Heart Rate Variability-Derived Features Based on Deep Neural Network for Distinguishing Different Anaesthesia States

Jian Zhan
Army Medical University

Zuo-xi Wu
Army Medical University

Zhen-xin Duan
Army Medical University

Gui-ying Yang
Army Medical University

Zhi-yong Du
Army Medical University

Xiao-hang Bao
Army Medical University

Hong Li (✉ lh78553@163.com)
Second Affiliated Hospital of Army Medical University  https://orcid.org/0000-0002-1338-0829

Research article

Keywords: depth of anaesthesia, heart rate variability, deep neural network, discrete wavelet transform

DOI: https://doi.org/10.21203/rs.3.rs-41792/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Estimating the depth of anaesthesia (DoA) is critical in modern anaesthetic practice. Multiple DoA monitors based on electroencephalograms (EEGs) have been widely used for DoA monitoring; however, these monitors may be inaccurate under certain conditions. In this work, the hypothesis that heart rate variability (HRV)-derived features based on a deep neural network can distinguish different anaesthesia states was investigated.

Methods: A novel method of distinguishing different anaesthesia states was developed based on four HRV-derived time and frequency domain features combined with a deep neural network. Four features were extracted from an electrocardiogram, including the HRV high-frequency power, low-frequency power, high-to-low-frequency power ratio, and sample entropy. Next, these features were used as inputs for the deep neural network, which used the expert assessment of consciousness level as the reference output. Finally, the deep neural network was compared with the logistic regression, support vector machine, and decision tree models. The datasets of 23 anaesthesia patients were used to assess the proposed method.

Results: The accuracies of the four models, in distinguishing the anaesthesia states, were 86.2% (logistic regression), 87.5% (support vector machine), 87.2% (decision tree), and 90.1% (deep neural network). The accuracy of deep neural network was higher than those of the logistic regression \((p < 0.05)\), support vector machine \((p < 0.05)\), and decision tree \((p < 0.05)\) approaches. Our method outperformed the logistic regression, support vector machine, and decision tree methods.

Conclusions: The incorporation of four HRV-derived time and frequency domain features and a deep neural network could accurately distinguish between different anaesthesia states; however, this study is a pilot of a feasibility study, providing a method to supplement DoA monitoring based on EEG features to improve the accuracy of DoA estimation.

Background

Both the central nervous and autonomic systems are related to the depth of anaesthesia (DoA) [1]. A DoA that is too shallow increases the risk of intraoperative awareness [2], and a DoA that is too deep can cause delayed recovery [3], cognitive dysfunction, and may increase the risk of death [4]. Therefore, accurate DoA monitoring is crucial to reduce the complications associated with overdose or insufficiency of anaesthetics and guarantee the safety and quality of anaesthesia.

However, the mechanisms of action of general anaesthetics are still not completely understood [5, 6], and there is currently no ‘gold standard’ for evaluating DoA [7]. DoA monitors based on electroencephalograms (EEG) signals, such as bispectral index (BIS), Narcotrend, and entropy have been widely used during surgery [8-10]. However, EEG signals only show the functions of the central nervous system and the indices based on these signals are not sufficiently accurate to assess DoA under certain conditions [11-15]. Therefore, it is essential to seek new methods of DoA monitoring to overcome the drawbacks of mainstream methods based on EEG signals [16] and improve DoA monitoring accuracy.
Electrocardiograms (ECGs) are internationally used in standard monitoring during general anaesthesia [17]. In addition, the heart rate variability (HRV) derived from an ECG is regulated by the central nervous and autonomic systems, and closely related to the DoA during surgery [18-20]. Therefore, HRV may be used as an important supplementary method of EEG monitoring in terms of DoA evaluation [21, 22].

Owing to the strong nonlinear characteristics of the EEG and ECG, nonlinear analysis methods may be used in studies of anaesthesia [23, 24]. Sample entropy (SampEn) is a typical nonlinear analysis method that was developed to study time domain features of HRV [25, 26] and provide an improved assessment of DoA during surgery [27, 28]. In addition, three frequency domain features of HRV, including the high-frequency power (HF), low-frequency power (LF), and ratio of high-to-low-frequency power (HF/LF), are related to the autonomic nervous system and have been implemented in anaesthesia research [29, 30].

Recently, several machine learning algorithms, including logistic regression [31], support vector machine [32], decision tree [33], artificial neural network [34], and deep neural network [35], have been utilized to assess DoA based on different time and frequency domain features of EEG signal. These results indicate that it is necessary to combine multiple time and frequency domain features to improve DoA assessment methods. Moreover, to our knowledge, there are currently few studies based on HRV-derived features combined with machine learning algorithms to identify different anaesthesia states. Thus, we propose the hypothesis that multiple time and frequency features of HRV based on a deep neural network could be used to distinguish different anaesthesia states and provide a key supplementary method for EEG monitoring in the assessment of DoA.

**Methods**

This study protocol was approved by the Institutional Ethics Committee of the Second Affiliated Hospital of the Army Medical University on March 25, 2020 (Chongqing, China, approval number: 2020-078-01). Patients were recruited from March 27, 2020 to April 29, 2020. Written informed consent was obtained from each patient. Twenty-three American Society of Anaesthesiology (ASA) physical status I or II adult patients, aged from 20 to 70 years old, scheduled to undergo elective laparoscopic abdominal surgery were recruited. Exclusion criteria included patients with neurological and cardiovascular diseases or a known allergy history of anaesthetics.

All patients underwent preoperative fasting for at least 8 h. The placement of the chest electrodes was the same for all participants. The five-leads were located at five different positions on the chest. The upper left position was at the junction of the midclavicular line on the left edge of the sternum and the first intercostal space. The lower left position was at the junction of the left midline of the clavicle and the level of the xiphoid process. The upper right position was at the junction between the midclavicular line on the right edge of the sternum and the first intercostal space. The lower right position was at the horizontal junction of the right clavicle midline and the xiphoid process, and the middle position was at the fourth intercostal space on the left edge of the sternum. After the electrodes were placed on the patient chest wall, anaesthesia was usually induced with intravenous midazolam, propofol, sufentanil,
and cisatracurium. Sevoflurane together with propofol and remifentanil were used to maintain anaesthesia. Table 1 summarises this information in detail. Physiological signals (such as ECG, BP, HR, and SpO2) were measured to guarantee the safety of the patients under different anaesthesia states. The attending anaesthetist adjusted the DoA accordingly, using these observed signals and their own experience. From the various monitoring feedback information observed, attending anaesthetists need to analyze, synthesize, and judge the vital function indicators of patients according to their own experience and to make timely adjustments and interventions as needed to keep the vital signs as normal or close to the normal physiological state as possible, to adjust the DoA and maintain it at an appropriate level.

In this study, ECG signals were recorded from twenty-three adult patients under general anaesthesia. The signals were recorded using a Philips MP60 monitor (Intellivue; Philips, Foster City, CA, USA). The operation time was 1—3 h. Raw ECG data were sampled at a 500-Hz sampling frequency.

**Expert assessment of consciousness level**

The expert assessment of consciousness level (EACL) is the average value of the DoA assessment score determined by five experienced anaesthesiologists (i.e., attending physicians) based on clinical recordings and their own experience [27]. An experienced anaesthesiologist trained for many years with rich clinical experience can be familiar with health risks evaluation and accurately assess the DoA through clinical signs, surgical stimulations, the dose of the anaesthetic agent, etc. combined with his or her own clinical experience. Thus, such an expert can perform anaesthesia-related operations proficiently and correctly handle various problems in anaesthesia even if he or she is not in the operating room during surgery. The states of general anaesthesia are classified as: anaesthesia induction, anaesthesia maintenance, and anaesthesia recovery. Anaesthesia induction means that the anaesthesia depth gradually increases, anaesthesia maintenance means that the anaesthesia depth is relatively stable, and anaesthesia recovery means that the anaesthesia depth gradually decreases. These three states represent the different states of anaesthesia depth. The obtained EACL value is a single number from 0–100, similar to the BIS (with 100 denoting ‘fully awake’ and 0 denoting ‘isoelectricity’). During surgery, the clinical information recorded included: (1) vital signs (e.g., HR, BP, SpO2), (2) anaesthetic events, including induction, intubation and extubation of anaesthesia, addition of muscle relaxant drugs, and airway management, (3) surgical events, including the start and end of the surgical procedure and the occurrence of noxious stimulus, (4) other clinical signs, including unusual responses, movement, and arousability under induction and recovery, and (5) any other related events, such as lacrimation, sweating, and patient demography.

**ECG preproccessing**

Body movements and medical device frequency noise are the main artifacts in ECG recordings. These artifacts seriously affect the analysis results of the ECG signals. Therefore, data preproccessing is essential for distinguishing different anaesthesia states, and can normalize and facilitate subsequent analysis. The specific process is detailed in additional file 1(1).
**Frequency domain algorithm**

Wavelet transform is a typical nonlinear analysis technique and one of the most useful methods for biological signal analysis, especially in cases of continuous signals with various frequency features [35, 36]. Therefore, in this study, discrete wavelet transform was used for the frequency domain analysis of the HRV power. The calculation formula for the HRV power is detailed in additional file 1(2). Entropy, as a nonlinear dynamic parameter that can measure the incidence of new information in a time series, can be described as a regularity or degree of randomness indicator. SampEn is an improved algorithm from entropy. The calculation formula for the SampEn is detailed in additional file 1(3).

**Machine learning algorithms**

Logistic regression is a classification algorithm used to predict the probability of classifying dependent variables. A support vector machine is a supervised learning algorithm that can be applied to classification problems. The calculation formula for the support vector machine approach is detailed in additional file 1(4). A decision tree is a multi-classification supervised learning algorithm. The calculation formula for the decision tree method is detailed in additional file 1(5). An artificial neural network is a nonparametric parallel computing model, which is similar to the neural structure of the human brain [37]. It usually consists of an input layer, a hidden layer, an output layer, and numerous interconnected nodes in multiple layers. The deep neural network developed from the artificial neural network was used in this study. The flowchart of the deep neural network construction is shown in Fig. 1. The deep neural network is detailed in additional file 1(6).

**Performance analysis**

The performance of four models was quantified based on the results of cross-validation using the precision, recall, and classification accuracy. Precision is defined as the ratio of the number of correct classifications of an anaesthesia state to the total number of classifications of the same type of anaesthesia state. Recall is defined as the ratio of the number of correct classifications of an anaesthesia state to the number of actual occurrences of this anaesthesia state. Classification accuracy is defined as the ratio of the total number of correctly identified anaesthesia states to the sum of all anaesthesia states. The calculation formulas for the precision, recall, and classification accuracy are detailed in additional file 1(7).

**Statistical analysis**

There are no standardized methods for sample size calculation based on machine learning algorithms. Thus, the sample size calculations in this pilot feasibility study were based on previous reports [32, 34]. The sample size is 23 cases in this study, including 46000 datasets. The average datasets per patient is 2000 datasets. Statistical analyses were performed using SPSS 22.0 (SPSS Inc., Chicago, IL) and Python (version 3.6.5) software. Data were expressed as mean (SD) or percentage, where appropriate. Ternary classification outcome parameters were expressed as events (percentages). Data are presented as tables,
box-and-whisker diagrams, and correlation graphs. In addition, we calculated the distribution of the four features in the three anaesthesia states. The Pearson's correlation coefficient between the EACL and the four features of the deep neural network model was also calculated to estimate the efficacy of the proposed method. The performances of four classification methods were compared: the logistic regression, support vector machine, decision tree, and proposed deep neural network methods. Owing to the small sample size in this study, the sample does not satisfy a normal distribution. Therefore, the four classification methods were compared using the Chi-square test. \( p < 0.05 \) was considered statistically significant.

**Results**

**Primary outcome**

Twenty-three adult patients were analysed in this study. The details of the selection procedure are shown in Fig. 2. Patient demographics and clinical characteristics are shown in Table 1. The deep neural network structure used in this study consisted of four layers: an input layer with four nodes, a hidden layer with ten nodes, a second hidden layer with seventeen nodes, and an output layer with one node.

The precision and recall values of 23 datasets of the anaesthesia induction, maintenance, and recovery states are listed in Table 2. In addition, the classification accuracies of the three different anaesthesia states were obtained through the calculation of the recall and precision. The deep neural network model yielded a classification accuracy of 90.1%, whereas the logistic regression, support vector machine, and decision tree approaches yielded classification accuracies of 86.2%, 87.5%, and 87.2%, respectively. The accuracy of the deep neural network was higher than those of the logistic regression (\( p < 0.05 \)), support vector machine (\( p < 0.05 \)), and decision tree (\( p < 0.05 \)) approaches. A comparison of the logistic regression, support vector machine, decision tree, and deep neural network methods is presented in Table 2. In addition, the precision and recall of the four models during the anaesthesia induction and recovery states were lower than those during the maintenance state.

**Secondary outcomes**

In this study, four features of the HRV were selected as the input of the deep neural network model. Specifically, these were the HF, LF, HF/LF ratio, and SampEn of the RR interval. The EACL was used as the reference output. Fig. 3 shows a clear correlation between the HF, LF, HF/LF, RR interval SampEn, and EACL. There are positive correlations between the HF (\( r = 0.221, p < 0.05 \)), LF (\( r = 0.238, p < 0.05 \)), and HF/LF (\( r = 0.106, p < 0.05 \)) and the EACL. There is a negative correlation between the RR interval SampEn and the EACL (\( r = -0.053, p < 0.05 \)). Therefore, these features can be used for the construction of the deep neural network model. Interestingly, the four features are mainly distributed in the EACL value range of 40-80. In addition, Fig. 4 shows the original ECG signal, filtered ECG signal, filtered RR interval, HF, LF, HF/LF ratio, and EACL in the same time period. The voltage of the filtered ECG signal was mainly between 0 and 2.5 mV. During the sampling period, the voltage of the ECG was relatively stable with no
significant changes. The filtered RR interval, HF, LF, and HF/LF ratio were significantly reduced before reaching a relatively stable level. The trend of change in the three frequency features was similar to that of the EACL.

**Exploratory outcomes**

Fig. 5 depicts the distribution characteristics of the four features under three different anaesthesia states. The HF during the anaesthesia induction state was significantly higher than that of the anaesthesia maintenance state ($p < 0.001$). The HF during the recovery state was significantly higher than that of the anaesthesia maintenance state ($p < 0.001$) and anaesthesia induction state ($p < 0.001$). Moreover, the LF gradually decreased during the three anaesthesia states. The HF/LF ratio during the anaesthesia recovery state was significantly higher than those of the anaesthesia induction and maintenance states ($p < 0.001$). Finally, the SampEn of the RR interval gradually increased under the three anaesthesia states.

**Discussion**

This study proposed a novel method for distinguishing different anaesthesia states based on four HRV-derived time and frequency domain features, combined with a deep neural network. In addition, this study compared the proposed deep neural network model with logistic regression, support vector machine, and decision tree in terms of classifying three anaesthesia states. The datasets of 23 anaesthesia patients under general anaesthesia were used for assessing the proposed method. We used the precision, recall, and accuracy for model performance assessment. Each of the four models provided high accuracy in classifying three anaesthesia states. However, the accuracy of the proposed method outperformed the three conventional methods. This suggests that, by testing the datasets from multiple HRV-derived features, it is possible to obtain a reliable anaesthesia states prediction based on machine learning algorithms.

Most of the researches have assessed the DoA based on EEG features and machine learning algorithms; however, few studies have distinguished different anaesthesia states using HRV-derived features based on machine learning algorithms. Several studies have been developed to predict the DoA using combinations of multiple EEG features and logistic regression [31], support vector machine [32], decision tree [33], and artificial neural network [34] respectively. We took a multidimensional approach using logistic regression, support vector machine, decision tree, and deep neural network methods and four HRV-derived features to distinguish different anaesthesia states. One of the major findings in this study is that, like EEG features, HRV-derived features based on machine learning algorithms can also distinguish different anaesthesia states. Moreover, Liu et al. used only the similarity and distribution index of HRV based on an artificial neural network to assess the DoA [21]. The similarity index of HRV could distinguish between the waking and isoflurane anaesthesia states [38]. Our findings are consistent with these results in that HRV-derived features selected can also be used to distinguish different anaesthesia states. However, we used multiple HRV-derived features and machine learning algorithms, which is different from the previous studies.
To assess the accuracy of these machine learning algorithms, we chose the EACL as the evaluation criterion for distinguishing different anaesthesia states. The EACL adopted in this study is a method of clinical evaluation by five experienced anaesthesiologists for evaluating the DoA. As the current DoA monitors such as the BIS are based on probabilistic approaches, the golden standard remains clinical assessment of the level of consciousness [39]. In addition, the current DoA monitors based on EEG features have accuracy limitations [11, 13–15]. To improve the accuracy of DoA estimation, previous studies used the EACL as the evaluation criterion for DoA assessment. Liu et al. used the EACL as the reference standard for output of an artificial neural network to assess the DoA [21]. Meanwhile, Jiang et al. used SampEn analysis of EEG signals based on an artificial neural network and EACL to model patient consciousness levels [27]. However, our study shows that HRV-derived features based on a deep neural network and EACL can be used to distinguish different anaesthesia states. Thus, in addition to the current DoA monitors, the EACL is also a reliable method of identifying anaesthesia states.

Our findings showed a clear correlation between the four HRV-derived features, and EACL. These HRV-derived features are also closely related to different anaesthesia states. As the HRV is controlled by the central nervous system, the DoA should be considered to assess the effects of anaesthetics on HRV [20]. To date, it is considered that the LF reflects the parasympathetic and sympathetic systems, whereas the HF and entropy are mediated primarily by the parasympathetic system [40, 41]. In addition, some previous studies have shown that HRV-derived features, including the entropy, HF, LF, and HF/LF, could reflect changes in the DoA. Propofol decreases entropy and HF in a BIS-dependent manner [20] and that propofol is related to the relative decrease in HF, increase in LF, and significant decrease in HF/LF during the anaesthesia induction state [42]. However, abrupt increases in LF and HF are related to the point at which patients became responsive to verbal commands during the anaesthesia recovery state [43], whereas our study showed that the HF increased and LF decreased. In addition, the results in this study indicated that the change of four HRV-derived features could reflect the change of anaesthesia states. Therefore, these HRV-derived features are reliable features of distinguishing anaesthesia states. However, the correlation between a single feature and the EACL was not strong, and the synergy between the four features can be improved to classify the different anaesthesia states. Thus, to implement the proposed method in clinical settings, different features need to be selected for subsequent research and improve the accuracy of prediction method.

The ideal DoA prediction method should have high accuracy and should not be influenced by interference from irrelevant signals. Our findings show that, with the help of multiple HRV-derived features and machine learning algorithms, it is feasible for distinguishing different anaesthesia states. In addition, the proposed method has several advantages. On the one hand, ECG signals are more stable and less susceptible to noise than EEG signals. On the other hand, in comparison with the EEG, the electrode sensors used for ECG signal acquisition are cheaper, rendering ECG a more cost-effective method. More importantly, our method may be a useful adjunct in monitoring DoA based on EEG features and is expected to assist anaesthesiologists in the accurate evaluation of the DoA.
Although promising, there are several limitations and a need for further improvement. First, we did not distinguish nociceptive effects and other physiological parameters such as hemodynamic and respiratory variables on HRV. However, our findings provide important references to guide future investigations. Second, we only explored four HRV-derived features as the inputs of the deep neural network in this feasibility study. We limited these features since they contain both time-frequency domain characteristics of HRV. In addition, cross-validation was used to train and test the model to avoid over-fitting, ensure model generalization, and improve the performance of the deep neural network. Additionally, we considered the impact of inter-clinician variability on the performance of the deep neural network model. To minimize personal error, the mean values of the DoA assessment score determined by five experienced anaesthesiologists were used as the reference standard for output of the deep neural network. Third, the DoA in this study was classified into the three anaesthesia states in the deep neural network model. It is necessary to explore new methods of DoA evaluation with higher precision, better performance, and more classifications (e.g., four or more states) in subsequent work. Fourth, the number of patients used in this study was limited. Increasing the number of patients could improve the performance of our proposed method. Besides, due to the emergence of agitation during the recovery period, the electrodes on the chest walls of eight patients fell off and the ECG data were interrupted, causing technical failure.

Conclusions

In conclusion, this study combined multiple HRV-derived features, including three frequency-domain features and one time-domain feature, with four machine learning algorithms to identify the three anaesthesia states. The proposed method could accurately distinguish between different anaesthesia states and outperformed better than that of the three traditional machine learning algorithms. Our method provides a useful reference for supplement of DoA assessment based on EEG features and is expected to assist anaesthesiologists in the accurate evaluation of the DoA. Other physiological signals, such as EEG, could be incorporated into the proposed method to improve the accuracy of DoA estimation further.

Abbreviations

DoA: depth of anaesthesia; EEG: electroencephalogram; HRV: heart rate variability; DWT: discrete wavelet transform; DNN: deep neural network; HF: high-frequency power; LF: low-frequency power; HF/LF: high-to-low-frequency power ratio; RR interval: the interval between R peaks in two adjacent heartbeats of the ECG; EACL: expert assessment of consciousness level; BIS: Bispectral index; ECG: electrocardiogram; HR: heart rate; BP: blood pressure; SpO\textsubscript{2}: peripheral oxygen saturation; SampEn: Sample entropy; ASA: American Society of Anaesthesiology; Hz: hertz; SD: Standard Deviation; BMI: body mass index

Declarations

Ethics approval and consent to participate

Ethical approval for Institutional Ethics Committee of the Second Affiliated Hospital of Army Medical
University prior to patient enrolment. Written informed consent was obtained from the patients.

**Consent for publication**
Not applicable.

**Availability of data and materials**
The datasets are not publicly available, but available from the corresponding author on reasonable request.

**Competing interests**
The authors declare no conflict of interest.

**Funding**
This study was supported by National Key Research and Development Project (2018YFC0117200) and Clinical Research Project of Army Medical University (No.CX2019LC114 and 2018JSLC0015).

**Authors’ contributions**
ZJ: study design, data analysis, writing paper. WZX: data analysis, writing paper. DZX: data collection. YGY: data collection, data analysis, manuscript revision. DZY and BXH: study design, manuscript revision. LH: study design, data analysis, writing paper, manuscript revision. All authors read and approved the final manuscript.

**Acknowledgements**
The authors would like to thank Editage (www.editage.cn) for English language editing and senior engineer Qin-yuan Yu and engineer Yi-wei Chen for providing guidance and help in machine learning algorithms, Chongqing Abacus Software Co., Ltd.

**References**
1. Lan JY, Abbod MF, Yeh RG, Fan SZ, Shieh JS. Intelligent modeling and control in anesthesia. J Med Biol Eng. 2012;32(5):293-308.
2. Avidan MS, Jacobsohn E, Glick D, Burnside BA, Zhang L, Villafranca A, et al. Prevention of intraoperative awareness in a high-risk surgical population. N Engl J Med. 2011;365(7):591-600.
3. Misal US, Joshi SA, Shaikh MM. Delayed recovery from anesthesia: A postgraduate educational review. Anesth Essays Res. 2016;10(2):164-72.
4. Harris M, Chung F. Complications of general anesthesia. Clin Plast Surg. 2013;40(4):503-13.
5. Hemmings HC Jr, Akabas MH, Goldstein PA, Trudell JR, Orser BA, Harrison NL. Emerging molecular mechanisms of general anesthetic action. Trends Pharmacol Sci. 2005;26(10):503-10.
6. Chau PL. New insights into the molecular mechanisms of general anaesthetics. Br J Pharmacol. 2010;161(2):288-307.
7. Awareness ASoATFoI. Practice advisory for intraoperative awareness and brain function monitoring: a report by the american society of anaesthesiologists task force on intraoperative awareness. Anesthesiology. 2006;104(4):847–64.

8. Bruhn J, Röpcke H, Hoef t A. Approximate entropy as an electroencephalographic measure of anesthetic drug effect during desflurane. Anesthesiology. 2000;92(3):715-26.

9. Chan MT, Cheng BC, Lee TM, Gin T ; CODA Trial Group. BIS-guided anesthesia decreases postoperative delirium and cognitive decline. J Neurosurg Anesthesiol. 2013;25(1):33-42.

10. Kreuer S, Biedler A, Larsen R, Altmann S, Wilhelm W. Narcotrend monitoring allows faster emergence and a reduction of drug consumption in propofol–remifentanil. Anesthesiology. 2003;99(1):34-41.

11. Bruhn J, Myles PS, Sneyd R, Struys MM. Depth of anaesthesia monitoring: what's available, what's validated and what's next? Br J Anaesth. 2006;97(1):85-94.

12. Ye S, Park J, Kim J, Jung J, Jeon A, Kim I, et al. Development for the evaluation index of an anaesthesia depth using the bispectrum analysis. Int J Biol & Med Sci. 2007(2):66.

13. Voss L, Sleigh J. Monitoring consciousness: the current status of EEG-based depth of anaesthesia monitors. Best Pract Res Clin Anaesthesiol. 2007;21(3):313-25.

14. Wu L, Zhao H, Weng H, Ma D. Lasting effects of general anesthetics on the brain in the young and elderly: "mixed picture" of neurotoxicity, neuroprotection and cognitive impairment. J Anesth. 2019;33(2):321-35.

15. Dahaba AA. Different conditions that could result in the bispectral index indicating an incorrect hypnotic state. Anesth Analg. 2005;101(3):765-73.

16. Ahmed MU, Li L, Cao J, Mandic DP. Multivariate multiscale entropy for brain consciousness analysis. Annu Int Conf IEEE Eng Med Biol Soc. 2011;2011:810-3.

17. Merry AF, Cooper JB, Soyannwo O, Wilson IH, Eichhorn JH. International standards for a safe practice of anesthesia 2010. Can J Anaesth. 2010;57(11):1027-34.

18. Hsu CH, Tsai MY, Huang GS, Lin TC, Chen KP, Ho ST, et al. Poincaré plot indexes of heart rate variability detect dynamic autonomic modulation during general anesthesia induction. Acta Anaesthesiol Taiwan. 2012;50(1):12-8.

19. Huhle R, Burghardt M, Zaunseder S, Wessel N, Koch T, Malberg H, et al. Effects of awareness and nociception on heart rate variability during general anaesthesia. Physiol Meas. 2012;33(2):207-17.

20. Kanaya N, Hirata N, Kurosawa S, Nakayama M, Namiki A. Differential effects of propofol and sevoflurane on heart rate variability. Anesthesiology. 2003;98(1):34-40.

21. Liu Q, Ma L, Chiu RC, Fan SZ, Abbod MF, Shieh JS. HRV-derived data similarity and distribution index based on ensemble neural network for measuring depth of anaesthesia. Peer J. 2017;5:e4067.

22. Lee BR, Won DO, Seo KS, Kim HJ, Lee SW. Classification of wakefulness and anesthetic sedation using combination feature of EEG and ECG. 2017 5th International Winter Conference on Brain-Computer Interface (BCI). IEEE. 2017:88-90.
23. Akbarian B, Erfanian A. Automatic seizure detection based on nonlinear dynamical analysis of EEG signals and mutual information. Basic Clin Neurosci. 2018;9(4):227-40.

24. Wei Q, Liu Q, Fan SZ, Lu CW, Lin TY, Abbod MF, et al. Analysis of EEG via multivariate empirical mode decomposition for depth of anesthesia based on sample entropy. Entropy. 2013;15(9):3458-70.

25. Udhayakumar RK, Karmakar C, Palaniswami M. Understanding Irregularity Characteristics of Short-Term HRV Signals Using Sample Entropy Profile. IEEE Trans Biomed Eng. 2018;65(11):2569-79.

26. Al-angari HM, Sahakian AV. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. IEEE Trans Biomed Eng. 2007;54(10):1900-4.

27. Jiang GJ, Fan SZ, Abbod MF, Huang HH, Lan JY, Tsai FF, et al. Sample entropy analysis of EEG signals via artificial neural networks to model patients’ consciousness level based on anesthesiologists experience. Biomed Res Int. 2015;2015:343478.

28. Huang JR, Fan SZ, Abbod MF, Jen KK, Wu JF, Shieh JS. Application of multivariate empirical mode decomposition and sample entropy in EEG signals via artificial neural networks for interpreting depth of anesthesia. Entropy. 2013;15(9):3325-39.

29. Komatsu T, Kimura T, Sanchala V, Shibutani K, Shimada Y. Effects of fentanyl-diazepam-pancuronium anesthesia on heart rate variability: a spectral analysis. J Cardiothorac Vasc Anesth. 1992;6(4):444-8.

30. Moak JP, Goldstein DS, Eldadah BA, Saleem A, Holmes C, Pechnik S, et al. Supine low-frequency power of heart rate variability reflects baroreflex function, not cardiac sympathetic innervation. Heart Rhythm. 2007;4(12):1523-9.

31. Ferreira AL, Nunes CS, Mendes JG, Amorim P. Usefulness of the Blink Reflex to Assess the Effect of Propofol During Induction of Anesthesia in Surgical Patients. XV Mediterranean Conference on Medical and Biological Engineering and Computing-MEDICON 2019. 2020:1057-63.

32. Shalbaf A, Shalbaf R, Saffar M, Sleigh J. Monitoring the level of hypnosis using a hierarchical SVM system. J Clin Monit Comput. 2020;34(2):331-8.

33. Lee HC, Ryu HG, Park Y, Yoon SB, Yang SM, Oh HW, Jung CW. Data driven investigation of bispectral index algorithm. Sci Rep. 2019;9(1):13769.

34. Gu Y, Liang Z, Hagihira S. Use of multiple EEG features and artificial neural network to monitor the depth of anesthesia. Sensors (Basel). 2019;19(11):2499.

35. Quian Quiroga R, Sakowitz OW, Basar E, Schürmann M. Wavelet transform in the analysis of the frequency composition of evoked potentials. Brain Res Brain Res Protoc. 2001;8(1):16-24.

36. Güler I, and Ubeyli ED. Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. J Neurosci Methods. 2005;148(2):113-21.

37. Bose NK, Liang P. Neural network fundamentals with graphs, algorithms, and applications. Mcgraw-Hill; 1996.

38. Huang HH, Lee YH, Chan HL, Wang YP, Huang CH, Fan SZ. Using a short-term parameter of heart rate variability to distinguish awake from isoflurane anesthetic states. Med Biol Eng Comput.
39. Pilge S, Schneider G. BIS and state entropy of the EEG - comparing apples and oranges. Br J Anaesth. 2015;115(2):164-6.
40. Kunze DL. Reflex discharge patterns of cardiac vagal efferent fibres. J Physiol. 1972;222(1):1-15.
41. Palazzolo JA, Estafanous FG, Murray PA. Entropy measures of heart rate variation in conscious dogs. Am J Physiol. 1998;274(4):H1099-105.
42. Win NN, Kohase H, Yoshikawa F, Wakita R, Takahashi M, Kondo N, et al. Haemodynamic changes and heart rate variability during midazolam-propofol co-induction. Anaesthesia. 2007;62(6):561-8.
43. Ireland N, Meagher J, Sleigh JW, Henderson JD. Heart rate variability in patients recovering from general anaesthesia. Br J Anaesth. 1996;76(5):657-62.

Tables

Table 1 Patients demographics and clinical characteristics.

| Parameters                                 | means (SD) |
|--------------------------------------------|------------|
| Age (year)                                 | 50.2(7.0)  |
| Height (cm)                                | 160.6(6.9) |
| Weight (kg)                                | 61.1(9.4)  |
| BMI (kg m^{-2})                            | 23.7(3.2)  |
| Duration of surgery (min)                  | 132.9(48.4)|
| Anaesthetic management                      | /          |
| Midazolam induction (mg)                    | 3.0(0.8)   |
| Propofol induction (mg)                     | 62.0(10.3) |
| Sufentanil induction (mg)                   | 20.2(2.7)  |
| Cis-atracurium (mg)                         | 13.1(1.9)  |
| Maintenance drugs infusion rate             | /          |
| Sevoflurane maintenance (Vol%)             | 1.7(0.4)   |
| Propofol maintenance (mg•kg^{-1}•h^{-1})    | 2.1(0.3)   |
| Remifentanil (µg•kg^{-1}•h^{-1})            | 0.1(0.04)  |
| Additional drugs administrated when         | /          |
| approaching the end of surgery              | /          |
| Sufentanil (mg)                             | 7.1(3.2)   |
| Atropine (mg)                               | 0.3(0.1)   |
| Neostigmine (mg)                            | 0.7(0.2)   |

Values are means (SD). BMI, body mass index.

Table 2 Comparison of logistic regression, support vector machine, decision tree, and deep neural network.

|                  | Precision of anaesthesia induction | Recall of anaesthesia induction | Precision of anaesthesia maintenance | Recall of anaesthesia maintenance | Precision of anaesthesia recovery | Recall of anaesthesia recovery | Classification accuracy |
|------------------|-----------------------------------|--------------------------------|-------------------------------------|-----------------------------------|---------------------------------|-------------------------------|-------------------------|
| LR               | 55.1%                             | 81.2%                          | 94.6%                               | 94.1%                             | 46.3%                           | 47.5%                         | 86.2%                   |
| SVM              | 55.7%                             | 80.1%                          | 95.1%                               | 94.6%                             | 47.1%                           | 46.8%                         | 87.5%                   |
| DT               | 56.1%                             | 80.9%                          | 95.6%                               | 94.8%                             | 47.3%                           | 47.1%                         | 87.2%                   |
| DNN              | 58.1%                             | 88.1%                          | 96.0%                               | 94.7%                             | 56.6%                           | 57.8%                         | 90.1%                   |

LR, logistic regression. SVM, support vector machine. DT, decision tree. DNN, deep neural network.