heart Rate) NIBPsys (systolic blood pressure) NIBPdia (diastolic blood pressure)

	Three electrodes are placed in the patient's chest to measure ECG derived parameters.	NIBPmean (mean blood pressure) RespiRate (Respiratory Rate) SPO2 (oxygen saturation)
Total Controlled Infusion-Total Intravenous Anesthesia System/TCI-TIVA Orchestra® (Fresnius Kabi)	It allows anesthetics to achieve a stable plasma or estimated effect-site concentration of propofol or Remifentanil. The drug is automatically delivered through a radial venous. The system calculates the infusion rate in order to maintain the predefined target concentration. To do so, demographic parameters from the patient are considered.	CpREMI – CpPROP (Plasmatic concentration) CeREMI – CePROP (Effect-site concentration) InfVolREMI – InfVolPROP (Infusied volum) InfRateREMI – InfRatePROP (infusion rate) Age Weight Height Gender
Conox® (Quantium Medical)	This non-invasive monitor collects information about the hypnotic and the analgesic effect, through qCON and qNOX indexes, respectively (see section 2.2). To do so, it uses three electrodes placed in the forehead, together with the BIS electrodes.	qCON (hypnotic effect index) qCONEMG (Electromiogram index) qCONBS (Burst Supression index) qCONSQI (Signal Quality Indicator) qCONqNOX (analgesic effect index) qCONZref qCONZpos 0-127 Hz (EEG spectrum at each frequency)

Table 4: *Monitoring systems and parameters obtained from the control tower of anesthesiology.*

The resulting files are Excel documents with different columns, one for each parameter introduced in the table above. Apart from these columns, three parameters are added: an ID patient column, a Time column (in seconds) and an Events column. Moreover, each row represents one second of the surgery.

5.2. Data preprocessing

After two months collecting data to feed the database, and once being familiarized with the anesthesia environment in the operating room, it comes the moment to start the modelling step. But, the data used to build the model is not that which I collected during my stay at CMA because at the beginning of the modelling stage I was simultaneously acquiring data in the control tower and training the model.

So, the data I used was already available from 85 prior patients with ID numbers between 568 and 665; thus, 85 excel files are the initial raw data of the project. Before proceeding with the modelling step, this data has to be processed in order to eliminate the unneeded columns and rows and to dichotomize the variables to prepare it for the Logistic Regression model.

This preprocessing is generated using two libraries from the Python program: pandas and numpy. Both of them are used as data analysis and manipulation tools. Furthermore, an available Schnider pharmacokinetic pharmacodynamic (PKPD) propofol model built with Monolix software is used to generate the BIS predictions from these patients.

First of all, the prediction of BIS is carried out. The Schnider model considers two compartments and an inhibitory effect among BIS. These two compartments are the central (plasmatic) and the effect compartment (effect-site). From the propofol infusion rate, the plasmatic concentration is calculated by Schnider model. Then, following a first order equation (Eq. 11) with a first-order rate constant Ke, the effect-site concentration is obtained.

$$\frac{\mathrm{dCe}}{\mathrm{dt}} = \mathrm{Ke} * (\mathrm{Cp} - \mathrm{Ce})$$

Equation 11: *First order equation to obtain Ce from Cp*

And finally, the BIS prediction is obtained from Equation 12,

$$\mathrm{BIS}_{\mathrm{predicted}} = 100 * \left(1 - \frac{\mathrm{Ce}^{\gamma}}{\mathrm{Ce50}^{\gamma} + \mathrm{Ce}^{\gamma}}\right) + \mathrm{e}$$

Equation 12: *BIS prediction expression, derived from Ce*

where Ce50 is the Ce associated with 50% of the maximum effect and e is the residual error, considered fix. The BIS inhibitory effect is introduced by Ce50 and γ, which is the steepness of the concentration versus the response relation and adds a sigmoid shape to this effect.

This prediction step has been run by Dr. Jaramillo because he has access to this Monolix model. In order to reduce the time to predict the BIS index, the initial data is reduced and the data passed to the Schnider model is 10 times smaller. Instead of predicting the BIS at every second, one row each ten is selected. Moreover, all the data is put in a single file to use it in Monolix.

The results of the parameters obtained with the model are shown in the table below, with no significant errors, as the relative standard errors are below 20% in all the cases:

	VALUE	STOCH. APPROX.	
		S.E.	R.S.E.(%)
Fixed Effects			
Ke_pop	0.13	0.0011	0.854
Ce50_pop	2.25	0.0063	0.279
gamma_pop	1.46	0.01	0.692
Error Model Parameters			
e	12.91	0.048	0.369

Table 5: *Parameters of the PKPD propofol-BIS model. Estimated standard error (S.E) and Relative Standar Error (R.S.E.) for each parameter are shown in the table.*

The "predictions vs observations" plot can be seen in Figure 7. As it is depicted, the error between them is not so high. They follow the ideal theoretical line represented in black, with some deviation. This deviation corresponds to the standard error, e, of the model, which is ±12.91.

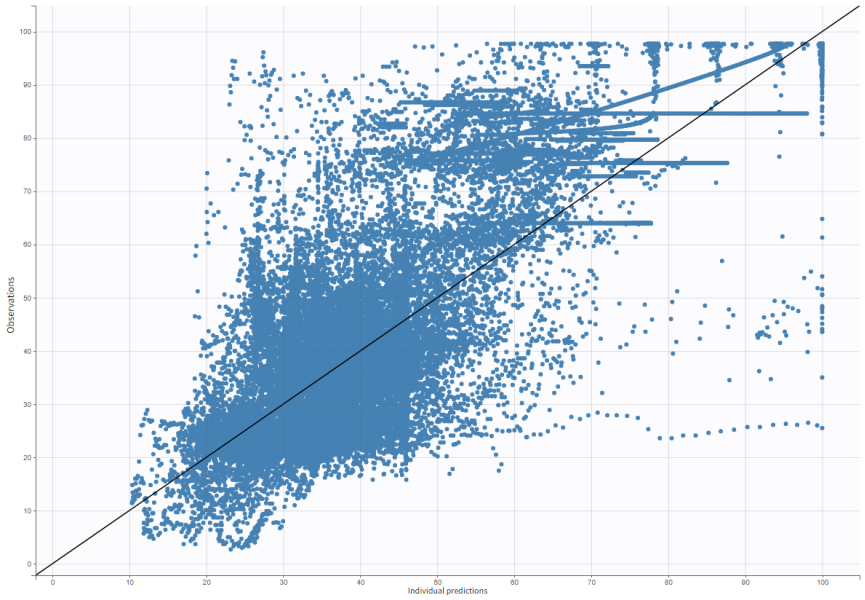

Figure 7*: Prediction vs observations plot.*

Figure 8 shows some examples of the real BIS and the predicted BIS. In these graphs, it can be seen clearer how the BIS predictions done with the propofol model follow the dynamics of the real BIS.

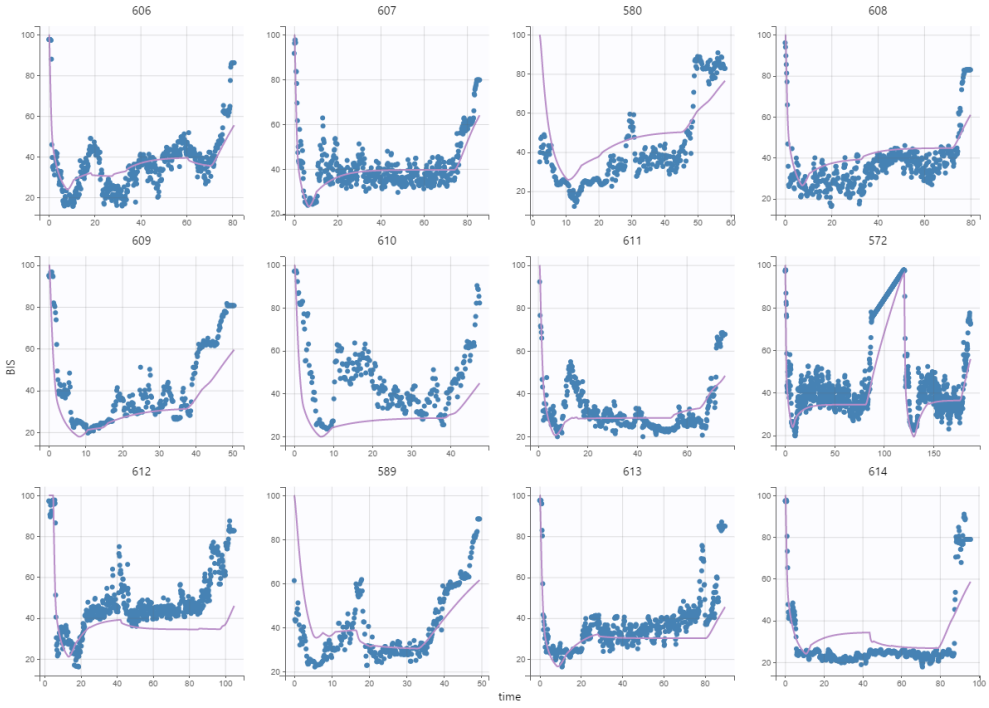

Figure 8: *Individual predictions for some individuals.*

Once the BIS predictions are generated, the preprocessing step has to be followed. To generate the main model, two sub-models are created. Due to the sigmoid shape of the Logistic Regression, the model cannot be built with the whole data. That is why 0-50 and 50-100 subgroups are obtained from the predicted BIS value. As the optimal range comprises very few values (10 values) from the total BIS range (100 values), this division is also made to improve the accuracy of the model.

First of all, a rearrangement of the file with which I work takes place. From all the initial columns introduced in Table 4, only a few of them are needed:

- BIS: recorded BIS. This column is binarized as:
    - BIS_up_45: All the BIS values above 45 are set to 1. Values from 0 to 45 are transformed into 0.
    - BIS_down_55: All the BIS values below 55 are set to 1. Values from 55 to 100 are transformed into 0.
- Pred_BIS: the BIS predicted by the propofol model.
- Pred_BIS_disc: PRED_BIS column dichotomized in order to obtain a 0 when Pred_BIS<50 (model A) and a 1 when Pred_BIS>50 (model B). This is only done to divide the data into two files, to generate each sub model.

This file will be used in the next section to build the first model, which only considers the effect of propofol.

When studying the effect introduced by the covariables, two more files are created which include the Ce_Remi column when the effect of Remifentanil wants to be analyzed, and the Ce_Remi + demographic data columns (age, height and weight) when the role of this information wants to be

studied. The last two cases imply a multiple logistic regression, while the propofol model is a simple logistic regression.

One important aspect to consider is the fact that I choose Ce_Remi as covariable and not Cp_Remi. The reason for that is because Cp_remi shows an initial peak, called the plasma overshoot [55], which aims to generate the desired Ce at equilibrium and calculated from the rate constant for drug transfer intro the brain. This initial peak could introduce some error in the results.

This rearrangement of the initial file is done to obtain a final result that enhances the probability to obtain a real BIS above 45 in the first half of the data (0<Pred_BIS<50 or Pred_BIS_disc= 0) and the probability to obtain a real BIS below 55 in the second half of the data (50<Pred_BIS<100 or Pred_BIS_disc=1).

Ideally, the final result would be as depicted in Figure 9: prediction values of BIS between 45 and 55 showing a probability of 1 to be in the optimal range.

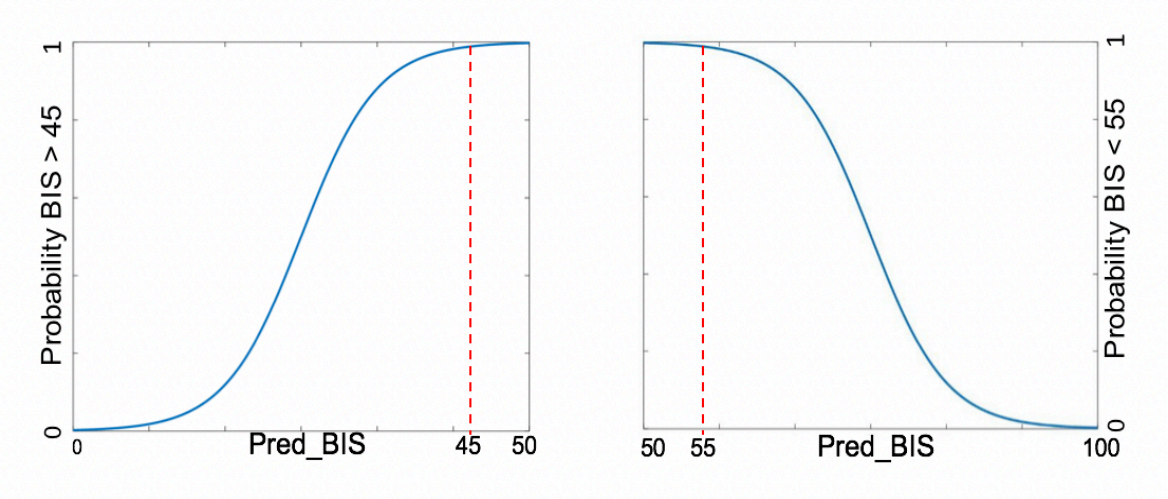

Figure 9: *Schematic of the logistic regression model to be trained and evaluated.*

So, at the end, three models are performed, which lead to six logistic regressions; each model has two sub-models, depending whether the value of the predicted BIS is between 0 and 50 or 50 and 100 which, at the same time, will depend on the initial concentration of propofol.

As it is shown in Figure 9, the density of points of the data is really high in almost all the BIS range. As the optimal range only comprises 10 points of 100, if the model was trained with this original data it would lead to overfitting. So, the way to avoid this fact is eliminating data from the original file, so that in the case of model A the difference in the number of data between 0 and 45 and 45 and 50 is not so high; and the same with the model B. To have an idea of this huge difference of points, there are 30636 points in model A and only 5184 range between 45 and 50. So, from these 30636 values, only 9000 are maintained, and the rest are eliminated. The same occurs with the predicted BIS in model B: there are approximately 6000 values but only 1500 are in the optimal range (between 50 and 55). So, some rows are also eliminated to avoid the model mislabeling due to the overfitting.

Moreover, the set of dependent (y) and independent (x) variables are defined in this part. In all the cases, the dependent variables are the binary variables obtained from the real BIS columns: BIS > 45 or BIS<55 in the model A and B, respectively, which, at the end, after joining model A and model B will represent the optimal depth of anesthesia range. The way of selecting the independent variables differs from the simple and the multiple logistic regression study. The simple case only has one independent variable, the Pred_BIS column, which is directly related with the propofol effect. However, the multiple models adds some columns: the Ce_Remi column in the second model and the Ce_Remi + age + height + weight columns in the third model.

In order to summarize this preprocessing step and to understand the different columns contained by each file, Table 6 has been created:

Model	Data	
Model 1 (A and B)	ID_patient	Bis_down_55
	Time	Pred_BIS
	Bis_up_45	Pred_BIS_disc
Model 2 (A)	Bis_up_45	
	Pred_BIS	
	Ce_REMI	
Model 2 (B)	Bis_down_55	
	Pred_BIS	
	Ce_REMI	
Model 3 (A)	Bis_up_45	Age
	Pred_BIS	Height
	Ce_REMI	Weight
Model 3 (B)	Bis_down_55	Age
	Pred_BIS	Height
	Ce_REMI	Weight

Table 6: *Columns maintained after the preprocessing step.*

Finally, just before running the model, data has to be separated into two groups: training and testing set. This division is needed to evaluate the performance of the model and to assess how good the predictions are done. In this case, 75% of the data is randomly assigned to the training group and the 25% resting forms the testing group. This is done with a function from the Python package Scikit-Learn, called train_test_split().

5.3. Model building

Once the data is processed, dichotomized and equilibrated to avoid overfitting, the model is built. In order to build the logistic regression models, Python offers many libraries; among others, the two most well-known libraries are StatsModels and Scikit-Learn. Although Scikit-Learn is popularly more related to ML and data science and StatsModels is matched with econometrics, generalized-linear-models, timeseries-analysis and regression-models both of them are useful in this step.

StatsModels is used to train the model, because it offers the possibility to generate multiple regression models, which are important when the effect of the covariables wants to be assessed. Furthermore, it reports the results in a more complete and understandable way. Moreover, Scikit-Learn is useful when testing the model, because it offers a direct function to generate the ROC curves.

5.3.1. Model training

The way of training the model is the same either in the simple logistic regression and in the multiple case. StatsModels provides a function which performs the logistic regression model: Logit(). But this library does not include the intercept β_0 by default, so the way to include it in the final results is by using the sm.add_constant() function. Then, the next step is fitting the model to the data with fit().

Once the model is trained, StatsModels provides a detailed summary of the results with the summary() function. The output contains the number of "Iterations" that have been done over the data, in order to optimize the model and the results. By default, the maximum iterations that can be performed are 35; after this number, the optimization fails.

Moreover, the summary includes two tables, one with the Logit Regression Results and another with the resulting coefficients of the model, its p-value and the corresponding confidence interval. From this second table, the most important parameters in which the attention will be focused are:

- Coef: the value of the different coefficients of the dependent variables which conform logistic regression model. The const term refers to the intercept value.
- P > |z|: this column shows the p-values for the different coefficients. If this value is above 0.05, the predictor is considered "not statistically significant".

The tables with the different results are included in the Appendix of the project. However, the summarized tables showing the results of the parameters obtained by the model and their significance are introduced in the tables below.

	Model A (propofol)		Model B (propofol)	
	Coefficient	p-values	Coefficient	p-values
β_0	-4.0704	0.000	5.0237	0.000
[propofol]	0.0978	0.000	-0.0797	0.000

Table 7: *Logistic Regression parameters from propofol model.*

First of all, the simplest model is analyzed. It considers only the effect introduced by the propofol concentration. Both parameters have a p-value highly low, which Python approximates to 0; they are depicted in Table 6. So, propofol can be considered to have an impact on predicting the optimal depth on anesthesia.

From this first model, the two resulting equations describing the probability to be in the optimal depth of anesthesia range are:

$$p(x) = \log\left(\frac{1}{1 + e^{-(-4.0704 + 0.0978 \cdot [propofol])}}\right)$$

$$p(x) = \log\left(\frac{1}{1 + e^{-(5.0237 - 0.0797 \cdot [propofol])}}\right)$$

***Equation 13, 14**: Probabilities to be in 45<BIS<50 and 50<BIS<55 for propofol model, respectively.*

Once the propofol effect is studied, a multiple logistic regression is done to analyze which effect is introduced by some covariables. Therefore, the second model built introduces this effect of the Remifentanil effect-site's concentration. The results of the new parameters are depicted in Table 8. As it is shown, propofol and remifentanil are good predictors as their p-value is below 0.05.

Model A (propofol + remifentanil)			Model B (propofol + remifentanil)	
	Coefficient	p-value	Coefficient	p-value
β_0	-1.3748	0.000	4.4349	0.000
[Propofol]	0.0655	0.000	-0.0865	0.000
[Ce_Remi]	-0.6491	0.000	1.1423	0.000

***Table 8:** Logistic Regression parameters from propofol + remifentanil model.*

In this case, the resulting probabilities to be in an optimal depth of anesthesia are described by Equations 8 and 9.

$$p(x) = \log\left(\frac{1}{1 + e^{-(-1.3748 + 0.0655 \cdot [propofol] - 0.6491 \cdot [Remifentanil])}}\right)$$

$$p(x) = \log\left(\frac{1}{1 + e^{-(-4.4349 - 0.0865 \cdot [propofol] + 1.1423 \cdot [Remifentanil])}}\right)$$

***Equation 15, 16:** Probabilities to be in 45<BIS<50 and 50<BIS<55 for propofol+remifentanil model, respectively.*

Finally, the third model is the most complete one, which apart from studying the effect of Propofol and Remifentanil, introduces the effect of the demographic variables. The results of the parameters obtained with the training set are depicted in Table 9. One more time, all the variables are good

predictors of the optimal depth of anesthesia as their p-values indicate they are statistically significant.

Model A (propofol + remifentanil + demographic variables)			Model B (propofol + remifentanil + demographic variables)	
	Coefficient	p-value	Coefficient	p-value
β_0	-1.8748	0.008	-2.8895	0.175
[Propofol]	0.0911	0.000	-0.0891	0.000
[Ce_Remi]	-0.6915	0.000	0.8551	0.000
Age	-0.0194	0.000	0.0380	0.000
Height	-0.0344	0.000	0.0310	0.012
Weight	0.0949	0.000	0.0135	0.031

Table 9: *Logistic Regression parameters from propofol + remifentanil + demographic variables model.*

And the equations which indicate the probability are introduced below.

$$p(x) = \log\left(\frac{1}{1 + e^{-(-1.87+0.09\cdot[propo]-0.69\cdot[Remi]-0.02\cdot Age-0.03\cdot Height+0.09\cdot Weight)}}\right)$$

$$p(x) = \log\left(\frac{1}{1 + e^{-(-2.89-0.09\cdot[propo]-0.85\cdot[Remi]-0.04\cdot Age-0.03\cdot Height+0.01\cdot Weight)}}\right)$$

Equation 17, 18*: Probabilities to be in 45<BIS<50 and 50<BIS<55 for propofol+Remifentanil+Demographic variables model, respectively. Due to the lack of space, the parameters have been rounded to 2 digits.*

5.3.2. Model testing

The last part of the model is testing it. This stage uses the testing group that was previously set. This data has not been used to train the model, so it is useful to test it and assess its performance. The function used to test the model and predict new events is predict(). These predictions are obtained as values between 0 and 1; the nearer the result is to 1, the higher the probability to belong to the optimal range, in this case. The results are shown below.

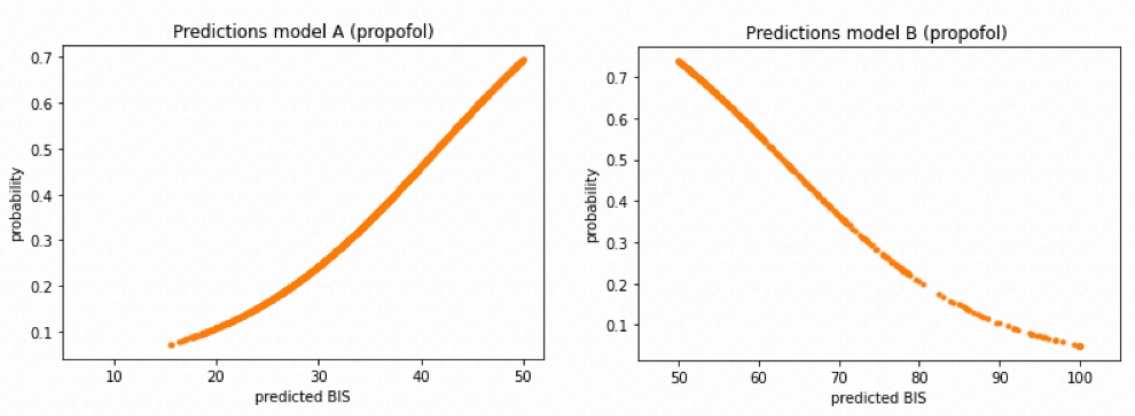

Figure 10: *Logistic Regression curve for the propofol model.*

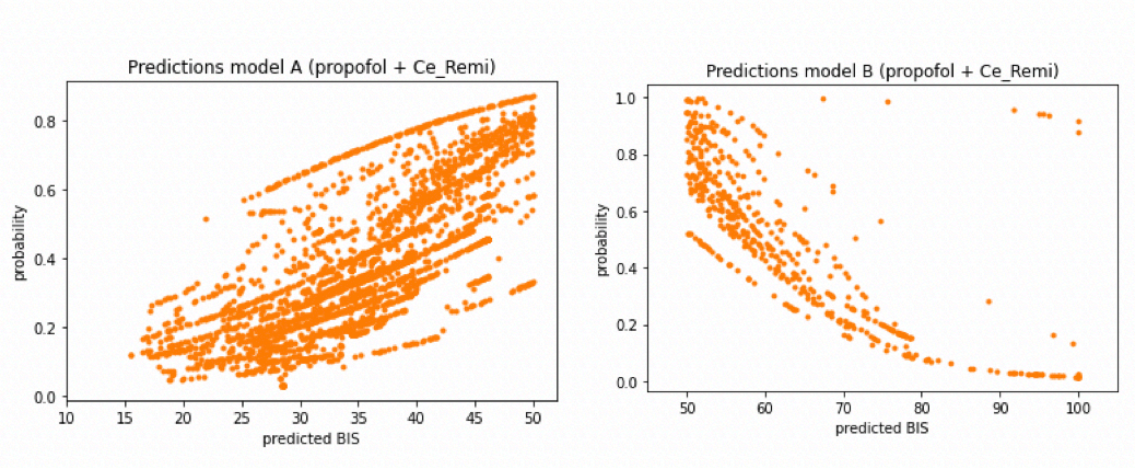

Figure 11: *Logistic Regression curve for the propofol + Ce_remi model.*

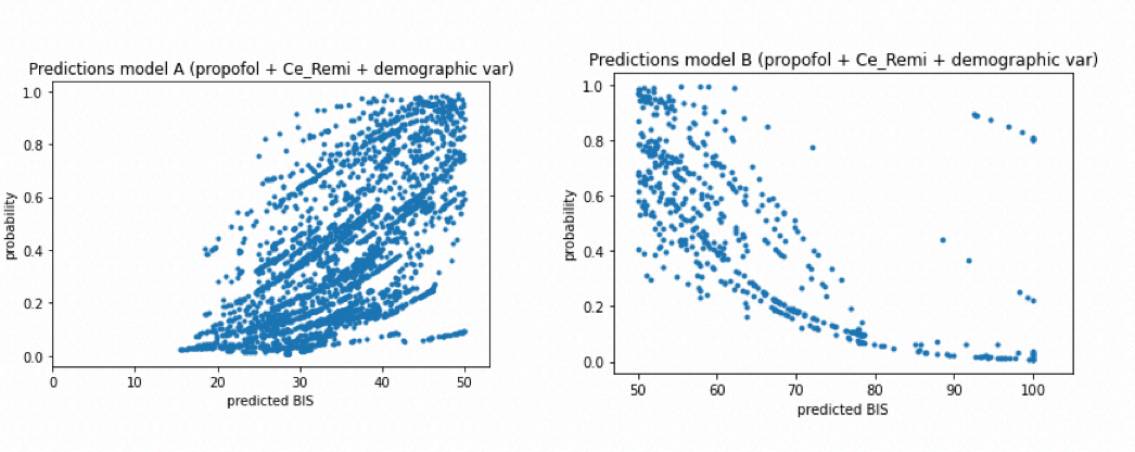

Figure 12: *Logistic Regression curve for the propofol + Ce_remi + demographic variables model.*

From these graphs we can see how the dispersion around the mean decreases; so, adding these factors could improve the model. However, to confirm it, the model has to be evaluated; this is introduced in the following section.

5.3.3. Model evaluation

Once the model is tested, its performance has to be assessed. As introduced in Section 1.2.3.2, the best way of assessing the performance of the models is by evaluating the ROC curves. This is because the main goal of this study is not to classify the new events but to obtain a probability to be in the optimal depth of anesthesia, so we will not decide one threshold which leads to a unique classifier.

So, as it has been done in the previous section, the different ROC curves are introduced following the order of the model to which they belong. Apart from the ROC curves, the results show the AUC-ROC in the legend and the best threshold analyzed by the Python software in order to obtain the best accuracy for each model. However, this threshold is not reliable as we will set one or another depending on our objectives: decreasing sensitivity or specificity.

The AUC-ROC for the propofol model are shown in the following figures and they have values of 0.67 and 0.76 for Model A and Model B, respectively. These values are acceptable but the effect of the covariables must be studied in order to see whether they introduce an improvement in the predictions.

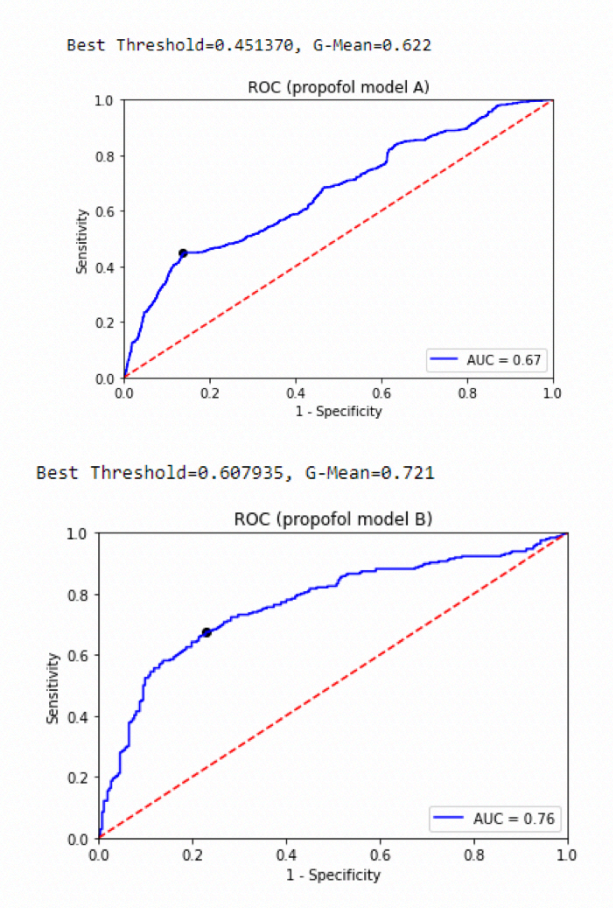

Figure 13: *ROC curves for propofol models.*

When the remifentanil effect is introduced, the ROC curves observed in figure 14 are obtained. It is observed how in both cases the area under these curves increases to 0.75 and 0.83 for models

A and B, respectively. So, remifentanil increases the predictive ability of the model and their values are more acceptable than in the previous case, where only the effect of propofol was considered.

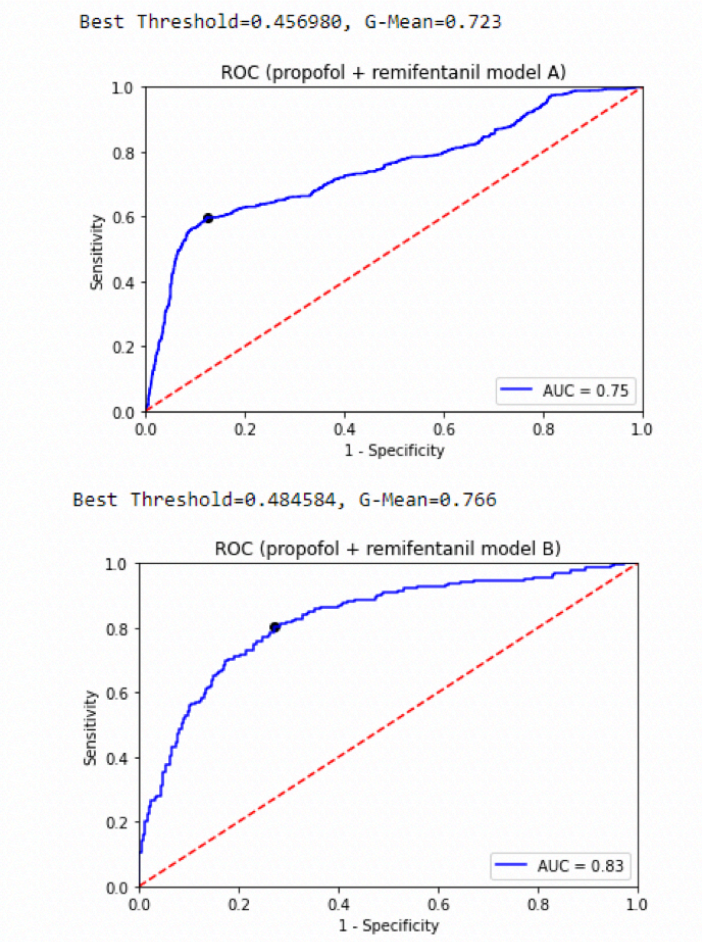

Figure 14: *ROC curves for propofol + Remifentanil models.*

Finally, when the demographic covariables are added to the model, the AUC-ROC still increases to values of 0.84 and 0.80. These values are considered excellent so, at the end, the effect of all the predictors together improves substantially the model from acceptable to excellent predictions.

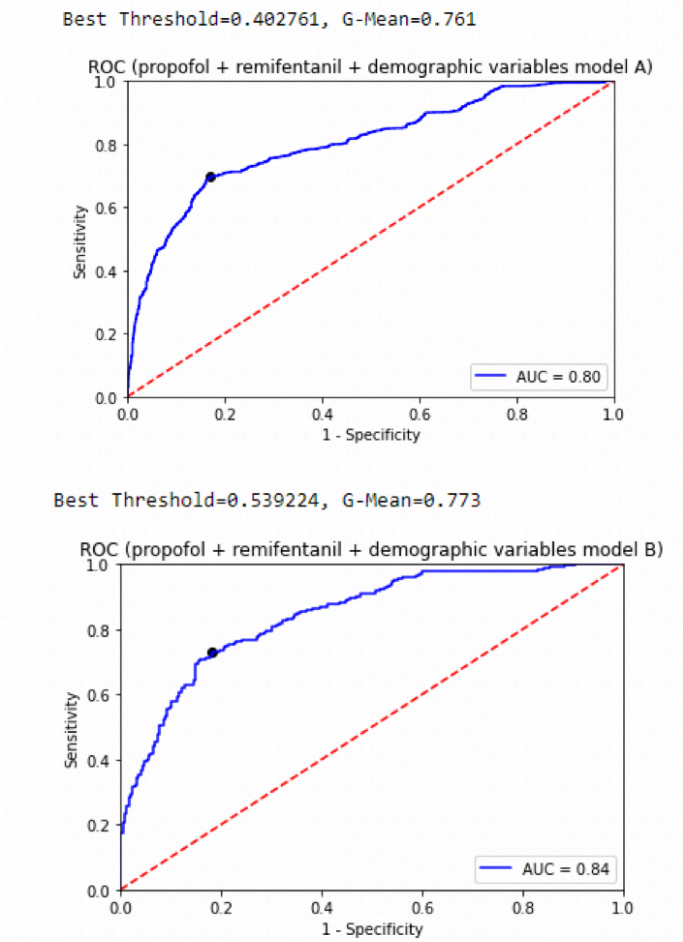

Figure 15: *ROC curves for propofol + remifentanil + demographic variables models.*

6. SCHEDULE OF EXECUTION

This section covers the schedule of execution of the project. First of all, a work-breakdown structure (WBS) is shown: a hierarchic decomposition of the different tasks to fulfill during the project is introduced. It is useful in order to have a clear idea of the main tasks. Then, a GANTT diagram is depicted, which allows to have a clearer schedule of the tasks.

6.1. Work-Breakdown Structure (WBS)

As we can see, apart from the three tasks introduced in the detail engineering section, the project required a prior preparation to familiarize with the topic and, at the end, the report is written and the final presentation is carried out.

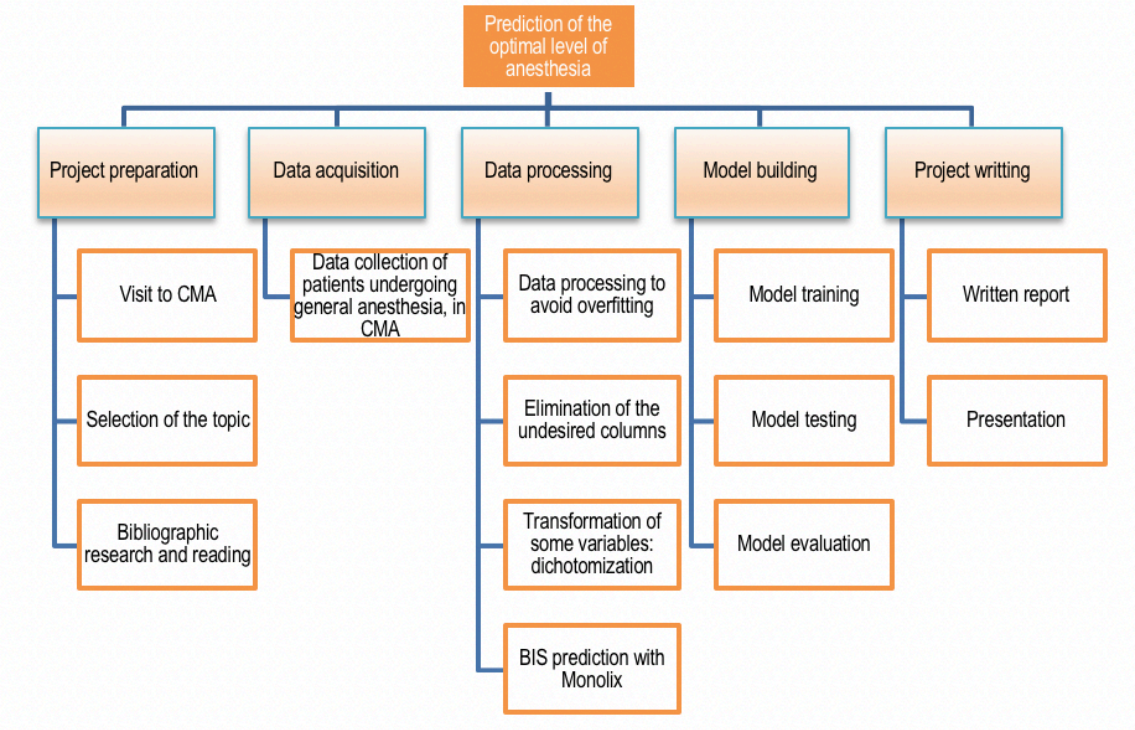

Figure 16: *WBS of the project*

6.2. GANTT Diagram

In order to generate the GANTT diagram from this WBS, Table 10 is needed. It shows the beginning and the end of each task. The data acquisition was thought to start at the end of January but, due to the third wave of Covid-19, this part finally started on February 18. The end of the model generation was on 27 April, when the supervisors approved the results and then the report was written and polished.

Activity	Description	Start date	Duration (days)	End date
A	Visit to CMA	21 dec	1	22 dec
B	Selection of the topic	2 feb	1	3 feb
C	Initial bibliographic research and reading	3 feb	16	18 feb
D	Data acquisition of patients undergoing general anesthesia	18 feb	34	24 march
E	Data processing to avoid overfitting	8 march	10	18 march
F	Elimination of the undesired columns	8 march	10	18 march
G	Transformation of some variables: dichotomization	8 march	10	18 march
H	BIS prediction with Monolix	18 march	6	24 march
I	Model training	24 march	19	12 apr
J	Model testing	12 apr	2	14 apr
K	Model evaluation	14 apr	14	27 apr
L	Written report	27 apr	48	14 june
M	Power Point Presentation	14 june	7	21 june

Table 10: *Required tasks to accomplish the project objectives.*

Figure 17 shows the GANTT diagram. It gives an idea of the different tasks which were carried out, the data delimited to accomplish them and their respective duration. Furthermore, this diagram provides an overall point of view of the project and if some of the activities can be done simultaneously.

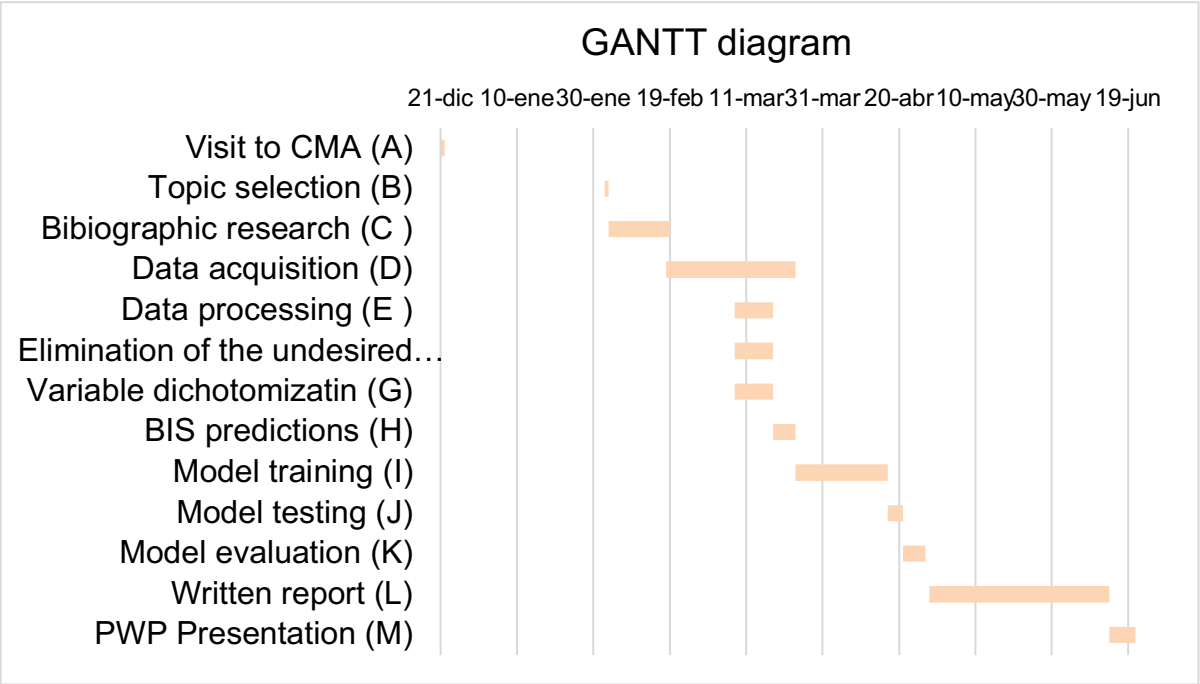

Figure 17: *GANTT diagram.*

7. TECHNICAL FEASIBILITY

The technical viability of the project refers to the technological characteristics involved in it. When talking about the technological aspects which are integrated in this project, we directly should think about the requirements of the different monitoring system devices and the software used during the programming step. However, the technical requirements of the monitoring systems used to feed the database are also introduced.

7.1. Monitoring systems' requirements

So, first of all the different technical requirements introduced by the use of the monitoring systems are commented below: [51, 55-57]

- BIS VISTA™: This monitoring system includes the monitor (P/N 185-0151), a power cord, a pole clamp, a BISx (P/N 185-0145-AMS) and a PIC (Patient interface cable which connects BISx to patient). Apart from these components, the sensors are required.
  Moreover, the monitor must operate in the proper environment and if the limits are exceeded the results can be affected. These environment limits are a temperature between 0 ºC and 40 ºC, a humidity between 15% and 95% and a pressure ranging from 360 mmHg to 800 mmHg.
  The system requires a power source of 100-240 VAC, 50-60 Hz and a maximum current consumption of 0.7 A. Furthermore, the monitor must be connected to ground to protect users and patients.
  The monitor contains a rechargeable lithium ion battery which lasts for 45 minutes when power cannot be supplied and the recharge time is about 6 hours. However, as most of the surgical procedures last from 30 minutes to 2 hours, it is advisable to charge the battery continuously by plugging it to the A/C power.
  The transfer of data can be done either by USB ports or by a RS-232 port.

- Infinity® Gamma: The monitor is equipped with a color screen. The operating ranges are approximately the same than in BIS VISTA.
  The power supply requires a VAC from 100-120 for a current of 0.8A and a 200-240 VAC for a 0.4A supply. The frequency is the same as in the previous monitor: 50-60 Hz.
  Moreover, depending on whether the battery is made of lead-acid or lithium-ion it will have a capacity of 75 minutes or 210 minutes, respectively. And its charging time will be 5.5 hours in the first case or 8 hours in the second case.

- TCI-TIVA Orchestra®: The workstation is formed by a Orchestra® Base Primea and 2 Orchestra® Module DPS. The power supply ranges from 95 to 240 V and 50-60 Hz and the battery has a minimum duration of 1 hour for the base and 5 hours for the module DPS.

- Conox®: The available information of this device is reduced. The components that integrate this medical monitoring device are the Conox screen, a USB with legal documentation, power supply and cables, patient cables and Conox sensors. The battery lasts for 2.5 hours and the power supply is 5V DC, 2A and 10W.

7.2. Software requirements

The second part of the project consisted on processing the data by the use of software. So, the different softwares used during the project and their technical requirements for the proper functioning are described below: [58-61]

- Python: The recommended hardware requirements to install Python are a modern operating system (Windows 7 or 10, Mac OS X 10.11 or higher, Linux RHEL 6/7), a 64-bit CPU (Intel /AMD architecture), 4GB of RAM and 5GB free of disk space.

- Rugloop II®: The system requirements of this software are very limited. They only involve a personal computer, a processor Intel Penitum III (PIII), a personal memory card or USB and a Windows 2000 or Windows XP, at least.

- Microsoft Excel: The system requirements are a Microsoft Windows 7 or later operating system, 1 Hz of processor speed, 1GB of minimum RAM size and 3GB of minimum Hard Drive Space. Moreover, a DVD-ROM and a DirectX 10.0 compatible graphic cards are required

- MonolixSuite2020R1: Although I have not personally used this software, it is an important part of the results. It shows a 64-bits structure, the minimum operating system is Windows 7 and 1 GB RAM or above is required. Moreover, a minimum screen resolution of 800x600 pixels is recommended.

7.3. SWOT analysis

These system requirements should be considered when developing the project. Once the project is developed, it is advisable to study its Strengths, Weaknesses, Opportunities and Threats (SWOT analysis). This analysis is helpful to identify the positive and negative aspects related to either the business competition and project planning and to create strategies to improve it. The SWOT analysis of this project is depicted in Figure 18:

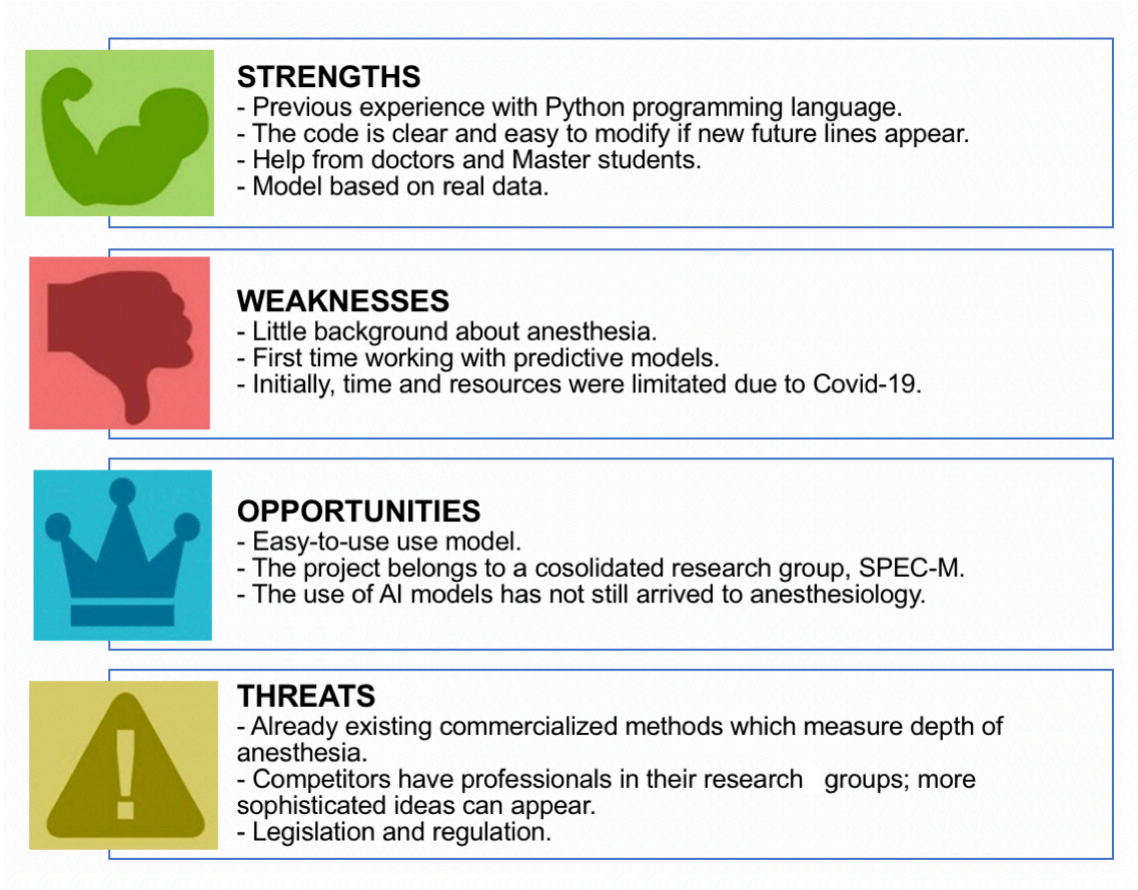

Figure 18: *SWOT Analysis.*

8. ECONOMIC FEASIBILITY

This section covers the different factors which suppose a cost during the project. The total cost of this project includes the salary of the people working on it, the computer with which the model has been done and the monitoring devices with which the database was fed.

The project has been realized by me, an undergraduate coursing biomedical engineer. But during the development of the project I have been supervised by Dr. Gambús and Dr. Jaramillo, both of them anesthesiologists. So, the breakdown of the total costs of the project is depicted in Table 11. In this section, the costs will only consider the medical devices and the software needed to build the model.

	Unity	Price/unity	Total
	WORKERS' SALARY		
Biomedical Engineer	300 hours	7€ /hour	2100 €
Supervisors	30 hours	40€ /hour	1200 €
	EQUIPMENT		
Asus F555L	1	700 €	700 €
BIS VISTA monitor	1	4750 €	4750 €
BIS VISTA electrodes	85 electrodes (1electrode/patient)	15 €	1275 €
	SOFTWARE LICENSES		
Python	1	Free	0
Monolix	1	Free	0
Microsoft 365 Personal	1	69	69 €
Rugloop II	1	269 €	269 €
TOTAL	-	-	10363 €

Table 11: *Total costs of the project.*

9. LEGAL ASPECTS

The development of this project involves some regulatory issues that have to be considered.

During the acquisition of data, we are dealing with data from patients from which we have their consent to collect their data and treat with it. Although it is introduced anonymously in the database, their medical history is available from the moment the patient enters to the pre-operative room. That is why the "Ley Orgánica 3/2018" has to be accomplished. This legislation is called "Ley de protección de datos personales y garantía de los derechos digitales" [62] which, from the clinical perspectives has something to do in this project. Specifically, all the stuff that takes part in the treatment of this data has to ensure its confidentiality.

The end product of this project is a Software as a Medical Device (SaMD). The Federal Food, Drug Administration (FDA) has made significant strides in the policies development for these products. SaMD are defined by the International Medical Device Regulators (IMDRF), the FDA's benefit-risk framework, as "*software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device."* [63] That is exactly what the Logistic Regression model intends; the medical purpose is assessing the probability of being in optimal depth of anesthesia, taking advantage of the results offered by the BIS VISTA, which would be the hardware.

It is important to know which class would the product belong to if it was launched into the market. The Regulation (EU) 2017/745 of the European Parliament and the council on medical devices [69] classifies this SaMD as class IIb; it intends to monitor vital physiological parameters, whose variations could result in immediate danger to the patient.

Once the SaMD product was validated and wanted to be commercialized, a marketing application should have to be submitted to FDA. This application should contain the submission type and data requirements based on the risk of this SaMD and if the permission was granted, the medical device could be distributed. If some later modifications were made to the design of the device, they should be reviewed under 510(k) notification. This notification software aims to assess if the changes introduced in the design are risky. [64]

10. CONCLUSIONS

Nowadays, basic and advanced monitoring systems are present in surgeries and they allow to have an idea of what is going on with the patient at every moment and if the injected doses of drugs can generate adverse effects. Moreover, they can indicate if some additional drug is needed in order to attenuate these effects.

Being able to assess the probability to be in the optimal range of anesthesia from some input parameters can have a great impact in the surgical room. Furthermore, if this optimal effect is predicted in real time many adversities would be avoided.

AI is gaining popularity in almost all the surrounding environments. However, it has not the order of the day yet in operating rooms. The aim of this project is to introduce the idea to include in them.

The results show the efficiency of Logistic Regression to predict whether the optimal depth of anesthesia is achieved only considering propofol concentrations. Nonetheless, propofol is not enough to predict it, as the resulting ROC curves do not show excellent results. But as covariables are included, the values of the ROC curves increase, thus showing an improvement in the model. Specifically, when remifentanil and demographic variables are considered, the model has a higher predictive capability.

Although the results show that the model is accurate, it could not be launched into the market yet. There are some limitations of the project and some aspects which would have to be studied before trying to approve it as a SaDM.

First of all, it should be validated with external data, acquired after having built the model. This blink data would allow to evaluate better the performance of the model. Moreover, this model has been trained with a non-representative set of data; the whole database is formed by women so, future lines of this project could include men in the model and evaluating if the performance increases.

Moreover, this project shows the probability to be in the optimal range. But something which could make the model more attractive would be adding the probability to be in an underdose and overdose stage, which would help even more the anesthesiologist. This could be done not by dichotomizing the real BIS variable as it has been done but creating three labels: 0, 1 or 2 if there is underdose, optimal depth of anesthesia or overdose, respectively.

As explained in some sections above, the depth of anesthesia can be monitored by using other parameters apart from the BIS index. Introducing these parameters in the model, such as qCON, could lead to a better performance of the model as the number of variables directly related with the optimal level of anesthesia would increase, and so the data with which the model is trained and its performance.

Leaving now the results and focusing on the total costs of this project, an approximation has been done in Section 6. However, it is not a fixed cost due to the price of BIS VISTA monitoring system; its price ranges from 4000€ to 5500€ [65] but sometimes the commercial company and the hospital reach an agreement and both of them take benefit from it. It could happen that the hospital wants to buy a new monitor and gives BIS monitors back to the company to make the supply more

attractive. Or sometimes the company is interested in getting rid of some sensors and gives the monitors for free or with a more affordable price.

To conclude, it has been proved that the optimal level of anesthesia depends mainly on the quantity of injected drugs but also on the demographic variables, thus leading to an individualized attention, and AI could be introduced in the surgical environment as a tool to help both anesthesiologists and doctors to manage the surgery the best way possible.

11. REFERENCES

[1]- Brown, E. N., Lydic, R. and Schiff, N. D. (2010) 'General Anesthesia, Sleep, and Coma', New England Journal of Medicine. Massachusetts Medical Society, 363(27), pp. 2638–2650. doi: 10.1056/nejmra0808281.

[2]- Brown, E. N., Pavone, K. J., & Naranjo, M. (2018). Multimodal General Anesthesia: Theory and Practice. Anesthesia and analgesia, 127(5), 1246–1258. https://doi.org/10.1213/ANE.0000000000003668

[3] - Siddiqui BA, Kim PY. Anesthesia Stages. [Updated 2021 Mar 7]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK557596/

[4]- Smith G, D'Cruz JR, Rondeau B, et al. General Anesthesia for Surgeons. [Updated 2021 Apr 8]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK493199/

[5]- Gambús P.L., Hendrikx J.F.A. Personalized Anaestheisa. Targeting Physiological Systems for Optimal Effect. Cambridge University Press, 2020.

[6] Facco E. Hypnosis and anesthesia: back to the future. Minerva Anestesiol. 2016 Dec;82(12):1343-1356. Epub 2016 Aug 30. PMID: 27575449

[7] Sedación y analgesia: una revisión (no date). Available at: https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1409-00902002000100001 (Accessed: 28 May 2021).

[8]- Hendrickx, J. F. A. et al. (2008) 'Is synergy the rule? a review of anesthetic interactions producing hypnosis and immobility', Anesthesia and Analgesia. Lippincott Williams and Wilkins, 107(2), pp. 494–506. doi: 10.1213/ane.0b013e31817b859e.

[9] Domingo Lopez, J., Boris Brandstrup, K. and Reina, M. C. RELAJANTES MUSCULARES.

[10] López Blanco, S. and Andresen, M. (2021) Relajantes musculares. Available at: http://publicacionesmedicina.uc.cl/TemasMedicinaInterna/relajantes_musculares.html (Accessed: 29 May 2021).

[11] López Blanco, S. and Andersen, M. (no date) Relajantes musculares. Available at: http://publicacionesmedicina.uc.cl/TemasMedicinaInterna/relajantes_musculares.html (Accessed: 29 May 2021).

[12] - Hemmerling, T. M. (2009) 'Automated anesthesia', Current Opinion in Anaesthesiology, pp. 757–763. doi: 10.1097/ACO.0b013e328332c9b4.

[13]- Anderson, B. J. and Houghton, J. (2019) 'Total Intravenous Anesthesia and Target-Controlled Infusion', in A Practice of Anesthesia for Infants and Children. Elsevier, pp. 177-198.e3. doi: 10.1016/B978-0-323-42974-0.00008-2.

[14] Hemmerling, T. M. (2009) 'Automated anesthesia', Current Opinion in Anaesthesiology, pp. 757–763. doi: 10.1097/ACO.0b013e328332c9b4.

[15]- Al-Rifai, Z. and Mulvey, D. (2016) 'Principles of total intravenous anaesthesia: practical aspects of using total intravenous anaesthesia', BJA Education. Elsevier Ltd, 16(8), pp. 276–280. doi: 10.1093/bjaed/mkv074.

[16]- Azuero, B. K. (2013) SEDANTES, ANALGÉSICOS Y RELAJANTES MUSCULARES.

[17]- Galeotti, G. (2009) Farmacocinética del propofol en infusión.

[18]- Mcneir, D. A., Mainous, E. G. and Trieger, N. (1988) Propofol as an Intavenous Agent in General Anesthesia and Conscious Sedaton. Anesth Prog 35:147-151

[19]- Brown, E. N., Purdon, P. L., & Van Dort, C. J. (2011). General anesthesia and altered states of arousal: a systems neuroscience analysis. Annual review of neuroscience, 34, 601–628. https://doi.org/10.1146/annurev-neuro-060909-153200

[20]- Pérez Vela J.L., Guillén F., Sánchez P., Hernández I. Rubio M., Renes E., Perales N. (2004) Analgesia con remifentanilo en el postoperatorio de cirugía cardíaca. Medicina intensica 38(2), 57-64.

[21]- Airway Management: Use of succinylcholine or Rocuronium. American Association of Nurse Anesthetists.

[22]- Jain A, Wermuth HR, Dua A, et al. Rocuronium. [Updated 2020 Dec 23]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539888/

[23]- Smith G, D'Cruz JR, Rondeau B, et al. General Anesthesia for Surgeons. [Updated 2020 Aug 25]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK493199/

[24]- Siddiqui BA, Kim PY. Anesthesia Stages. [Updated 2021 Mar 7]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK557596/

[25]- Hight, D. F., Dadok, V. M., Szeri, A. J., García, P. S., Voss, L., & Sleigh, J. W. (2014). Emergence from general anesthesia and the sleep-manifold. Frontiers in systems neuroscience, 8, 146. https://doi.org/10.3389/fnsys.2014.00146

[26]- Tomic, J. I. E. et al. (2018) 'Recomendación Clínica: Disponibilidad y Uso de Monitorización Perioperatoria', Revista Chilena de Anestesia. Sociedad de Anestesiologia de Chile, 47(2), pp. 137–144. doi: 10.25237/revchilanestv47n02.12.

[27]- Esteban Álvarez-Reséndiz, G. et al. (no date) Monitoreo anestésico básico DEFINICIÓN DE MONITOREO, Supl. 1 Abril-Junio. Available at: http://www.medigraphic.com/rmawww.medigraphic.org.mx (Accessed: 29 May 2021).

[28]- Bartels, K., Esper, S. A. and Thiele, R. H. (2016) 'Blood pressure monitoring for the anesthesiologist: A practical review', Anesthesia and Analgesia. Lippincott Williams and Wilkins, 122(6), pp. 1866–1879. doi: 10.1213/ANE.0000000000001340.

[29]- Monitoring During Anesthesia | Michigan Medicine (no date). Available at: https://www.uofmhealth.org/health-library/rt1592 (Accessed: 29 May 2021).

[30]- Irma Guadalupe García-Colmenero, D. et al. (2018) Electroencefalografía para el anestesiólogo, consideraciones clínicas, Supl. 1 Abril-Junio. Available at: http://www.medigraphic.com/rmawww.medigraphic.org.mxwww.medigraphic.org.mx (Accessed: 29 May 2021).

[31]- Hagihira, S. (2015) 'Changes in the electroencephalogram during anaesthesia and their physiological basis', British Journal of Anaesthesia. Oxford University Press, pp. i27–i31. doi: 10.1093/bja/aev212.

[32]- Purdon, P. L. et al. (2015) 'Clinical electroencephalography for anesthesiologists', Anesthesiology. Lippincott Williams and Wilkins, pp. 937–960. doi: 10.1097/ALN.0000000000000841.

[33]- The Introductory Guide to EEG (Electroencephalography) - EMOTIV (no date). Available at: https://www.emotiv.com/eeg-guide/ (Accessed: 29 May 2021).

[34]- Hajat, Z., Ahmad, N. and Andrzejowski, J. (2017) 'The role and limitations of EEG-based depth of anaesthesia monitoring in theatres and intensive care', Anaesthesia. Blackwell Publishing Ltd, pp. 38–47. doi: 10.1111/anae.13739.

[35]- Lopez D.D. (2016) Bispectral index monitoring basics. Nursing made incredibly Easy! 14(2) pp.15-16

[36]- Espino Timón C. (2017) Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso.

[37]- Das, S., Dey, A. and Roy, N. (2015) Applications of Artificial Intelligence in Machine Learning: Review and Prospect, International Journal of Computer Applications.

[38]- Supervised vs. Unsupervised Learning: What's the Difference? | IBM (no date). Available at: https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning (Accessed: 29 May 2021).

[39]- Poma, S. and Mercedes, C. MODELO DE REGRESIÓN LOGÍSTICA.

[40]- Sperandei S. (2014). Understanding logistic regression analysis. Biochemia medica, 24(1), 12–18. https://doi.org/10.11613/BM.2014.003 AHORA 40 antes 45

[41]- Domínguez-Almendros, S., Benítez-Parejo, N. and Gonzalez-Ramirez, A. R. (2011) 'Logistic regression models', Allergologia et Immunopathologia. Elsevier, 39(5), pp. 295–305. doi: 10.1016/j.aller.2011.05.002.

[42]- Introduction to Machine Learning Model Evaluation | by Steve Mutuvi | Heartbeat (no date). Available at: https://heartbeat.fritz.ai/introduction-to-machine-learning-model-evaluation-fa859e1b2d7f (Accessed: 29 May 2021).

[43]- Mandrekar, J. N. (2010) 'Receiver operating characteristic curve in diagnostic test assessment', Journal of Thoracic Oncology. Lippincott Williams and Wilkins, 5(9), pp. 1315–1316. doi: 10.1097/JTO.0b013e3181ec173d.

[44]- Sperandei S. (2014). Understanding logistic regression analysis. Biochemia medica, 24(1), 12–18. https://doi.org/10.11613/BM.2014.003

[45]- Musizza, B., & Ribaric, S. (2010). Monitoring the depth of anaesthesia. Sensors (Basel, Switzerland), 10(12), 10896–10935. https://doi.org/10.3390/s101210896

[46]- Ben-Israel, N., Kliger, M., Zuckerman, G. et al. Monitoring the nociception level: a multi-parameter approach. J Clin Monit Comput 27, 659–668 (2013). https://doi-org.sire.ub.edu/10.1007/s10877-013-9487-9

[47]- Kang, H. (2015) 'Intraoperative nociception monitoring', Anesth Pain Med, 10, pp. 227–234. doi: 10.17085/apm.2015.10.4.227.

[48]- Vallverdú Ferrer M. (2014) Medición de la nocicepción con señal electroencefalográfica

[49]- Ledowski, T. (2019) 'Objective monitoring of nociception: a review of current commercial solutions', British Journal of Anaesthesia. Elsevier Ltd, pp. e312–e321. doi: 10.1016/j.bja.2019.03.024.

[50]- Li, Xuhong et al. (2021) 'Interpretable Deep Learning: Interpretation, Interpretability, Trustworthiness, and Beyond'. Available at: http://arxiv.org/abs/2103.10689 (Accessed: 29 May 2021).

[51]- BIS VISTATM Monitoring System OPERATING MANUAL (no date). Available at: www.aspectmedical.com (Accessed: 29 May 2021).

[52]- Infinity® Gamma XL. Available at: https://www.draeger.com/en_me/Products/Infinity-Gamma-XL (Accessed: 29 May 2021).

[53]- Conox® - Fresenius Kabi España. Available at: https://www.fresenius-kabi.com/es/productos/conox (Accessed: 29 May 2021).

[54]- Al-Rifai MBChB MPharm FRCA, Z. and Mulvey MBBS FRCA, D. (2015) 'Principles of total intravenous anaesthesia: basic pharmacokinetics and model descriptions'. doi: 10.1093/bjaceaccp/mkv021.

[55]- Drager Infinity Gamma XL Patient Monitor - Avante Health Solutions. Available at: https://avantehs.com/p/drager-infinity-gamma-xl-patient-monitor/11705 (Accessed: 29 May 2021).

[56]- Orchestra® Base Primea - Fresenius Kabi España. Available at: https://www.fresenius-kabi.com/es/productos/orchestra-base-primea (Accessed: 29 May 2021).

[57]- qCon 2000: a Smart Anaesthesia Monitoring - Quantium Medical. Available at: https://quantiummedical.com/products/qcon2000/ (Accessed: 29 May 2021).

[68]- Enthought Python Minimum Hardware Requirements – Enthought Knowledge Base (no date). Available at: https://support.enthought.com/hc/en-us/articles/204273874-Enthought-Python-Minimum-Hardware-Requirements (Accessed: 29 May 2021).

[59]- rugloopII© MO. Available at: http://www.demed.be/rugloopIIMoniOnly.htm (Accessed: 29 May 2021).

[60]- Microsoft Excel 2016 - media Specs – CNET. Available at: https://www.cnet.com/products/microsoft-excel-2016-media-06508501/ (Accessed: 29 May 2021).

[61]- [Win64] Monolix Suite 2020R1 – Lixoft. Available at: https://lixoft.com/download/win64-monolix-suite-2020r1/ (Accessed: 29 May 2021).

[62]- BOE.es - BOE-A-2018-16673 Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Available at: https://www.boe.es/buscar/doc.php?id=BOE-A-2018-16673 (Accessed: 29 May 2021).

[63]- Software as a Medical Device (SaMD) | FDA. Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd (Accessed: 29 May 2021).

[69]- REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - of 5 April 2017 - on medical devices, amending Directive 2001/ 83/ EC, Regulation (EC) No 178/ 2002 and Regulation (EC) No 1223/ 2009 and repealing Council Directives 90/ 385/ EEC and 93/ 42/ EEC (no date).

[64]- Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)-Discussion Paper and Request for Feedback. Available at: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm514737.pdf. (Accessed: 29 May 2021).

[65]- BISTM Brain Monitoring: Know the Costs and Understand the Benefits | Blog | HCP Resources | Medtronic (no date). Available at: https://hcpresources.medtronic.com/blog/bis-brain-monitoring-technology-know-the-costs-and-understand-the-benefits (Accessed: 29 May 2021).

APPENDIX

I. Python Code

a. Preprocessing

The first step is predicting BIS with the Monolix software. The initial data contained in different files, one per patient, is reduced 10 times and is introduced in the same Excel file, which will be used by Dr. Jaramillo to generate the predictions.

```
#Initial data is reduced 10 times: only 1 row each 10 is maintained.
import os
import pandas as pd
import numpy as np

paths=r"C:\Users\jreyort\Desktop\Marta\TFG\2.data_BIS"

files=os.listdir(paths)

#files=pd.read_excel(path, "Hoja1")
for index in range(len(files)):
    file= pd.read_excel("2.data_BIS/"+files[index], "Sheet1")

    all_ind=[]
    for element in range(len(file)):
        all_ind.append(element)
    ind_reduced=all_ind[0:len(all_ind):10]

    a=file.iloc[ind_reduced,0]#id
    b=file.iloc[ind_reduced,1]#time
    c=file.iloc[ind_reduced,2]#bis
    e=file.iloc[ind_reduced,4] #cp remi
    f=file.iloc[ind_reduced,5] #ce remi
    g=file.iloc[ind_reduced,6] #cp prop
    h=file.iloc[ind_reduced,7] #ce prop
    i=file.iloc[ind_reduced,8] #age
    j=file.iloc[ind_reduced,9] #height
    k=file.iloc[ind_reduced,10] #weight
    l=file.iloc[ind_reduced,11] #gender
    m=file.iloc[ind_reduced,12] #inf remi
    l=file.iloc[ind_reduced,13] #inf propo

    dat= {'ID_patient': a, 'Time(s)': b, 'BIS': c, 'CpREMI': e,
'CeREMI': f, 'CpPROP': g, 'CePROP': h, 'age': i, 'height': j,
'weight': k, 'gender': l, 'InfRateREMI': m, 'InfRatePROPO': l}
    df=pd.DataFrame(data=dat)
    df.to_excel("2.data_BIS/data_new_"+str(index)+".xlsx", index =
False, header=True)
```

```
#data is all grouped in the same dataframe

paths=r"C:\Users\jreyort\Desktop\Marta\TFG\4.first_data"

files=os.listdir(paths)

appended_data = []
for index in range(len(files)):
    file= pd.read_excel("4.first_data/"+files[index], "Sheet1")
    # store DataFrame in list
    appended_data.append(file)
appended_data = pd.concat(appended_data)
# write DataFrame to an excel sheet
appended_data.to_excel('appended_first_data.xlsx')
```

Once the BIS predictions are generated, the files that will be used to build the model are created. First of all, two binary columns are created from the real BIS; depending on whether BIS<45 and BIS>45 or BIS<55 and BIS>55. Then, 6 files are created, for models A and B of the three cases studied (propofol, propofol + remifentanil and propofol + remifentanil + demographic variables). The dynamics followed in all the cases are the same, so only model 1A is commented on the code.

```
import pandas as pd
import numpy as np
import os
import numpy as np

#we get the file with the BIS predictions (columns: id, time, bis,
predicted_bis)
path=r"C:\Users\jreyort\Desktop\Marta\TFG"
file_pred= pd.read_excel("predictions.xlsx", "predictions")

#original data for models 1, 2 and 3
id_patient=file_pred.iloc[:,0]
time=file_pred.iloc[:,1]
bis=file_pred.iloc[:,2]
pred=file_pred.iloc[:,3]

bis_45_100=[] #binary data for BIS<45 (0) and BIS>=45 (1), with
respect to REAL BIS
for value_1 in range(len(bis)):
    if 0<bis[value_1]<=44.9:
        bis_45_100.append(0)
    elif 44.9<bis[value_1]<=100:
        bis_45_100.append(1)
    else:
        bis_45_100.append(np.nan)

bis_0_55=[] #binary data for BIS>55 (0) and BIS<=55 (1), with respect
to REAL BIS
for value_2 in range(len(bis)):
    if 0<bis[value_2]<=54.9:
        bis_0_55.append(1)
    elif 54.9<bis[value_2]<=100:
        bis_0_55.append(0)
```

```
    else:
        bis_0_55.append(np.nan)

pred_bis_disc=[]
for value_3 in range(len(pred)):
    if 0<=pred[value_3]<=50:
        pred_bis_disc.append(0)
    elif 50<pred[value_3]<=100:
        pred_bis_disc.append(1)
    else:
        pred_bis_disc.append(np.nan)

data_new= {'ID_patient': id_patient, 'Time(s)': time, 'BIS': bis,
"BIS_up_45": bis_45_100, "BIS_down_55": bis_0_55, "Pred_BIS":pred,
'Pred_BIS_disc': pred_bis_disc}
df=pd.DataFrame(data=data_new)
df.to_excel("data_model.xlsx", index = False, header=True) #the file
is saved as an Excel

#We now create and save the files we will use to build the multiple
logistic regression models
file_old= pd.read_excel("data_pacientes.xlsx", "Sheet1")
file_model= pd.read_excel("data_model.xlsx", "Sheet1")

#New file with ce_remi column
bis=file_model.iloc[:,2]
bis_45_100=file_model.iloc[:,3]
bis_0_55=file_model.iloc[:,4]
pred=file_model.iloc[:,5]
pred_bis_disc=file_model.iloc[:,6] #no cal
ce_remi=file_old.iloc[:,7]

data_cp_remi= {'BIS': bis, "BIS_up_45": bis_45_100, "BIS_down_55":
bis_0_55, "Pred_BIS":pred, 'Pred_BIS_disc': pred_bis_disc, "Ce_ReMI":
ce_remi}
df=pd.DataFrame(data=data_cp_remi)
df.to_excel("data_model_cp_remi.xlsx", index = False, header=True)

#New file with ce_remi + demographic data column
bis=file_model.iloc[:,2]
bis_45_100=file_model.iloc[:,3]
bis_0_55=file_model.iloc[:,4]
pred=file_model.iloc[:,5]
pred_bis_disc=file_model.iloc[:,6]
ce_remi=file_old.iloc[:,7]
age=file_old.iloc[:,10]
height=file_old.iloc[:,11]
weight=file_old.iloc[:,12]

data_cp_remi_demo= {'BIS': bis, "BIS_up_45": bis_45_100,
"BIS_down_55": bis_0_55, "Pred_BIS":pred, 'Pred_BIS_disc':
pred_bis_disc, "Ce_ReMI": ce_remi, "Age": age, "Height": height,
"Weight": weight}
df=pd.DataFrame(data=data_cp_remi_demo)
df.to_excel("data_model_remi_demo.xlsx", index = False, header=True)

#---------------------------------
```

```
#MODEL 1 (PROPOFOL)

#MODEL 1a
paths=r"C:\Users\jreyort\Desktop\Marta\TFG"
file= pd.read_excel("data_model.xlsx", "Sheet1")

#selecting only the data from the model A (pred_bis<50)
bis_50_down = file.loc[:, 'Pred_BIS_disc'] == 0
df_bis_down = file.loc[bis_50_down]
print("Longitud de los datos: ", len(df_bis_down))

#original data from model A saved to excel
df_bis_down.to_excel("data_down50.xlsx", index = False, header=True)

#originally, there are 30636 recordings but only 5184 are between
45<pred_BIS<50
file2= pd.read_excel("data_down50.xlsx", "Sheet1")
bis1 = file2.loc[:, 'BIS_up_45'] == 1
bis1_1 = file2.loc[bis1]

#we reduce the data between 0<pred_BIS<45 to 9000, to balance it with
the prior one related to the optimal level
bis2 = file2.loc[:, 'BIS_up_45'] == 0
bis2_2 = file2.loc[bis2]
bis2_2.iloc[0:9000,:]

#we concatenate the selected data, 0<pred_BIS<45 and 45<pred_BIS<50
result = pd.concat([bis1_1, bis2_2.iloc[0:9000,:]])

#we shuffle the data
result_shuffled=result.sample(frac=1).reset_index(drop=True)
result_shuffled.to_excel("data_down50_new_model1.xlsx", index = False,
header=True)

#the resulting file is saved as an Excel
file_new2= pd.read_excel("data_down50_new_model1.xlsx", "Sheet1")

#------------------------------
#MODEL 1B, the same as in model A
paths=r"C:\Users\jreyort\Desktop\Marta\TFG"
file= pd.read_excel("data_model.xlsx", "Sheet1")

#selecting only the data from the model B (pred_bis>50)
bis_50_up = file.loc[:, 'Pred_BIS_disc'] == 1
df_bis_up = file.loc[bis_50_up]
print("Longitud de los datos: ", len(df_bis_up))

#original data from model B saved to excel
df_bis_up.to_excel("data_up50.xlsx", index = False, header=True)

#originally, there are 6118 recordings but only 1016 are between
50<pred_BIS<55
file2b= pd.read_excel("data_up50.xlsx", "Sheet1")
bis1b = file2b.loc[:, 'BIS_down_55'] == 1
bis1_1b = file2b.loc[bis1b]
```

```
#we reduce the data between 0<pred_BIS<45 to 9000, to balance it with
the prior one related to the optimal level
bis2b = file2b.loc[:, 'BIS_down_55'] == 0
bis2_2b = file2b.loc[bis2b]
bis2_2b.iloc[0:1016,:]

resultb = pd.concat([bis1_1b, bis2_2b.iloc[0:1016,:]])

result_shuffledb=result.sample(frac=1).reset_index(drop=True)

result_shuffledb.to_excel("data_up50_new_model1.xlsx", index = False,
header=True)

file_new1= pd.read_excel("data_up50_new_model1.xlsx", "Sheet1")


#-------------------------------------
#MODEL 2 (CE_REMI)
paths=r"C:\Users\jreyort\Desktop\Marta\TFG"
file_model2= pd.read_excel("data_model_cp_remi.xlsx", "Sheet1")

#MODEL 2a
bis_50_down_model2 = file_model2.loc[:, 'Pred_BIS_disc'] == 0
df_bis_down_model2 = file_model2.loc[bis_50_down_model2]
print("Longitud de los datos: ", len(df_bis_down_model2)) #30636
recordings

df_bis_down_model2.to_excel("data_down50_model2.xlsx", index = False,
header=True)

#the data is reduced to avoid overfitting. We eliminate only some data
from the first part: 0<pred_bis<45
file2_model= pd.read_excel("data_down50_model2.xlsx", "Sheet1")
bis1_model2 = file2_model.loc[:, 'BIS_up_45'] == 1
bis1_1_model2 = file2_model.loc[bis1_model2]
len(bis1_1_model2) #5184 recordings in 45<pred_bis<50

bis2_model2 = file2_model.loc[:, 'BIS_up_45'] == 0
bis2_2_model2 = file2_model.loc[bis2_model2]
bis2_2_model2.iloc[0:9000,:]

result_model2 = pd.concat([bis1_1_model2,
bis2_2_model2.iloc[0:9000,:]])

result_shuffled_model2=result_model2.sample(frac=1).reset_index(drop=T
rue)

#only the desired columns to build the model are selected; the rest
are eliminated
result_shuffled_model2.drop(['BIS', 'BIS_down_55', 'Pred_BIS_disc'],
axis = 'columns', inplace=True)

#the file is saved
result_shuffled_model2.to_excel("data_down50_new_model2.xlsx", index =
False, header=True)

#--------------------------------
```

```
#MODEL 2b
bis_50_up_model2 = file_model2.loc[:, 'Pred_BIS_disc'] == 1
df_bis_up_model2 = file_model2.loc[bis_50_up_model2]
print("Longitud de los datos: ", len(df_bis_up_model2)) #length: 6118

df_bis_up_model2.to_excel("data_up50_model2.xlsx", index = False,
header=True)

file2_model= pd.read_excel("data_up50_model2.xlsx", "Sheet1")
bis1_model2 = file2_model.loc[:, 'BIS_down_55'] == 1
bis1_1_model2 = file2_model.loc[bis1_model2]
len(bis1_1_model2) #only 1016 recordings are between 50<pred_bis<55,
from 6118

bis2_model2 = file2_model.loc[:, 'BIS_down_55'] == 0
bis2_2_model2 = file2_model.loc[bis2_model2]
bis2_2_model2.iloc[0:1016,:]

result_model2 = pd.concat([bis1_1_model2,
bis2_2_model2.iloc[0:1016,:]])

result_shuffled_model2=result_model2.sample(frac=1).reset_index(drop=T
rue)

#only the desired columns to build the model are selected
result_shuffled_model2.drop(['BIS', 'BIS_up_45', 'Pred_BIS_disc'],
axis = 'columns', inplace=True)

result_shuffled_model2.to_excel("data_up50_new_model2.xlsx", index =
False, header=True)

file_new1_model2= pd.read_excel("data_up50_new_model2.xlsx", "Sheet1")


#--------------------------
#MODEL 3 (CE_REMI + DEMO)
paths=r"C:\Users\jreyort\Desktop\Marta\TFG"
file_model3= pd.read_excel("data_model_remi_demo.xlsx", "Sheet1")

#MODEL 3a
bis_50_down_model3 = file_model3.loc[:, 'Pred_BIS_disc'] == 0
df_bis_down_model3 = file_model3.loc[bis_50_down_model3]
print("Longitud de los datos: ", len(df_bis_down_model3))

df_bis_down_model3.to_excel("data_down50_model3.xlsx", index = False,
header=True)

file3_model= pd.read_excel("data_down50_model3.xlsx", "Sheet1")
bis1_model3 = file3_model.loc[:, 'BIS_up_45'] == 1
bis1_1_model3 = file3_model.loc[bis1_model3]

bis2_model3 = file3_model.loc[:, 'BIS_up_45'] == 0
bis2_2_model3 = file3_model.loc[bis2_model3]
bis2_2_model3.iloc[0:9000,:]

result_model3 = pd.concat([bis1_1_model3,
bis2_2_model3.iloc[0:9000,:]])
```

```
result_model3

result_shuffled_model3=result_model3.sample(frac=1).reset_index(drop=T
rue)

#we eliminate the undesired columns, leaving only the ones needed for
the model
result_shuffled_model3.drop(['BIS', 'BIS_down_55', 'Pred_BIS_disc'],
axis = 'columns', inplace=True)

result_shuffled_model3.to_excel("data_down50_new_model3.xlsx", index =
False, header=True)

file_new2_model3= pd.read_excel("data_down50_new_model3.xlsx",
"Sheet1")

#--------------------------------
#MODEL 3b
paths=r"C:\Users\jreyort\Desktop\Marta\TFG"
file_model3= pd.read_excel("data_model_remi_demo.xlsx", "Sheet1")

bis_50_up_model3 = file_model3.loc[:, 'Pred_BIS_disc'] == 1
df_bis_up_model3 = file_model3.loc[bis_50_up_model3]

df_bis_up_model3.to_excel("data_up50_model3.xlsx", index = False,
header=True)

file3_model= pd.read_excel("data_up50_model3.xlsx", "Sheet1")
bis1_model3 = file3_model.loc[:, 'BIS_down_55'] == 1
bis1_1_model3 = file3_model.loc[bis1_model3]

bis2_model3 = file3_model.loc[:, 'BIS_down_55'] == 0
bis2_2_model3 = file3_model.loc[bis2_model3]
bis2_2_model3.iloc[0:1016,:]

result_model3 = pd.concat([bis1_1_model3,
bis2_2_model3.iloc[0:1016,:]])

result_shuffled_model3=result_model3.sample(frac=1).reset_index(drop=T
rue)

result_shuffled_model3.drop(['BIS', 'BIS_up_45', 'Pred_BIS_disc'],
axis = 'columns', inplace=True)

result_shuffled_model3.to_excel("data_up50_new_model3.xlsx", index =
False, header=True)

file_new1_model3= pd.read_excel("data_up50_new_model3.xlsx", "Sheet1")
```

b. Model building

The last step is building the model. As in the previous section, the dynamics for building the model is pretty similar in all the cases, so only the first case is commented.

```
#import libraries
import pandas as pd
import numpy as np
```

```
import os
import matplotlib.pyplot as plt
import numpy as np
from sklearn.model_selection import train_test_split
import statsmodels.api as sm
from sklearn.metrics import confusion_matrix, accuracy_score,
roc_curve,auc, roc_auc_score

#MODEL 1A PROPOFOL (pred_BIS<50)
file_new_1b= pd.read_excel("data_down50_new_model1.xlsx", "Sheet1")

#selecting dependent and independent variables
X_1b = file_new_1b[["Pred_BIS"]] #pred bis < 50
y_1b = file_new_1b['BIS_up_45'] #bis real>45

#dividing data in test and train sets
X_train_1b, X_test_1b, y_train_1b, y_test_1b =
train_test_split(X_1b.values.reshape(-1,1), y_1b.values.reshape(-1,1),
train_size   = 0.75, shuffle= True, random_state=0)

#model training
X_train_1b=sm.add_constant(X_train_1b, prepend=True)
#building the model
modelo_1b=sm.Logit(endog=y_train_1b, exog=X_train_1b,)
#fitting the model
modelo_1b=modelo_1b.fit()
print(modelo_1b.summary())


#model testing
X_test_1b = sm.add_constant(X_test_1b, prepend=True)
predicciones_test_1b= modelo_1b.predict(exog = X_test_1b)
plt.figure()
plt.plot(X_test_1b,predicciones_test_1b,'.')
plt.title("Predictions model A (propofol)")
plt.xlim([5,52])
plt.xlabel("predicted BIS")
plt.ylabel("probability")

#model validation
fpr_1b, tpr_1b, threshold_1b = roc_curve(y_test_1b,
predicciones_test_1b)
auroc_1b = auc(fpr_1b, tpr_1b)
gmeans_1b = np.sqrt(tpr_1b * (1-fpr_1b))
ix_1b = np.argmax(gmeans_1b)
print('Best Threshold=%f, G-Mean=%.3f' % (threshold_1b[ix_1b],
gmeans_1b[ix_1b]))
plt.plot(fpr_1b, tpr_1b, 'b', label = 'AUC = %0.2f' % auroc_1b)
plt.title('ROC (propofol model A)')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('Sensitivity')
plt.xlabel('1 - Specificity')
plt.scatter(fpr_1b[ix_1b], tpr_1b[ix_1b], marker='o', color='black',
label='Best')
```

```
plt.show()
print('roc_auc_score for Logistic Regression: ',
roc_auc_score(y_test_1b, predicciones_test_1b))

#------------------------
#MODEL 1B PROPOFOL (pred_BIS>50). Same steps as before
file_new_1a= pd.read_excel("data_up50_new_model1.xlsx", "Sheet1")

#selecting dependent and independent variables
X_1a = file_new_1a[["Pred_BIS"]] #pred bis > 50
y_1a = file_new_1a['BIS_down_55'] #bis real<55

#dividing data in test and train sets
X_train_1a, X_test_1a, y_train_1a, y_test_1a =
train_test_split(X_1a.values.reshape(-1,1), y_1a.values.reshape(-1,1),
train_size   = 0.75, shuffle= True, random_state=0)

#model training, building and fitting
X_train_1a=sm.add_constant(X_train_1a, prepend=True)
modelo_1a=sm.Logit(endog=y_train_1a, exog=X_train_1a,)
modelo_1a=modelo_1a.fit()
print(modelo_1a.summary())


#model testing and validation
X_test_1a = sm.add_constant(X_test_1a, prepend=True)
predicciones_test_1a = modelo_1a.predict(exog = X_test_1a)
fpr_1a, tpr_1a, threshold_1a = roc_curve(y_test_1a,
predicciones_test_1a)
auroc_1a = auc(fpr_1a, tpr_1a)
gmeans_1a = np.sqrt(tpr_1a * (1-fpr_1a))
ix_1a = np.argmax(gmeans_1a)
print('Best Threshold=%f, G-Mean=%.3f' % (threshold_1a[ix_1a],
gmeans_1a[ix_1a]))
plt.plot(fpr_1a, tpr_1a, 'b', label = 'AUC = %0.2f' % auroc_1a)
plt.title('ROC (propofol model B)')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('Sensitivity')
plt.xlabel('1 - Specificity')
plt.scatter(fpr_1a[ix_1a], tpr_1a[ix_1a], marker='o', color='black',
label='Best')
plt.show()
print('roc_auc_score for Logistic Regression: ',
roc_auc_score(y_test_1a, predicciones_test_1a))

plt.figure()
plt.plot(X_test_1a,predicciones_test_1a,'.')
plt.title("Predictions model B (propofol) ")
plt.xlim([45,105])
plt.xlabel("predicted BIS")
plt.ylabel("probability")

#-------------------------------------------------------
#MODEL 2A PROPOFOL + REMI (pred_BIS<50)
```

```
file_new_2b= pd.read_excel("data_down50_new_model2.xlsx", "Sheet1")

#the variables of the model are selected
X_2b = file_new_2b.drop(columns= "BIS_up_45") #pred bis < 50 + Ce_remi
y_2b = file_new_2b['BIS_up_45'] #bis real>45

X_train_2b, X_test_2b, y_train_2b, y_test_2b = train_test_split(X_2b,
y_2b.values.reshape(-1,1), train_size   = 0.75, shuffle= True,
random_state=0)

X_train_2b=sm.add_constant(X_train_2b, prepend=True)
modelo_2b=sm.Logit(endog=y_train_2b, exog=X_train_2b,)
modelo_2b=modelo_2b.fit()
print(modelo_2b.summary())

X_test_2b = sm.add_constant(X_test_2b, prepend=True)
predicciones_test_2b = modelo_2b.predict(exog = X_test_2b)

fpr_2b, tpr_2b, threshold_2b = roc_curve(y_test_2b,
predicciones_test_2b)
auroc_2b = auc(fpr_2b, tpr_2b)
gmeans_2b = np.sqrt(tpr_2b * (1-fpr_2b))
ix_2b = np.argmax(gmeans_2b)
print('Best Threshold=%f, G-Mean=%.3f' % (threshold_2b[ix_2b],
gmeans_2b[ix_2b]))
plt.plot(fpr_2b, tpr_2b, 'b', label = 'AUC = %0.2f' % auroc_2b)
plt.title('ROC (propofol + remifentanil model A)')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('Sensitivity')
plt.xlabel('1 - Specificity')
plt.scatter(fpr_2b[ix_2b], tpr_2b[ix_2b], marker='o', color='black',
label='Best')
plt.show()
print('roc_auc_score for Logistic Regression: ',
roc_auc_score(y_test_2b, predicciones_test_2b))

plt.figure()
plt.plot(X_test_2b,predicciones_test_2b,'.')
plt.title("Predictions model A (propofol + Ce_Remi)")
plt.xlim([10,52])
plt.xlabel("predicted BIS")
plt.ylabel("probability")

#----------------------
#MODEL 2B PROPOFOL + REMI (pred_BIS>50)
paths=r"C:\Users\jreyort\Desktop\Marta\TFG"
file_new_2a= pd.read_excel("data_up50_new_model2.xlsx", "Sheet1")

X_2a = file_new_2a.drop(columns= "BIS_down_55") #pred bis > 50
y_2a = file_new_2a['BIS_down_55'] #bis real<55

X_train_2a, X_test_2a, y_train_2a, y_test_2a = train_test_split(X_2a,
y_2a.values.reshape(-1,1), train_size   = 0.75, shuffle= True,
random_state=0)
```

```
X_train_2a=sm.add_constant(X_train_2a, prepend=True)
modelo_2a=sm.Logit(endog=y_train_2a, exog=X_train_2a,)
modelo_2a=modelo_2a.fit()
print(modelo_2a.summary())

X_test_2a = sm.add_constant(X_test_2a, prepend=True)
predicciones_test_2a = modelo_2a.predict(exog = X_test_2a)
fpr_2a, tpr_2a, threshold_2a = roc_curve(y_test_2a,
predicciones_test_2a)
auroc_2a = auc(fpr_2a, tpr_2a)
gmeans_2a = np.sqrt(tpr_2a * (1-fpr_2a))
ix_2a = np.argmax(gmeans_2a)
print('Best Threshold=%f, G-Mean=%.3f' % (threshold_2a[ix_2a],
gmeans_2a[ix_2a]))

plt.plot(fpr_2a, tpr_2a, 'b', label = 'AUC = %0.2f' % auroc_2a)
plt.title('ROC (propofol + remifentanil model B)')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('Sensitivity')
plt.xlabel('1 - Specificity')
plt.scatter(fpr_2a[ix_2a], tpr_2a[ix_2a], marker='o', color='black',
label='Best')
plt.show()

print('roc_auc_score for Logistic Regression: ',
roc_auc_score(y_test_2a, predicciones_test_2a))
plt.figure()
plt.plot(X_test_2a,predicciones_test_2a,'.')
plt.title("Predictions model B (propofol + Ce_Remi)")
plt.xlim([45,105])
plt.xlabel("predicted BIS")
plt.ylabel("probability")

#-----------------------
#MODEL 3A PROPOFOL + REMI + DEMOGRAPHIC VAR (pred_BIS<50)
file_new_3b= pd.read_excel("data_down50_new_model3.xlsx", "Sheet1")

X_3b = file_new_3b.drop(columns= "BIS_up_45") #pred bis < 50 + Ce_remi
+ demographic data
y_3b = file_new_3b['BIS_up_45'] #bis real>45

X_train_3b, X_test_3b, y_train_3b, y_test_3b = train_test_split(X_3b,
y_3b.values.reshape(-1,1), train_size   = 0.75, shuffle= True,
random_state=0)

X_train_3b=sm.add_constant(X_train_3b, prepend=True)
modelo_3b=sm.Logit(endog=y_train_3b, exog=X_train_3b,)
modelo_3b=modelo_3b.fit()
print(modelo_3b.summary())

X_test_3b = sm.add_constant(X_test_3b, prepend=True)
predicciones_test_3b = modelo_3b.predict(exog = X_test_3b)
```

```
fpr_3b, tpr_3b, threshold_3b = roc_curve(y_test_3b,
predicciones_test_3b)
auroc_3b = auc(fpr_3b, tpr_3b)
gmeans_3b = np.sqrt(tpr_3b * (1-fpr_3b))
ix_3b = np.argmax(gmeans_3b)
print('Best Threshold=%f, G-Mean=%.3f' % (threshold_3b[ix_3b],
gmeans_3b[ix_3b]))
plt.plot(fpr_3b, tpr_3b, 'b', label = 'AUC = %0.2f' % auroc_3b)
plt.title('ROC (propofol + remifentanil + demographic variables model
A)')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('Sensitivity')
plt.xlabel('1 - Specificity')
plt.scatter(fpr_3b[ix_3b], tpr_3b[ix_3b], marker='o', color='black',
label='Best')
plt.show()
print('roc_auc_score for Logistic Regression: ',
roc_auc_score(y_test_3b, predicciones_test_3b))
plt.figure()
plt.plot(X_test_3b["Pred_BIS"],predicciones_test_3b,'.')
plt.title("Predictions model A (propofol + Ce_Remi + demographic
var)")
plt.xlim([0,53])
plt.xlabel("predicted BIS")
plt.ylabel("probability")

#-----------------------------
#MODEL 3B PROPOFOL + REMI + DEMOGRAPHIC VAR (pred_BIS>50)
file_new_3a= pd.read_excel("data_up50_new_model3.xlsx", "Sheet1")

X_3a = file_new_3a.drop(columns= "BIS_down_55") #pred bis > 50 +
ce_remi+ demographic data
y_3a = file_new_3a['BIS_down_55'] #bis real<55

X_train_3a, X_test_3a, y_train_3a, y_test_3a = train_test_split(X_3a,
y_3a.values.reshape(-1,1), train_size   = 0.75, shuffle= True,
random_state=0)

X_train_3a=sm.add_constant(X_train_3a, prepend=True)
modelo_3a=sm.Logit(endog=y_train_3a, exog=X_train_3a,)
modelo_3a=modelo_3a.fit()
print(modelo_3a.summary())

X_test_3a = sm.add_constant(X_test_3a, prepend=True)
predicciones_test_3a = modelo_3a.predict(exog = X_test_3a)

fpr_3a, tpr_3a, threshold_3a = roc_curve(y_test_3a,
predicciones_test_3a)
auroc_3a = auc(fpr_3a, tpr_3a)
gmeans_3a = np.sqrt(tpr_3a * (1-fpr_3a))
ix_3a = np.argmax(gmeans_3a)
print('Best Threshold=%f, G-Mean=%.3f' % (threshold_3a[ix_3a],
gmeans_3a[ix_3a]))
plt.plot(fpr_3a, tpr_3a, 'b', label = 'AUC = %0.2f' % auroc_3a)
```

```
plt.title('ROC (propofol + remifentanil + demographic variables model
B)')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('Sensitivity')
plt.xlabel('1 - Specificity')
plt.scatter(fpr_3a[ix_3a], tpr_3a[ix_3a], marker='o', color='black',
label='Best')
plt.show()
print('roc_auc_score for Logistic Regression: ',
roc_auc_score(y_test_3a, predicciones_test_3a))
plt.figure()
plt.plot(X_test_3a["Pred_BIS"],predicciones_test_3a,'.')
plt.title("Predictions model B (propofol + Ce_Remi + demographic
var)")
plt.xlim([47,105])
plt.xlabel("predicted BIS")
plt.ylabel("probability")
```

II. Parameters of the model

This second section shows the results obtained for each model. They were already introduced in the detail engineering but here appear in a more detailed way.

Model A: propofol

```
Optimization terminated successfully.
         Current function value: 0.606315
         Iterations 5
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                10638
Model:                          Logit   Df Residuals:                    10636
Method:                           MLE   Df Model:                            1
Date:                Sat, 29 May 2021   Pseudo R-squ.:                 0.07572
Time:                        14:13:09   Log-Likelihood:                -6450.0
converged:                       True   LL-Null:                       -6978.4
Covariance Type:            nonrobust   LLR p-value:                8.007e-232
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -4.0704      0.119    -34.181      0.000      -4.304      -3.837
x1             0.0978      0.003     30.487      0.000       0.092       0.104
==============================================================================
```

Model B: propofol

```
Optimization terminated successfully.
         Current function value: 0.594338
         Iterations 6
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                 1524
Model:                          Logit   Df Residuals:                     1522
Method:                           MLE   Df Model:                            1
Date:                Wed, 26 May 2021   Pseudo R-squ.:                  0.1425
Time:                        23:13:20   Log-Likelihood:                -905.77
converged:                       True   LL-Null:                       -1056.3
Covariance Type:            nonrobust   LLR p-value:                 1.993e-67
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          5.0237      0.343     14.631      0.000       4.351       5.697
x1            -0.0797      0.006    -14.411      0.000      -0.091      -0.069
==============================================================================
```

Model A: Propofol + remifentanil

```
Optimization terminated successfully.
         Current function value: 0.568429
         Iterations 6
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                10638
Model:                          Logit   Df Residuals:                    10635
Method:                           MLE   Df Model:                            2
Date:                Sat, 29 May 2021   Pseudo R-squ.:                  0.1357
Time:                        14:17:03   Log-Likelihood:                -6046.9
converged:                       True   LL-Null:                       -6996.0
Covariance Type:            nonrobust   LLR p-value:                     0.000
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -1.3748      0.145     -9.476      0.000      -1.659      -1.090
Pred_BIS       0.0655      0.003     19.540      0.000       0.059       0.072
Ce_ReMI       -0.6491      0.023    -28.112      0.000      -0.694      -0.604
==============================================================================
```

Model B: Propofol + remifentanil

```
Optimization terminated successfully.
         Current function value: 0.518136
         Iterations 6
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                 1524
Model:                          Logit   Df Residuals:                     1521
Method:                           MLE   Df Model:                            2
Date:                Sat, 29 May 2021   Pseudo R-squ.:                  0.2524
Time:                        14:15:25   Log-Likelihood:                -789.64
converged:                       True   LL-Null:                       -1056.3
Covariance Type:            nonrobust   LLR p-value:                1.564e-116
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          4.4349      0.409     10.831      0.000       3.632       5.237
Pred_BIS      -0.0865      0.007    -12.915      0.000      -0.100      -0.073
Ce_ReMI        1.1423      0.095     12.059      0.000       0.957       1.328
==============================================================================
```

Model A: Propofol + remifentanil + demographic variables

```
Optimization terminated successfully.
         Current function value: 0.496724
         Iterations 6
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                10638
Model:                          Logit   Df Residuals:                    10632
Method:                           MLE   Df Model:                            5
Date:                Sat, 29 May 2021   Pseudo R-squ.:                  0.2465
Time:                        14:20:56   Log-Likelihood:                -5284.2
converged:                       True   LL-Null:                       -7012.6
Covariance Type:            nonrobust   LLR p-value:                     0.000
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -1.8748      0.706     -2.656      0.008      -3.258      -0.491
Pred_BIS       0.0911      0.004     23.074      0.000       0.083       0.099
Ce_ReMI       -0.6915      0.025    -27.193      0.000      -0.741      -0.642
Age           -0.0194      0.002     -9.553      0.000      -0.023      -0.015
Height        -0.0344      0.004     -7.931      0.000      -0.043      -0.026
Weight         0.0949      0.003     33.960      0.000       0.089       0.100
==============================================================================
```

Model B: Propofol + remifentanil + demographic variables

```
Optimization terminated successfully.
         Current function value: 0.502538
         Iterations 7
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                 1524
Model:                          Logit   Df Residuals:                     1518
Method:                           MLE   Df Model:                            5
Date:                Sat, 29 May 2021   Pseudo R-squ.:                  0.2750
Time:                        14:19:23   Log-Likelihood:                -765.87
converged:                       True   LL-Null:                       -1056.3
Covariance Type:            nonrobust   LLR p-value:                2.730e-123
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -2.8895      2.131     -1.356      0.175      -7.066       1.287
Pred_BIS      -0.0891      0.007    -12.783      0.000      -0.103      -0.075
Ce_ReMI        0.8551      0.091      9.412      0.000       0.677       1.033
Age            0.0380      0.006      6.467      0.000       0.026       0.049
Height         0.0310      0.012      2.509      0.012       0.007       0.055
Weight         0.0135      0.006      2.156      0.031       0.001       0.026
==============================================================================
```