Computers in Biology and Medicine 118 (2020) 103645

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Machine learning based method for the evaluation of the Analgesia Nociception Index in the assessment of general anesthesia

José M. Gonzalez-Cava [a,*], Rafael Arnay [a], Ana León [b], María Martín [b], José A. Reboso [b], José Luis Calvo-Rolle [c], Juan Albino Mendez-Perez [a]

[a] Departamento de Ingeniería Informática y Sistemas. Universidad de La Laguna (ULL), 38200, La Laguna, Tenerife, Spain
[b] Hospital Universitario de Canarias, San Cristóbal de La Laguna, Tenerife, Spain
[c] Department of Industrial Engineering. University of A Coruña, Ferrol, A Coruña, Spain

ARTICLE INFO

Keywords:
Anesthesia
Analgesia assessment
Analgesia nociception index
Machine learning
Opioid titration
Support vector machine

ABSTRACT

Measuring the level of analgesia to adapt the opioids infusion during anesthesia to the real needs of the patient is still a challenge. This is a consequence of the absence of a specific measure capable of quantifying the nociception level of the patients. Unlike existing proposals, this paper aims to evaluate the suitability of the Analgesia Nociception Index (ANI) as a guidance variable to replicate the decisions made by the experts when a modification of the opioid infusion rate is required. To this end, different machine learning classifiers were trained with several sets of clinical features. Data for training were captured from 17 patients undergoing cholecystectomy surgery. Satisfactory results were obtained when including information about minimum values of ANI for predicting a change of dose. Specifically, a higher efficiency of the Support Vector Machine (SVM) classifier was observed compared with the situation in which the ANI index was not included: accuracy: 86.21% (83.62%–87.93%), precision: 86.11% (83.78%–88.57%), recall: 91.18% (88.24%–91.18%), specificity: 79.17% (75%–83.33%), AUC: 0.89 (0.87–0.90) and kappa index: 0.71 (0.66–0.75). The results of this research evidenced that including information about the minimum values of ANI together with the hemodynamic information outperformed the decisions made regarding only non-specific traditional signs such as heart rate and blood pressure. In addition, the analysis of the results showed that including the ANI monitor in the decision making process may anticipate a dose change to prevent hemodynamic events. Finally, the SVM was able to perform accurate predictions when making different decisions commonly observed in the clinical practice.

1. Introduction

Pain can be defined as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" [1]. The presence of subjective psychological aspects regarding pain makes it difficult to find efficient methods and techniques for pain measurement and treatment. It has been one of the main problems when trying to define general protocols for the delivery of analgesics. The absence of pain is specifically an important issue during surgeries, in which physicians should ensure an accurate level of analgesia. During anesthesia, nociception may be considered a pain measurement as it derives from the activation of nociceptors due to physiological processes [2]. One of the main trends in analgesia has been the evaluation of the nociception-antinociception balance. As a matter of fact, different devices have been recently presented as reliable tools to measure nociception [3].

Nowadays, however, there is not any accepted standard practice in order to supply analgesic drug during anesthesia. Traditional protocols for the delivery of opioids have been based on indirect signs, such as movement, presence of tachycardia, sweat or lacrimation [4]. As a result, the decision-making process during the anesthesia practice mainly relies on the expertise of the anesthesiologist. According to the US Institute of Medicine, 80% of patients who undergo surgeries report postoperative pain, even reaching extreme pain levels [5]. Inadequate levels of analgesia in patients undergoing surgery may result in risk of overdosing, risk of post-operative hyperalgesia and may increase the time of recovery after the surgery [6]. In addition, the presence of acute pain during surgery is related to the development of chronical pain [7].

Recent approaches for opioid titration have included the use of new monitors able to measure the nociceptive activity during surgery.

* Corresponding author. Avda. Astrofísico Francisco Sánchez s/n, 38206, La Laguna, Tenerife, Spain.
E-mail address: jgonzalc@ull.edu.es (J.M. Gonzalez-Cava).

https://doi.org/10.1016/j.compbiomed.2020.103645
Received 17 October 2019; Received in revised form 21 January 2020; Accepted 3 February 2020
Available online 5 February 2020

Different research has been conducted using the Analgesia Nociception Index (ANI) monitor [8–12]. This monitor makes a Heart Rate Variability (HRV) analysis to measure the effect of the Respiratory Sinus Arrhythmia (RSA). Very promising results have been reached when using the ANI monitor as guidance variable for opioids titration [9,10]. A more sensitivity variable to stimuli and a lower drug consumption have been observed compared to traditional variables [13–15]. However, more research is needed to ensure the reliability of the ANI index for the analgesia management.

Considering the current problem in this field, the main objective of this research is the evaluation of the ANI monitor as a device capable of providing valuable information for the guidance of the analgesic drug titration during anesthesia. Particularly, unlike previous proposals, we present a new point of view to study whether it is possible to replicate the decisions of the experts in critical situations where a modification of dose is required. To this end, we have analyzed if the use of the ANI index in the decision making process can outperform the assessment of opioids traditionally based on non-specific signs such as heart rate and arterial pressure. This research presents an evaluation of the performance reached by different machine learning classifiers to predict the changes of dose when including different features in the decision making process.

Regarding the main objective of the study, our hypothesis is that a variation in the remifentanil infusion rate evidences a bad analgesia level of the patient. Therefore, dichotomous qualitative variables representing "Increment of drug" and "Decrement of drug" decisions will be used for the data labelling. Then, the performance of the different predictors will be discussed. Finally, the performance of the synthesized two-class classifier will be analyzed under different scenarios observed in the clinical practice. Thus, unlike previous research, this study constitutes a new alternative for the evaluation of the ANI monitor, specifically focused on the decision making process for the analgesic drug titration. The presented methodology based on machine learning turns this approach into a first step towards the development of standardization of the analgesia management during anesthesia. Furthermore, this proposal could be considered for the development of oncoming intelligent controllers for the automation of analgesia.

1.1. Related works

Monitoring nociception is a challenge to lower the incidence of acute postoperative pain and the move towards a more automated approach to analgesia and anesthesia [16]. Main trend focuses on the development of new devices for the evaluation of the nociception-antinociception balance. These monitors claim a reduction of the postoperative pain together with a lower consumption of the analgesic drug compared with the traditionally used vital signs, including blood pressure and heart rate [17]. These devices are based on the detection of clinical signs related to the reaction to nociception. Among all commercially available options, the Analgesia Nociception Index has been widely studied in the clinical practice. Previous research aimed a clinical validation of the ANI monitor under different conditions. Some studies have been performed to evaluate the post-operative pain in awaken patients with the ANI monitor [18–20]. Pain intensity assessed in a 0–10 numerical rating scale by the patients has been compared with the measurement of ANI during the postoperative period. Some controversial have been found when comparing the different studies. Main source of conflicts may come from the presence of arousal and emotions affecting the sympatho-vagal balance in awaken patients, constituting an important source of statistical artifact in the evaluation of pain intensity [21]. Including the subjective post-operative pain evaluation of the patient may constitute, therefore, a source of conflicts in this kind of studies.

To deal with this problem, the ANI monitor has been evaluated throughout the surgery. In Refs. [22,23], the evolution of the ANI as well as heart rate and systolic blood pressure were recorded during the anesthetic process. Patients received tetanic stimulation to study the capability of the different clinical signs to reflect the nociceptive stimuli. The results evidenced opposing conclusions when evaluating the reactivity of ANI for the detection of the stimuli compared with other hemodynamic variables. Some other studies have resulted in inconclusive results when performing ANI-guided analgesia. Although the intraoperative opioid consumption was reduced, no effect was observed in the reduction of opioid-related side-effects [24]. In conclusion, no evidence exists for a clinically relevant benefit of ANI monitoring so far.

In light of the above, this research is not focused on the development of an algorithm for the automation of analgesia, but on a new point of view to overcome main difficulties found in previous research for the evaluation of the ANI monitor. Specifically, the main objective of this study lies in the evaluation of the Analgesia Nociception Index as a feedback variable to replicate the decisions of the clinician when an abnormal level of analgesia is detected. Thus, main novelties of the study are:

i) This proposal will not be based on a clinical validation of the ANI monitor as a tool capable of measuring the analgesic level, but on the analysis as a tool capable of providing valuable information when a change of dose is needed. This new scheme lies in the application of machine learning techniques for the analysis.
ii) The analysis is free from the subjectivity introduced by the post-operative evaluation of the patients. Unlike some of the previous proposals, only data recorded during the surgery involving both clinical variables and actions performed by the clinicians will be studied. As a result, the effects of arousal or emotions are diminished.
iii) The data acquisition process proposed in this methodology is not based on an invasive scheme. Instead of applying painful stimuli for the analysis of ANI as included in previous proposals, this study is fully based on the data obtained during real surgeries. Consequently, potential damage introduced by external painful stimulus is avoided.

2. Methods

2.1. Monitoring nociception: The Analgesia Nociception Index

The Analgesia Nociception Index (ANI) developed by Mdoloris Medical System [25], is based on the analysis of the parasympathetic component of the autonomic nervous system regarding the respiratory sinus arrhythmia. This is a consequence of a diminution of the RR intervals during inspiration. ANI uses specific electrocardiogram (ECG) electrodes placed on the chest or in the back to measure the heart rate variability. The spectral analysis of ECG results in a dimensionless score (0–100) displayed every second. ANI_i appears in yellow on the monitor and it is straightforward influenced by the reactions of the patients to the actions of the surgeon. Moreover, the monitor displays an additional value, ANI_m, which results from a 2 min averaging of ANI_i. ANI_m is supposed to be related to effects of analgesia on patients and, therefore, to be of interest for the titration of opioids. This device can be used with unconscious as well as conscious patients. For unconscious patients under general anesthesia, keeping ANI_m in the 50–70 range is related to an adequate analgesia, avoiding unwanted hemodynamic events. In case that ANI_m decreases below 50, hemodynamic reactivity in the next 10 min has been observed. Finally, ANI_m values over 70 makes it possible to decrease opioids administration without any risk.

2.2. Clinical protocol and data collection

This study was approved by the Ethics Committee for the Clinical Research of the Hospital Universitario de Canarias. Written informed consent was obtained from the patients enrolled in the study. A Total Intravenous Anesthesia (TIVA) with propofol (hypnotic drug) and remifentanil (analgesic drug) was performed for induction and

maintenance of general anesthesia. Two syringe pumps Graseby 3500 were used. Intravenous remifentanil infusion of 0.2 μg/kg/min started 7 min before induction. A propofol intravenous bolus of 1.5 mg/kg at the maximum syringe pump rate (1200 ml/min) was supplied. A Bispectral Index monitor (BIS) was used as guidance variable for propofol titration. The propofol dose was changed manually during the surgery to maintain BIS values between 40 and 60, with a target of 50. Remifentanil dose was adjusted by the clinician as a consequence of the hemodynamic response, defined as the variation of more than 20% of the heart rate (HR) and/or blood pressure (BP) for 5 min. Additionally, the anesthesiologist could change the remifentanil dose to prevent the effects of surgical stimuli during the process. Changes of remifentanil up to 0.05–0.1 μg/kg/min were allowed. Further details for remifentanil drug titration in anesthesia can be found in Ref. [26]. The clinical protocol also included a post-operative evaluation of the patients. This evaluation consisted of detecting postoperative complications related to the opioid administration such as nausea, vomiting, shaking or fatigue, the needs of dosing any other drug to minimize analgesia effect, or the time spent in Post-Anesthesia Care Unit (PACU).

To obtain the dataset for the analysis, information about remifentanil dose, heart rate, blood pressure and ANI were recorded during the surgery. For HR and BP monitoring, BeneView T8 or iPM 12, both developed by Shezhen Mindray Bio-Medical Electronics Co., were used in this study. Non-invasive blood pressure cuffs (NIBP) and ECG electrodes and cables compatible with the monitors were used for BP and HR respectively. The use of one monitor or the other depended on the availability in the operating room scheduled for the surgery. For the Analgesia Nociception Index monitoring, ANI monitor developed by Mdoloris Medical Systems (software version 1.1) was used. ECG for ANI computation was obtained by two electrodes placed on the patient's chest. A PC ran a real time application developed in Matlab for the data acquisition. Information from the ANI monitor (ANI_m and ANI_i) was recorded automatically at a sample time of 5 s in the PC via a USB port. Two anesthesiologists took part in each intervention. Anesthesiologist 1 was in charge of the drug supply task. Anesthesiologist 2 oversaw the acquisition process and registered the variations of HR, BP and remifentanil changes in the Matlab application every 5 min as recommended in Ref. [27]. A detailed scheme of the collecting system developed in the operating theatre is shown in Fig. 1.

The main key to this study lies in the information displayed by the ANI monitor was not available for anesthesiologist 1 to avoid biasing their decisions. Apart from the changes in remifentanil, anesthesiologist 2 also wrote down the occurrence of relevant clinical events such as surgical stimuli or the titration of additional drugs. As clinicians were not familiar with the program, not only a PC, but also a data collection notebook was used for data recording of ANI, HR, BP every 5 min as well as the description of the main clinical events.

2.3. Data preprocessing

Before applying the machine learning algorithms, the dataset recorded was preprocessed according to the following steps:

1. *Identification of the goal information*: This study aimed to analyze those remifentanil changes that were based on the hemodynamic response of patients. As a matter of fact, changes due to the anticipation to surgical stimuli warned by the surgeon should not be taken into account for the analysis. According to the timeline presented in Fig. 2, only changes between T_5 and T_6 were considered.
2. *Validation of hemodynamic data*: To check the hemodynamic data, both anesthesiologists involved in the data collection process inspected the data recorded in the notebook after each surgery. According to the criteria of the experts, if an abnormal record of HR or BP was detected, that part of the intervention would be discarded for the analysis. Then, the information recorded in the computer was compared with the data collection notebook. In case of any divergence, notebook data prevailed against PC.
3. *Validation of ANI data*: The presence of some artifacts during the intervention worsened the quality of the signal. In those cases in which the monitor was not capable of processing the information

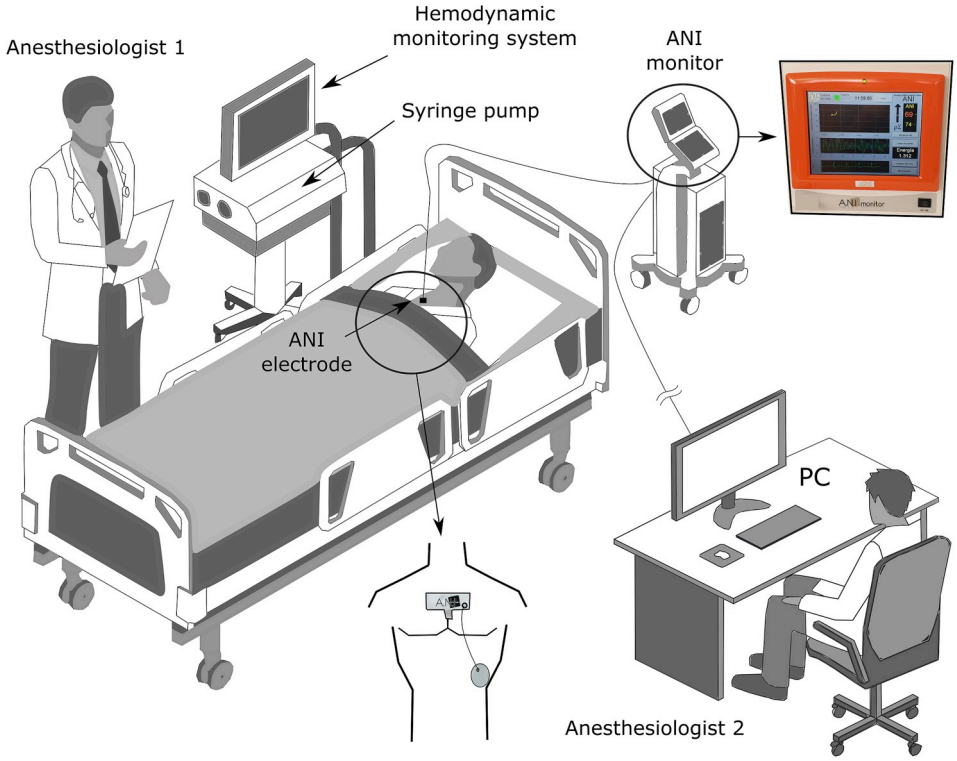

Fig. 1. Scheme of the data acquisition process during the surgeries.

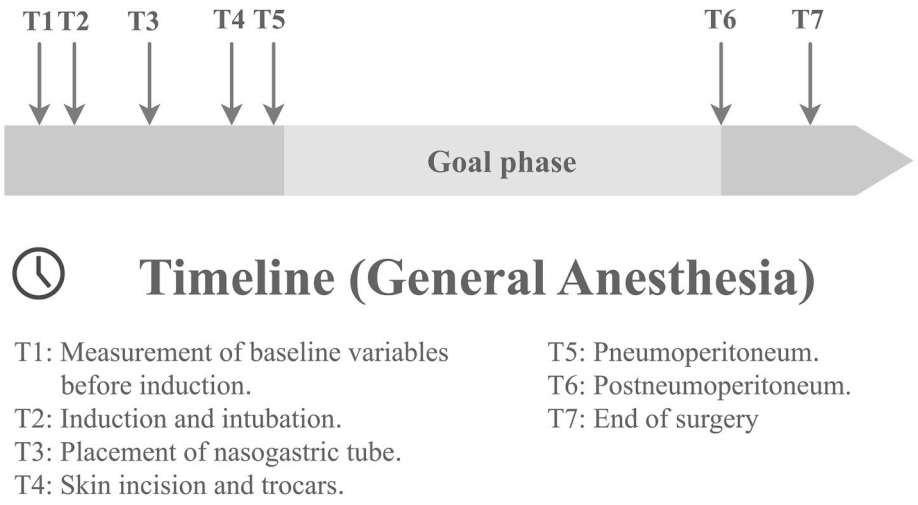

Fig. 2. Timeline of the general anesthetic process in patients enrolled in this study.

from the sensor, it displayed a zero value for both ANI_i and ANI_m. It was observed that these situations remained for less than 2 min during the goal phase. Considering the dynamics of the signal, a linear interpolation was applied to reconstruct the index during this time frame. New values at instant k, $k = \{0,T, 2T, \ldots,T_{error}\}$, were approximated using the values of the line that joined the two points immediately anterior (ANI_0) and posterior (ANI_{end}) to the failure during the time the error persisted (T_{error}) at each sample time $T = 5s$ as:

$$ANI_k = ANI_0 + \frac{ANI_{end} - ANI_0}{T_{error}} \cdot k$$

4. *Matching the information from both sources*: In case of any complication during the surgery, the anesthesiologist 2 helped the anesthesiologist 1 with the anesthetic process except for the analgesic titration. Under this condition, and trying to avoid missing data, it was specified that the hemodynamic data would be only registered in the notebook. As a result, this information had to be included afterwards in the digital record for the analysis. The main problem lied in the divergences on the timestamp. Unlike the timestamp in the program, expressed as the time spent since the beginning of the surgery (expressed in seconds), the anesthesiologist linked each manual record with the time of day (expressed as HH:MM). To deal with this issue, a protocol for merging the information was designed. As the ANI values were automatically captured every 5 s in the PC and, given the fact that the anesthesiologist 2 also wrote down the ANI values every 5 min by hand, it was possible to match this information to synchronize both sources. Note that the same ANI_i and ANI_m tuples could have happened more than once along the surgery. To face this problem, different tentative times in the PC format were first assigned to each event recorded in the notebook. To do that, the timestamp of the "Incision" event was considered as the reference as this event was always recorded in both sources. Finally, to assign each manual record with the appropriate record in the PC, the Euclidean distance including ANI_i, ANI_m and time from both sources was computed. Those manual and PC values that resulted in the minimum Euclidean distance were linked for the analysis.
5. *Labelling the dataset*. After the evaluation of the hemodynamic state of the patient every 5 min, the anesthesiologist must have increased, decreased or kept the dose of remifentanil to ensure an appropriate analgesic state. A variation in the remifentanil infusion rate evidences an inadequate level of analgesia. According to the main objective of this study, only those records associated with changes of the drug were considered for training the classifiers. Therefore, qualitative variables were defined as "Increment of drug" or "Decrement of drug" for these situations.
6. *Creating the csv file*: Each change of dose was kept as a different record. They included information about the evolution of HR, BP, ANI_i, ANI_m and remifentanil dose during the 10 min before a change. All the records were saved in a csv file to extract the information according to the input feature proposal.

2.4. Machine learning approach

According to the main objective of this study, different machine learning algorithms were trained to analyze whether including the Analgesia Nociception Index could outperform the decision making process only based on hemodynamic variables. Initially, an automated training was carried out to search the best classification model type among all the methods included in the Classification Learner toolbox provided by MATLAB2017a. Those methods corresponded to Decision Trees, Discriminant Analysis, Logistic Regression, Support Vector Machines, Nearest Neighbor Classifiers and Ensemble Classifiers. Different parameters were tested for each method as presented in Ref. [28]. After a preliminary analysis of the results (see Appendix), the four methods that performed best were proposed in order to carry out a more exhaustive analysis:

1) K-Nearest Neighbors (KNN) [29]: number of neighbors was set to 1 and the Euclidean distance was used.
2) Decision Tree (DT) [30]: Maximum number of splits was set to 4 and the GDI (Gini Diversity Index) was used as the split criterion.
3) Linear Discriminant Analysis (LDA) [31]: a linear discriminator with a Gamma parameter set to 0 was used.
4) Support-Vector Machine (SVM) [32]: a linear kernel function with a box constraint set to 1 and a kernel scale automatically selected was used.

First, the aim was to study the capability of the different outcoming models to predict the change of remifentanil in this problem. Specifically, this analysis focused on "Increment of drug" and "Decrement of drug" decisions. To deal with overfitting, all these methods were subjected to a cross-validation procedure. In order to determine the suitability of the different methods, the following performance indicators were computed from the cross-validation process [33]: accuracy; specificity; precision; recall; Kappa index [34] and Area Under the Curve (AUC).

As only those changes of dose made in the goal phase of the surgery would be analyzed, a low number of records was expected. To face this

issue, we proposed a 3-fold cross validation repeated 100 times considering all the data obtained from the acquisition process. On the one hand, splitting the total amount of data in a low number of folds shows the prediction capability of each model. Higher performances would imply that a model is capable of learning the general behavior from a wide range of different situations, despite of the limitations in the number of training data. On the other hand, due to the low number of records, and trying to avoid the expected high variability in the results as a consequence of the fold configuration, the cross validation process was repeated 100 times. Each record was randomly assigned to a fold at each iteration. The results were finally averaged to study the variability. In this sense, the more similar the results among iterations, the more robust the model in terms of the generalization capability. Once the best algorithm was identified, the performance focused on the input feature proposals was studied.

2.5. Feature proposal

To determine the impact of including the Analgesia Nociception Index in the decision making process, different feature vectors have been proposed to train the models. These variables include not only information about the evolution of ANI during the last five to 10 min, but also about the traditional parameters considered in the standard clinical practice. First, the effect of including ANI derived information to predict a dose change is aimed. Particularly, it is important to determine if the performance of the machine learning algorithms purely based on hemodynamic information can be outperformed by including information from the ANI monitor. In the absence of any other clinical information, it will evidence the potential of using the ANI in the operating theatre to detect an inadequate level of analgesia compared with the hemodynamic information. In addition, as the ANI monitor has not been widely used in the clinical practice, vague criteria to interpret the information displayed by the monitor has been defined. Considering the variations observed in the ANI signal during a 5-min period and, in order to perform a reliable comparison with the hemodynamic information recorded every 5 min, different features were extracted from the raw ANI_i and ANI_m that summarize the main characteristics during this time period. As a matter of fact, this study proposes four different feature vectors based on information captured from the ANI for training the different machine learning algorithms:

Feature vector proposal 1: Hemodynamic information: This feature proposal aims to represent the standard clinical practice, in which only hemodynamic information is considered. This includes information about the current values of the systolic pressure, diastolic pressure, heart rate and remifentanil infusion rate, as well as this information of the last 5 and 10 min before a change of drug dose.

Feature vector proposal 2: Minimum ANI information: This feature proposal includes not only the hemodynamic variables and remifentanil infusion rate as presented in the proposal 1, but also the information of the ANI_m averaged in the last 5 min as well as minimum ANI_i and ANI_m values registered during the last 5 min and 10 min before a change of remifentanil dose. As ANI_m has been presented as an indicator of the general analgesic state of the patient, this proposal aims to find possible correlations between the average of ANI_m and the change performed by the clinician. Furthermore, including minimum values reached by the ANI would make it easier to identify patterns related to an inadequate analgesia level, together with the evaluation of the monitor capability to predict hemodynamic events as presented in previous works.

Feature vector proposal 3: Maximum ANI information: Unlike feature proposal 2, maximum values of ANI will be evaluated instead. As a result, hemodynamic variables together with remifentanil infusion rate, the average of ANI_m in the last 5 min and maximum ANI_i and ANI_m values in the last 5 and 10 min will be included.

Feature vector proposal 4: ANI information: This input proposal is only based on the remifentanil infusion rate and on the ANI monitor. The main purpose is to evaluate whether only including information about the ANI could outperform the prediction purely based on the hemodynamic information. To this end, maximum, minimum and mean values of ANI_i and ANI_m for different time span has been considered. Additionally, to determine the ANI_m trend during the last 5 min, the slope of the regression line that better fits the ANI_m evolution in the last 5 min has been computed. This feature adds information not only about the current trend, but also could be considered as a tentative prediction of the future evolution of ANI_m. The feature vector proposal together with each feature included is presented in Table 1.

The same clinical dataset captured for the training phase will be used for this analysis. Finally, three different scenarios representing real situations during anesthesia according to the expert's criteria will be proposed for the evaluation of the classifier:

1. *Urgent changes*: Those changes that must be undoubtedly performed according to the clinical signs of the patient. An urgent change should be carried out when an (absolute) variation of the arterial systolic pressure $> 25\%$ occurs in the last 5 min.
2. *Non-urgent changes*: These records represent those changes that are not based on a strong variation in the hemodynamic activity. These changes could have been based on external factors not recorded in the study such as movements, a compensation of a previous change, or even a prediction of a possible hemodynamic event in the next minutes regarding the recent hemodynamic evolution.
3. *Keep the current dose*: These records represent an absence of change motivated by an accurate analgesic state of the patient. During the data collection phase in this study the clinicians did not report any insight to motivate a no-change of dose decision. To deal with this issue, a conservative criterion to include only those situations in which a no-change of drug could be potentially justified by an appropriate analgesic state of the patient is defined:

- Only those decisions that resulted from the update of the hemodynamic information every 5 min should be included.
- A record in which the dose was kept should be included in the analysis only if the decision was made, at least, 10 min after the last change of dose. The main objective is to minimize the impact of those clinical factors depending on pharmacological considerations. If so, it can be assumed that the decision of keeping the dose of drug is mainly due to an appropriate level of analgesia.
- Only those records followed by no changes of dose during the next 10 min are evaluated. Previous research claims that the ANI monitor is capable of anticipating the appearance of hemodynamic events in the next 10 min. If no evidence of nociception is inferred in the next 10 min and, consequently, no change of dose has been made, it can be assumed that the values displayed by the ANI monitor are compatible with an appropriate analgesic level.

Table 1
Description of the feature vector proposal used for training the classifiers. SP_k: Systolic Pressure, DP_k: Diastolic Pressure, HR_k: Heart $Rate_k$, $Remi_k$: Remifentanil infusion rate, $ANI_{i,k}$: Instantaneous value of ANI as recorded from the monitor, $ANI_{m,k}$: Mean value of ANI as recorded from the monitor. Subscript k indicates the time span (in minutes) considered for the variable as described in the text. $\overline{ANI}$, $\underline{ANI}$ and $\widetilde{ANI}$ represents maximum, minimum and mean values respectively of the corresponding variable. ANI_{mtrend} is the trend of the ANI_m.

Feature vector proposal	Features
1:Hemodynamic information	SP, SP_5, SP_{10} DP, DP_5 DP_{10} HR, HR_5, HR_{10} Remi, $Remi_5$, $Remi_{10}$
2: Minimum ANI	SP, SP_5, SP_{10}, DP, DP_5 DP_{10}, HR, HR_5, HR_{10}, Remi, $Remi_5$, $Remi_{10}$, $\widetilde{\underline{ANI}}_{m5}$ $\underline{ANI}_{m5}$, $\underline{ANI}_{m10}$, $\underline{ANI}_{i5}$, $\underline{ANI}_{i10}$
3: Maximum ANI	SP, SP_5, SP_{10}, DP, DP_5 DP_{10}, HR, HR_5, HR_{10}, Remi, $Remi_5$, $Remi_{10}$, $\widetilde{\overline{ANI}}_{m5}$ $\overline{ANI}_{m5}$, $\overline{ANI}_{m10}$, $\overline{ANI}_{i5}$, $\overline{ANI}_{i10}$
4: Only ANI	Remi, $Remi_5$, $Remi_{10}$, $\underline{ANI}_{m5}$, $\underline{ANI}_{m10}$, $\underline{ANI}_{i5}$, $\overline{ANI}_{m5}$, $\overline{ANI}_{m10}$, $\overline{ANI}_{i5}$, $\widetilde{\overline{ANI}}_{m5}$, $\widetilde{\overline{ANI}}_{m10}$, $\widetilde{\overline{ANI}}_{i5}$, ANI_{mtrend}

As the trained classifier will be dichotomous, i.e. only increments or decrements can be predicted, the analysis of the posterior probability associated with each prediction is proposed for the evaluation of the classifier [35].

3. Results

17 subjects (4 males, 13 females, age: 59 ± 11.6 years, weight: 77.47 ± 14.79 kg, height: 164.94 ± 7.39 cm) were enrolled in the study. No abnormal situation was reported during the post-operative evaluation of analgesia in any of the 17 surgeries included in the study. A total of 58 changes of remifentanil infusion rate (34 increments vs. 24 decrements) were performed during the goal phase of the surgery. The mean number of dose changes per patient was 3.4 ± 2.4. An example of the clinical data recorded for this study during the surgeries is shown in Fig. 3.

Table 2 shows the performance of the different classifiers trained in this study after applying a 3-fold cross validation repeated one hundred times for each input proposal as described in section 2.4. Regardless of the input combination used for the training dataset, the best results were reached when training the Support Vector Machines. Specifically, considering the input proposal number 2, Kappa index and Area Under the Curve increased 0.24 and 0.12 respectively when applying SVM compared with the performance reached by LDA. Under the same conditions, a decrement of the Standard Deviation obtained from the one hundred iterations was also observed. In light of the results, we applied this model to perform further analysis of the data.

Fig. 4 shows a graphical comparison of the performance reached by the SVM models depending on the input feature vectors. Given the fact that the predictor 1 represents the standard situation in which only the hemodynamic information has been considered, the performance of the other classifiers is compared to determine the impact of including ANI in the decision making process. A better performance was observed when training the predictor 2. Considering the accuracy as the capability of matching predictions with the real decisions, the model based on both hemodynamic information and minimum ANI index-derived features reached a score of 86.21% (83.62%–87.93%). This result outperformed the accuracy reached by the predictor number 1 of 82% (79.31%–84.48%), based only on hemodynamic information. Including only the

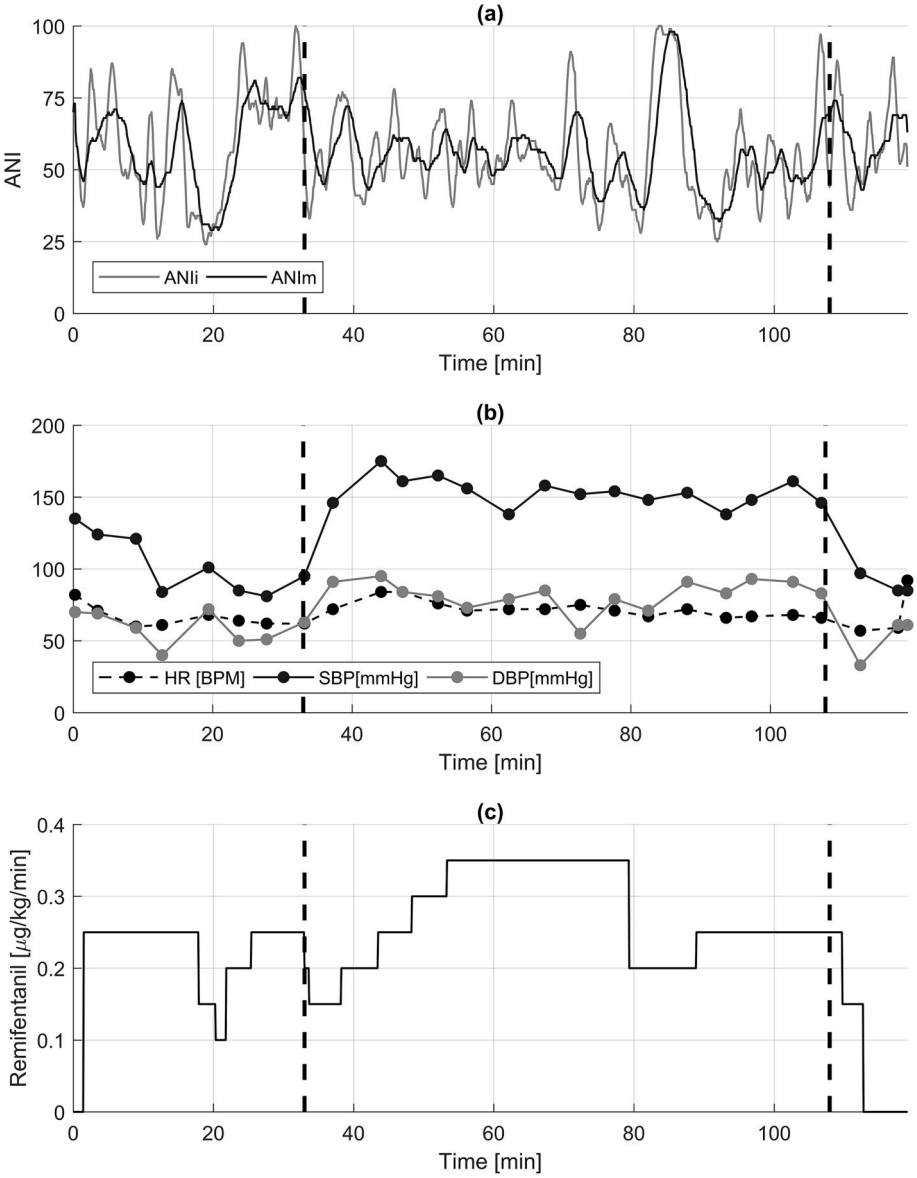

Fig. 3. Evolution of the clinical variables recorded during the study: (a) represents the evolution of instantaneous ANI (ANIi) and mean ANI (ANIm). (b) corresponds to the evolution of the hemodynamic variables measured every 5 min. (c) shows the changes of the remifentanil infusion rate during the surgery. Dashed lines represent the beginning and end, respectively, of the target period considered for this study (T_5 and T_6 respectively).

Table 2
Classifier performance indices expressed as Mean ± SD for different machine learning algorithms and features. KNN: K-Nearest Neighbors, DT: Decision Tree, LDA: Linear Discriminant Analysis, SVM: Support Vector Machine. Feature Vector as enumerated in Table 1.

Method	Feature Vector	Accuracy	Specificity	Precision	Recall	Kappa	AUC
KNN	1	72.84± 4.33	61.71± 6.37	74.96± 3.44	80.71± 5.83	0.43± 0.089	0.71 ± 0.043
	2	64.45± 4.03	58.00± 6.20	70.00± 3.62	69.00± 4.53	0.27± 0.083	0.63± 0.042
	3	68.29± 4.30	61.33± 6.65	72.91± 3.75	73.21± 5.88	0.35± 0.087	0.67± 0.043
	4	55.09± 6.09	56.63± 10.63	64.10± 6.62	54.00± 6.19	0.10± 0.13	0.55± 0.065
DT	1	69.95± 5.78	70.42± 10.23	77.33± 6.46	69.62± 7.72	0.39± 0.12	0.73± 0.065
	2	68.52± 5.58	66.21± 10.97	75.04± 5.84	70.15± 8.78	0.36± 0.11	0.71± 0.07
	3	68.66± 5.84	68.54± 9.97	75.89± 6.14	68.74± 7.66	0.37± 0.12	0.71± 0.069
	4	54.91± 6.2	47.79± 9.43	61.82± 4.97	59.94± 10.63	0.11± 0.078	0.55± 0.069
LDA	1	78.09± 4.09	71.42± 7.11	80.54± 4.12	82.79± 3.97	0.55± 0.086	0.80± 0.038
	2	72.90± 5.29	67.29± 7.64	77± 4.61	76.85± 6.57	0.44± 0.11	0.76± 0.050
	3	72.02± 4.58	67.29± 6.79	76.64± 4.01	75.35± 6.47	0.43± 0.091	0.75± 0.043
	4	52.88± 5.81	42.38± 8.38	59.66± 4.75	60.29± 8.04	0.027± 0.12	0.53± 0.06
SVM	1	81.14± 3.87	73.08± 6.22	82.14± 3.57	86.82± 4.29	0.61± 0.081	0.87± 0.025
	2	84.45± 3.58	79.13± 6.50	85.82± 3.85	88.21± 3.53	0.68± 0.075	0.88± 0.023
	3	80.78± 4.67	72.33± 8.17	81.80± 4.64	86.74± 4.61	0.60± 0.99	0.87± 0.032
	4	59.91± 4.65	39.00± 7.69	63.46± 3.46	74.68± 5.9	0.14± 0.99	0.60± 0.046

ANI index information decreased the accuracy to 60% (56.90%–62.07%).

For the evaluation of the performance depending on each kind of decision, i.e. increments and decrements of remifentanil infusion rate, precision and recall as well as specificity were analyzed. On the one hand, precision and recall indexes can be regarded as measures to quantify the performance of the classifier when predicting increments of the drug. In this sense, classifier 2 was capable of predicting 91.18% (88.24%–91.18%) of the increments of drug correctly, compared to 88.24% (85.29%–91.18%), 85.29% (82.35%–88.24%) and 76.47% (70.59%–79.41%) reached by classifiers 1, 3 and 4 respectively. Moreover, not only a higher median value, but also a narrower interquartile range reached by model 2 evidences the generalization capacity to predict increments of remifentanil regardless of the specific fold considered for the training. In addition, similar conclusions were reached when studying the precision score.

On the other hand, the capability of predicting the decrements of the infusion rate was also studied. Specificity values of 75% (70.83%–79.17%), 79.17% (75%–83.33%), 70.83% (66.67–79.17%) and 37.5% (33.33%–41.67%) were respectively reached by the different predictors. In this case, introducing information about the minimum values of ANI index also allowed improving the decision making process focused on decrements of the drug. Despite of the acceptable results reached by some of the classifiers, it is important to note that lower values are reached when comparing specificity with recall scores. It could be due to the training dataset included more samples involving increments of drug rather than decrements. Notwithstanding the fact, regarding the clinical scenario described in this analysis, to predict the increments of drug accurately is a critical decision in order to avoid chronic pain and short times of recovery after the surgery. Considering the low range of remifentanil infusion rate proposed in the clinical protocol, not decreasing the analgesic dose could rarely provoke a damage for patients. In fact, clinicians tend to overdose analgesic drug to prevent painful situations for patients in the recovery phase.

Furthermore, the general behavior of the classifiers was tested through AUC and Kappa index. Unlike the previous analysis, AUC scores for predictors 1 and 2 slightly differs, 0.87 (0.86–0.89) vs. 0.89 (0.87–0.90) respectively. Conversely, Kappa index firmly showed that better results were reached when considering not only hemodynamic evolution, but also minimum ANI information during the last 10 min. In addition, both AUC and Kappa index evidenced a poor performance when the remifentanil changes only depended on the information derived from the ANI monitor.

In light of the above, combining features involving both hemodynamic evolution and minimum ANI values during the 10 min before a change of dose outperformed those decisions made by the anesthesiologist only based on hemodynamic information. Moreover, introducing information about maximum ANI was not enough for improving the decisions based on traditional clinical criteria. It is important to highlight that using only ANI as a guiding variable for remifentanil dosage during anesthesia worsened the decision-making process.

Finally, the capability of model 2 for predicting each individual change included in the dataset was analyzed. For this purpose, the results of the predictions after the cross-validation process performed in each iteration was computed. Success rate per change considering the 100 iterations is depicted in Fig. 5. It showed that 76% of changes were predicted according to the clinician's criteria achieving a success rate within 90%–100%. Specifically, 55% of the predictions were always right regardless of the iteration. Despite of these promising results, there were at least 7 situations with a success rate under 60%. To deepen in the limitations of this proposal, these situations were analyzed.

On the one hand, changes number 11, 34, 42 and 57 were characterized by a similar trend in terms of hemodynamic variables and ANI evolution that resulted in discrepancies when comparing the predictor

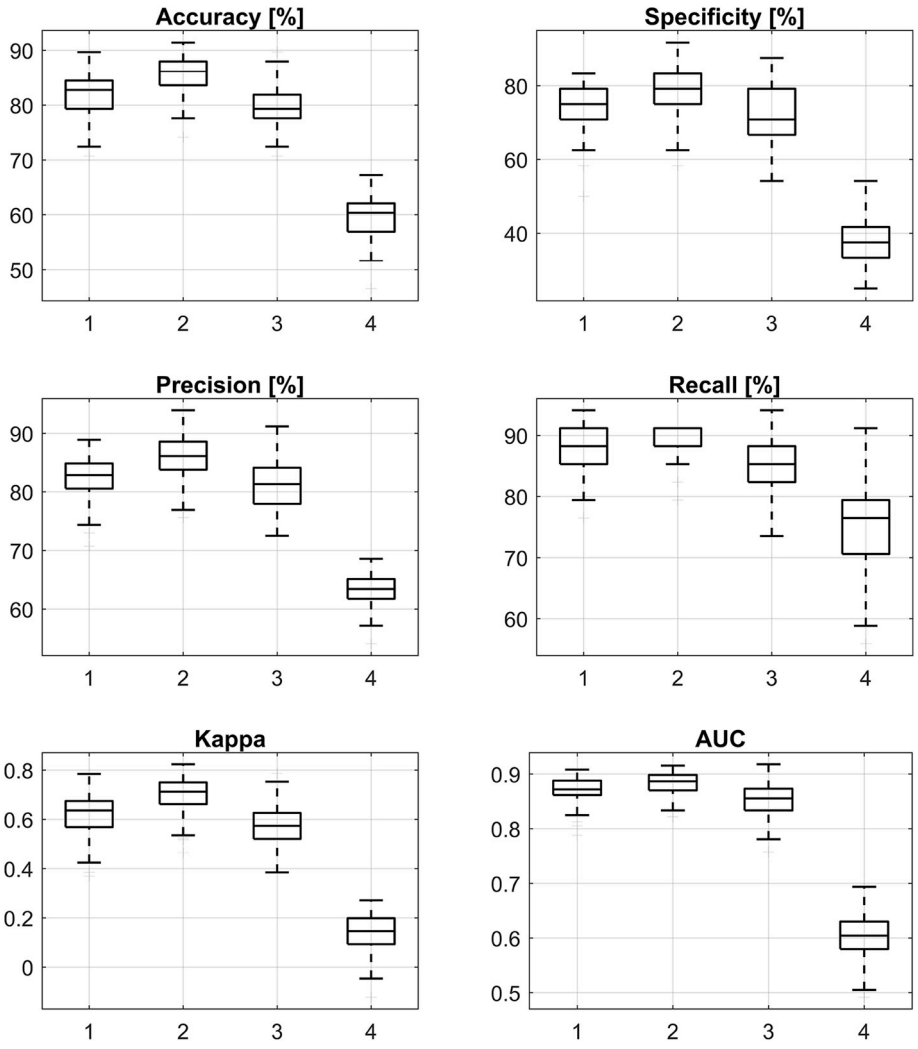

Fig. 4. Boxplots representing the performance of the SVM models for the different input proposals.

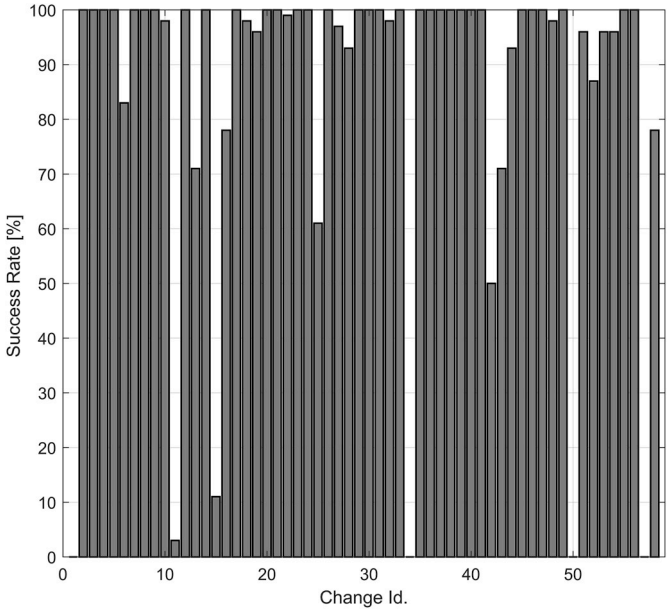

Fig. 5. Success Rate expressed as the % of times that the SVM predictions matched each decision of the clinician for the features proposal number 2.

response with the clinician's decisions. Small changes in heart rate as well as in blood pressure were observed during the previous 10 min. These slight variations in the hemodynamic variables could evidence a questionable decision. In fact, some of these changes might have been based on the previous remifentanil infusion rate as it was set in the limits of drug allowed in the clinical protocol (0.05 μg/kg/min and 0.4 μg/kg/min). As a matter of fact, a boundary value of remifentanil could result in a slanted decision when small hemodynamic variations were observed. In addition, these four cases were related to ANI values that reported an unacceptable level of analgesia according to the instructions of the monitor. In order to check the effectiveness of the decision made by the clinician in these four decisions, changes in the remifentanil infusion rate performed in the next 10 min after the change were also analyzed. It was observed that the clinician had to amend the decision due to the presence of post-risky hemodynamic events registered. In conclusion, despite of existing a divergence with the decision of the clinician, considering both hemodynamic and ANI information could have avoided the hemodynamic reactivity in patients.

On the other hand, change number 50 seems to derive from an error when acquiring data during the intervention. Probably, there was a delay between the real change of dose and the instant in which the second anesthesiologist recorded the change in the computer. As a consequence, the variations in both hemodynamic variables and ANI evolution considered for the training process could have been affected by the pharmacological effect of the real change of dose. It is important

to highlight that it was a punctual situation due to the manual acquisition of remifentanil that did not distort the procedure defined for the data collection. Finally, changes number 1 and 15 did not follow any of the patterns analyzed so far. Both cases consisted of variations in hemodynamic variables directly opposed to the information displayed by the ANI index. In this sense, these changes might have been based on additional information not recorded in the program, such as a possible anticipation of a surgical stimulus warned by the surgeon. Likewise, it is important to point out that heart rate and blood pressure are not exclusive measures of the sympathetic-parasympathetic balance of the nervous system. As a result, other physiological events beyond analgesia could have provoked these hemodynamic variations. Consequently, new studies should be carried out to analyze the possible uncertainties that may affect the process in order to deal with them.

3.1. Analysis of the two-class classifier in different clinical situations

According to the criteria defined in subsection 2.6, 26 urgent cases, 32 non-urgent cases and 16 situations in which the dose was kept were identified throughout the 17 interventions. The performance of the two-class classifier was evaluated in the three proposed scenarios. The posterior probability of the *predicted class* has been computed for the analysis. Consequently, probabilities ranging from 0.5 to 1 were obtained. Fig. 6 summarizes main results of the classification when an urgent or non-urgent change of remifentanil was needed. For the study of those situations where the remifentanil dose was kept and, given the fact that the trained classifier was dichotomous, Fig. 7 merely shows the posterior probability distribution for each prediction.

It was observed that 96% of the urgent cases were accurately predicted by the SVM classifier. In addition, 81% of those cases in which an urgent change was needed were correctly classified with a posterior probability greater than 0.8. The 84% of the non-urgent changes were correctly classified. Particularly, 72% of the non-urgent labelled changes resulted in an accurate classification with a posterior probability greater than 0.65. In those situations in which the dose was not changed, 62% of the records were predicted as increments or decrements of drug with a posterior probability lower than 0.65. Thus, the current classifier was not confident enough to decide whether these cases corresponded to an increment or a decrement of the remifentanil infusion rate. This evidences that these records included information that was not compatible with the patterns learned for the classification in any of the two classes. It was also observed that those predictions that resulted in a posterior

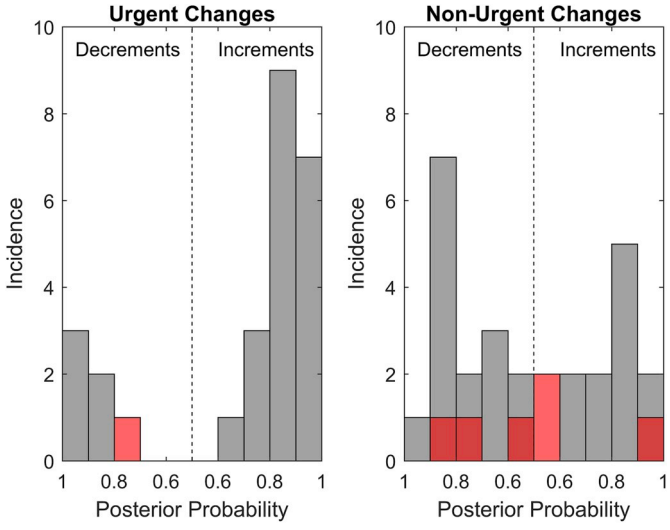

Fig. 6. Posterior probability distribution when predicting urgent/non-urgent increments and decrements of remifentanil. Red bars represent misclassification.

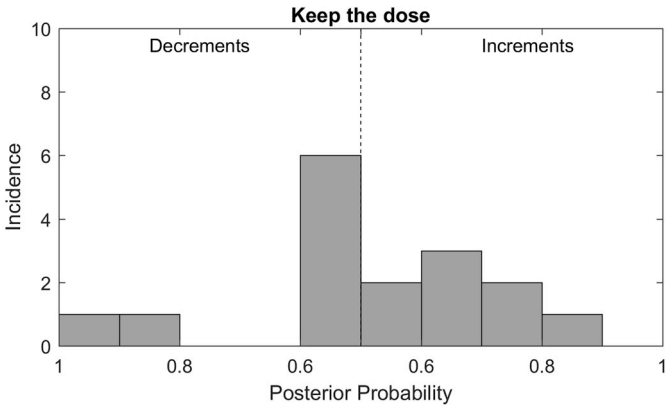

Fig. 7. Posterior probability distribution when classifying those cases in which the dose was kept.

probability greater than 0.65 matches the change performed by the clinician in the next 10–15 min.

Considering the previous results of the classifier in the different scenarios, we have studied the suitability of the current classifier to predict the three possible actions during anesthesia: increment, decrement or no change of dose. The synthesis of the three-class classifier was based on the posterior probability analysis. To this end, the original classifier would predict a "Keep the dose" action if the resulting posterior probability was lower than 0.65. The proposed three-class classifier reached in an accuracy of 77% when predicting the different actions. The confusion matrix that summarizes the general performance is shown in Table 3. These results may evidence that the original classifier is not only capable of distinguishing patterns belonging to critical cases (increments or decrements of drug) but also detects information that does not belong to any of these two categories that could be related to no-change of drug.

4. Discussion

In this paper, we have studied the suitability of the Analgesia Nociception Index to provide valuable information to guide the analgesic drug titration during general anesthesia. Specifically, we have proposed a new scheme based on machine learning classifiers to study the effect of introducing the ANI monitor to replicate the decisions of the experts in critical situations where a modification of dose is required.

This study was based on the clinical data captured from 17 patients undergoing cholecystectomy surgeries. The results of this research evidences that considering the minimum values of ANI during the last 10 min let outperform the decisions made regarding the traditional criteria considering only the hemodynamic information. An accuracy of 86.21% (83.62%–87.93%) was reached when using a SVM. Similar performances have been reached when applying machine learning methods in drug titration. In proton therapy, an accuracy ranging from 0.75 to 0.88 was reached when applying a leave-one-out study to predict the discrepancies between planned and real dose [36]. In cancer drug therapies, the prediction score for optimal drug dose reached a 0.8 of accuracy [37]. In addition, a median kappa index of 0.71 turns our classifier into a good tool for prediction [38]. It was observed that this

Table 3
Confusion matrix of the three-class classifier.

		Real Action		
		Increment	Decrement	Keep
Predicted Action	Increment	29	1	4
	Decrement	3	18	2
	Keep	2	5	10

monitor does not only help in the decision making process, but also is capable of predicting and avoiding errors in the drug titration process.

To complete this study, we have also analyzed the behavior of the classifier for the predictions of different situations observed in the clinical practice. The results pointed out that urgent changes of remifentanil were mostly accurately predicted with a high confidence level. In addition, the two-class classifier presented a low confidence level when evaluating those cases in which the dose of drug was kept. This fact may evidence that the feature proposal based on hemodynamic information together with minimum ANI values provides enough information to distinguish, not only between critical situations where an increment or decrement of drug is needed, but also to detect those situations where an appropriate level of analgesia is observed.

The main difference with respect to previous published works lies in the methodology used. Specifically, the proposed analysis based on machine learning together with a non-invasive scheme for the analysis of the monitor constitutes one of the major novelties of this work. Most of the previous research has been based on the validation of the ANI monitor in the clinical stage. Main trend has been based on establishing a correlation between values of ANI recorded during the surgery and post-operative pain reported by the patients through VAS scores. In Ref. [18], an association between acute postoperative pain and ANI scores was reported with very high negative predicted values of higher ANI scores (>57) for determining acute pain. However, several posterior studies reported no relation between both variables [39,40]. One important weakness of these previous studies lies in the presence of subjectivity when including the patient evaluation of the postoperative pain since the perception of pain differs from persons [41]. For the evaluation of the ANI monitor in this study, subjectivity of patients has not affected the results. Thus, the suitability of the monitor has been evaluated through its capability to report valuable information in order to replicate the actions of the anesthesiologist during the surgery. Furthermore, unlike previous studies, our strategy is not based on an invasive scheme, preventing harmful situations for the patient derived from the application of painful stimuli.

Although this study constitutes a new point of view for the evaluation of ANI, some of the conclusions reached can be compared with previous published works. Some of the results analyzed in our study have evidenced the capability of anticipating hemodynamic events during the surgery. Particularly, it was observed that the evolution of the ANI index may warn the appearance of a hemodynamic event in the next 10–15 min despite of the absence of change in heart rate or blood pressure. In Ref. [42], sensitivity and specificity (88% and 83%) evidenced the predictive capability of ANI to anticipate hemodynamic changes in the next 5 min. Other studies, however, have detected a low probability of ANI to detect hemodynamic reaction [43].

This study presents some limitations that should be taken into account. First, only 58 records involving changes of remifentanil were recorded throughout the 17 surgeries. This may seem a low number of records for the application of machine learning techniques. However, the analysis of the performance presented in Section 3 considering not only mean values, but also the deviations resulting from the 100 iteration of a 3 cross-validation overcomes this difficulty. From the clinical perspective, 58 cases may have not described all the possible situations observed in the clinical practice in which a change of dose is required. To deal with this issue, future research should be conducted to include a higher number of records. Notwithstanding that fact, note that the current SVM classifier was capable of predicting accurately urgent changes of doses with a high confidence interval. Consequently, new urgent cases based on hemodynamic information may be also predicted by the classifier. In addition, the performance of the monitor in other kind of surgeries should be also analyzed. Note that this study has been only focused on those decisions made according to the analgesic state of the patient rather than on other factors depending on the kind of intervention. To this end, a specific phase of the surgery was analyzed. Consequently, low variations of the results are expected when trying the classifier in new kind of surgeries. Further studies should be conducted to this end. Finally, collecting new cases will be also helpful to study possible slight variations introduced in the decisions making process towards a more personalized titration. These variations are widely known as interpatient and intrapatient variability. New studies should include more cases to detect relevant events in the ANI monitor that may be correlated with this factor. Additionally, increasing the number of records will also make possible the proposal of a new analysis based on a time series forecasting problem including information of the ANI for remifentanil prediction.

It is important to highlight that the main objective of this research was not the proposal of an automatic system to supply drug during anesthesia, but a new approach to study the capability of the ANI monitor to replicate decisions of the anesthesiologist when a bad level of analgesia is detected. Thus, this research constitutes a new step towards the development of a closed-loop solution for remifentanil titration. The main aim will be the design of an application capable of guiding the decision making process in the operating room towards an integral automation of the anesthetic process.

5. Conclusion

The main goal of this research was the analysis of the information provided by the Analgesia Nociception Index as a valuable tool to replicate the actions of the anesthesiologist in remifentanil analgesia. A non-invasive clinical scheme together with the use of machine learning algorithms for the analysis are presented. The results evidenced that (i) including data of the minimum ANI values recorded during the last 10 min outperforms those decisions only based on hemodynamic information; (ii) ANI may be capable of anticipating the need of a change of dose before the appearance of a hemodynamic event and (iii) the resulting SVM performs accurate predictions under different situations commonly observed in the clinical practice, particularly when an urgent change of dose must be made. As far as we are concerned, this is the first study in which the actions of the clinicians based on hemodynamic information has been objectively correlated with the information displayed by the ANI monitor during anesthesia. Despite of more research is needed to test the suitability of ANI, including a higher number of patients and types of surgeries, the promising results will motivate the development of an intelligent structure based on the information provided by the ANI monitor for a closed-loop control of remifentanil analgesia.

Declaration of competing interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgements

Jose M. Gonzalez-Cava's research was supported by the Spanish Ministry of Science, Innovation and universities (www.ciencia.gob.es) under the "Formación de Profesorado Universitario" grant FPU15/03347. This research was partially supported through the "Fundación Canaria de Investigación Sanitaria" (FUNCANIS) [ref: PIFUN23/18].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2020.103645.

Appendix

This appendix presents the results of the preliminary study performed to determine the machine learning algorithms and parametrization finally included in this study. To this end, all the methods included in the Classification Learner toolbox provided by MATLAB2017a were tested. Parametrization used for each method was presented in Ref. [28]. For the evaluation, a 3-fold cross-validation was repeated 5 times per each classifier type. The resulting mean accuracy of the 5 tests is presented in Table A1. Finally, the 4 classifiers with the highest performance were proposed for the study. Only one classifier per group was considered for the study.

Table A.1
Accuracy obtained for the preliminary study when performing a 3-fold cross-validation repeated 5 times. Feature vectors as presented in subsection 2.5. Classifiers finally selected for the analysis are highlighted.

Classifier Group	Classifier Type	Feature vectors				Mean
		1	2	3	4	
Decision Trees	Complex Tree	70,0	67,9	69,0	53,1	65,0
	Medium Tree	70,0	67,9	69,0	53,1	65,0
	Simple Tree	70,7	68,6	69,7	53,1	65,5
Discriminant Analysis	Linear Discriminant	77,6	72,1	67,6	52,1	67,3
	Quadratic Discriminant	62,7	0,0	0,0	47,6	27,6
Logistic Regression	Logistic Regression	66,7	64,6	62,1	53,8	61,8
Support Vector Machines	Linear SVM	79,0	83,5	77,6	55,8	74,0
	Quadratic SVM	74,8	75,5	79,3	45,5	68,8
	Cubic SVM	74,8	70,3	75,8	47,3	67,1
	Fine Gaussian SVM	69,3	58,6	58,6	57,2	60,9
	Medium Gaussian SVM	77,9	76,2	76,6	54,5	71,3
	Coarse Gaussian SVM	60,3	59,3	58,6	58,9	59,3
Nearest Neighbor	Fine KNN	73,5	63,8	64,2	49,0	62,6
	Medium KNN	65,4	63,9	65,6	39,5	58,6
	Coarse KNN	51,0	58,6	58,6	51,0	54,8
	Cosine KNN	64,2	65,4	63,3	49,2	60,5
	Cubic KNN	64,8	63,6	66,6	48,0	60,8
	Weighted KNN	64,6	64,2	65,4	45,0	59,8
Ensemble Classifiers	Boosted Trees	58,6	58,6	58,6	58,6	58,6
	Bagged Trees	63,4	63,8	61,4	55,1	60,9
	Subspace Discriminant	67,2	65,8	65,2	47,8	61,5
	Subspace KNN	64,3	61,7	60,5	52,1	59,6
	RUSBoosted Trees	61,2	55,9	58,8	47,2	55,8

References

[1] R. Gorczyca, R. Filip, E. Walczak, Psychological aspects of pain, Ann. Agric. Environ. Med. 20 (1) (2013) 23–27.
[2] B. Guignard, Monitoring analgesia, Best Pract. Res. Clin. Anaesthesiol. 20 (1) (2006) 161–180.
[3] Stockle Pierre-Andre, Intraoperative pain monitoring: update and future directions, Anesthesiol. Reanim. 4 (1) (2018) 204–214.
[4] M. Gruenewald, C. Ilies, Monitoring the nociception-anti-nociception balance, Best Pract. Res. Clin. Anaesthesiol. 27 (2) (2013) 235–247.
[5] L.S. Simon, Relieving pain in America: a blueprint for transforming prevention, Care, education, and research, J. Pain Palliat. Care Pharmacother. 26 (2) (2012) 197–198.
[6] K. Kuusniemi, R. Poyhia, Present-day challenges and future solutions in postoperative pain management: results from PainForum 2014, J. Pain Res. 9 (2016) 25–36.
[7] T.J. Gan, Poorly controlled postoperative pain: prevalence, consequences, and prevention, J. Pain Res. 10 (2017) 2287–2298.
[8] M. Jeanne, M. Delecroix, J. De Jonckheere, A. Keribedj, R. Logier, B.T. Tavernier, Variations of the analgesia nociception index during propofol anesthesia for total knee replacement, Clin. J. Pain 30 (12) (2014) 1084–1088.
[9] E. Boselli, R. Logier, L. Bouvet, B. Allaouchiche, Prediction of hemodynamic reactivity using dynamic variations of Analgesia/Nociception Index (Delta ANI), J. Clin. Monit. Comput. 30 (6) (2016) 977–984.
[10] G. Daccache, M. Jeanne, D. Fletcher, The analgesia nociception index: tailoring opioid administration, Anesth. Analg. 125 (1) (Jul. 2017) 15–17.
[11] J.M. Gonzalez-Cava, et al., A machine learning based system for analgesic drug delivery, in Advances in Intelligent Systems and Computing 649 (2018).
[12] J.-L. Casteleiro-Roca, E. Jove, J.M. Gonzalez-Cava, J.A. Méndez Pérez, J.L. Calvo-Rolle, F. Blanco Alvarez, Hybrid model for the ANI index prediction using Remifentanil drug and EMG signal, Neural Comput. Appl. (2018) 1–10.
[13] F. Julien-Marsollier, et al., Evaluation of the analgesia nociception index for monitoring intraoperative analgesia in children, Br. J. Anaesth. 121 (2) (2018) 462–468.
[14] N. Dundar, A. Kus, Y. Gurkan, K. Toker, M. Solak, Analgesia nociception index (ani) monitoring in patients with thoracic paravertebral block: a randomized controlled study, J. Clin. Monit. Comput. 32 (3) (2018) 481–486.
[15] L.K. Kommula, S. Bansal, G.S. Umamaheswara Rao, Analgesia nociception index monitoring during supratentorial craniotomy, J. Neurosurg. Anesthesiol. 31 (1) (2019) 57–61.
[16] T. Ledowski, Objective monitoring of nociception: a review of current commercial solutions, Br. J. Anaesth. 132 (2) (2019) 312–321.
[17] Y. Jiao, B. He, X. Tong, R. Xia, C. Zhang, X. Shi, Intraoperative Monitoring of Nociception for Opioid Administration: a Meta-Analysis of Randomized Controlled Trials, *Minerva Anestesiol.*, 2019.
[18] E. Boselli, et al., Prospective observational study of the non-invasive assessment of immediate postoperative pain using the analgesia/nociception index (ANI), Br. J. Anaesth. 111 (3) (Sep. 2013) 453–459.
[19] Q. Yan, H.Y. An, Y. Feng, Pain assessment in conscious healthy volunteers: a crossover study evaluating the analgesia/nociception index, Br. J. Anaesth. 118 (4) (2017) 635–636.
[20] G. Chanques, et al., Analgesia nociception index for the assessment of pain in critically ill patients: a diagnostic accuracy study, Br. J. Anaesth. 119 (4) (2017) 812–820.
[21] T. Ledowski, J. Bromilow, M.J. Paech, H. Storm, R. Hacking, S.A. Schug, Skin conductance monitoring compared with Bispectral Inde ® to assess emergence from total i.v. anaesthesia using propofol and remifentanil, Br. J. Anaesth. 97 (6) (2006) 817–821.
[22] M. Jeanne, C. Clément, J. De Jonckheere, R. Logier, B. Tavernier, Variations of the analgesia nociception index during general anaesthesia for laparoscopic abdominal surgery, J. Clin. Monit. Comput. 26 (4) (2012) 289–294.
[23] M. Gruenewald, J. Herz, T. Schoenherr, C. Thee, M. Steinfath, B. Bein, Measurement of the nociceptive balance by analgesia nociception index and surgical pleth index during sevoflurane-remifentanil anesthesia, Minerva Anestesiol. 81 (5) (2015) 480–489.
[24] L. Le Gall, et al., Benefits of intraoperative analgesia guided by the Analgesia Nociception Index (ANI) in bariatric surgery: an unmatched case-control study, Anaesth. Crit. Care Pain Med. 38 (1) (2019) 35–39.

[25] Mdoloris Medical Systems, Ani - analgesia nociception index [Online]. Available, https://www.mdoloris.com/de/technologies/ani-analgesia-nociception-index/, 2016. Accessed: 23-Dec-2019.
[26] C. Stroumpos, M. Manolaraki, G.A. Paspatis, Remifentanil, a different opioid: potential clinical applications and safety aspects, Expet Opin. Drug Saf. 9 (2) (2010) 355–364.
[27] J.Y. Min, H.I. Kim, S.J. Park, H. Lim, J.H. Song, H.J. Byon, Adequate interval for the monitoring of vital signs during endotracheal intubation, BMC Anesthesiol. 17 (1) (2017) 110.
[28] MathWorks, Classification learner app [Online]. Available, https://es.mathworks.com/help/stats/choose-a-classifier.html?lang=en, 2019. Accessed: 10-Jan-2020.
[29] Z. Zhang, Introduction to machine learning: k-nearest neighbors, Ann. Transl. Med. 4 (11) (2016).
[30] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning Data Mining, Inference, and Prediction, second ed., Springer Series in Statistics, 2009.
[31] J.H. Friedman, Regularized discriminant analysis, J. Am. Stat. Assoc. 84 (405) (1989) 165–175.
[32] M.A. Hearst, Support vector machines, IEEE Intell. Syst. Their Appl. 13 (4) (1998) 18–21.
[33] A. Tharwat, Classification assessment methods, in: Applied Computing and Informatics, 2018.
[34] M.L. McHugh, Interrater reliability: the kappa statistic, Biochem. Med. 22 (3) (2013) 276–282.
[35] J.C. Platt, Probabilistic outputs for Support vector machines and comparisons to regularized likelihood methods, in: Advances In Large Margin Classifiers, 2000.
[36] P. Gueth, et al., Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy, Phys. Med. Biol. 58 (13) (2013) 4563–4577.
[37] C. Huang, R. Mezencev, J.F. McDonald, F. Vannberg, Open source machine-learning algorithms for the prediction of optimal cancer drug therapies, PloS One 12 (10) (2017).
[38] D.G. Altman, Practical Statistics for Medical Research, Chapman and Hall, Stat. Med, 1991.
[39] R. Issa, et al., Evaluation of the analgesia nociception index (ANI) in healthy awake volunteers, Can. J. Anesth. 64 (8) (2017) 828–835.
[40] N. Parker, T. Ledowski, W.S. Tiong, C. Lee, B. Wong, T. Fiori, Analgesia nociception index: evaluation as a new parameter for acute postoperative pain, Br. J. Anaesth. 111 (4) (2013) 627–629.
[41] H.M. McCormack, D.J. Horne, S. Sheather, Clinical applications of visual analogue scales: a critical review, Psychol. Med. 18 (4) (1988) 1007–1019.
[42] E. Boselli, L. Bouvet, G. Bégou, S. Torkmani, B. Allaouchiche, Prediction of hemodynamic reactivity during total intravenous anesthesia for suspension laryngoscopy using Analgesia/Nociception Index (ANI): a prospective observational study, Minerva Anestesiol. 81 (3) (2015) 288–297.
[43] T. Ledowski, L. Averhoff, W.S. Tiong, C. Lee, Analgesia Nociception Index (ANI) to predict intraoperative haemodynamic changes: results of a pilot investigation, Acta Anaesthesiol. Scand. 58 (1) (2014) 74–79.