EEG signal processing in anaesthesia. Use of a neural network technique for monitoring depth of anaesthesia

O. Ortolani1*, A. Conti2, A. Di Filippo1, C. Adembri1, E. Moraldi1, A. Evangelisti3, M. Maggini4 and S. J. Roberts5

1Dipartimento Area Critica Medico Chirurgica, Sez. di Anestesiologia e Rianimazione, Università di Firenze, Italy. 2Dipartimento Di Patologia e Biologia Molecolare e Cellulare, Universita’ Federico II, Napoli, Italy. 3Dipartimento di Sistemi e Informatica, Universita’ di Firenze, Italy. 4Dipartimento di Ingegneria dell’Informazione, Universita’ di Siena, Italy. 5Department of Engineering Science, University of Oxford, Oxford, UK

*Corresponding author: Via Carducci 42, I-80121 Naples, Italy

Background. The Bispectral Index (BIS) is a proprietary index of anaesthesia depth, which is correlated with the level of consciousness and probability of intraoperative recall. The present study investigates the use of a neural network technique to obtain a non-proprietary index of the depth of anaesthesia from the processed EEG data.

Methods. Two hundred patients, who underwent general abdominal surgery, were recruited for our trial. For anaesthesia we used a total i.v. technique, tracheal intubation, and artificial ventilation. Fourteen EEG variables, including the BIS, were extracted from the EEG, monitored with an EEG computerized monitor, and then stored on a computer. Data from 150 patients were used to train the neural network. All the variables, excluding the BIS, were used as input data in the neural network. The output targets of the network were provided by anaesthesia scores ranging from 10 to 100 assigned by the anaesthesiologist according to the observer’s assessment of alertness and sedation (OAA/S) and other clinical means of assessing depth of anaesthesia. Data from the other 50 patients were used to test the model and for statistical analysis.

Results. The artificial neural network was successfully trained to predict an anaesthesia depth index, the NED (neural network evaluated depth), ranging from 0 to 100. The correlation coefficient between the NED and the BIS over the test set was 0.94 (P<0.0001).

Conclusion. We have developed a neural network model, which evaluates 13 processed EEG parameters to produce an index of anaesthesia depth, which correlates very well with the BIS during total i.v. anaesthesia with propofol.

Br J Anaesth 2002; 88: 644–8

Keywords: anaesthesia, depth; anaesthetics i.v., propofol; analgesics opioid, remifentanil; brain, electroencephalography; monitoring, bispectral index; monitoring, electroencephalography

Accepted for publication: January 3, 2002

Evaluation of the processed EEG seems to provide important information on the depth of anaesthesia. Bispectral analysis is an EEG frequency domain method of analysis that examines the correlation of phase between signal components. More specifically, the bispectral analysis may be used to quantify the amount of synchronization in the EEG. The combination of the time domain (burst suppression), frequency domain (beta ratio), and high order spectral variables (bispectral analysis) can provide a numerical index (Bispectral Index—BIS) on a normalized scale ranging between 0 and 100. Indeed, this can be achieved using a proprietary algorithm (Aspect Medical Systems, Natick, USA). The relationship between the BIS and the hypnotic effects provided by propofol, alfentanil, nitrous oxide, sevoflurane, and isoflurane have been described recently. The authors concluded that the BIS is correlated to
Table 1 Processed EEG variables monitored during anaesthesia. All absolute power variables are reported in dB with respect to 0.0001 mV²

<table>
<thead>
<tr>
<th>Score</th>
<th>Arbitrary value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1=Alert</td>
<td>100</td>
</tr>
<tr>
<td>2=Light sedation with execution of verbal commands</td>
<td>80</td>
</tr>
<tr>
<td>3=Sedation with movements after light pain stimuli</td>
<td>70</td>
</tr>
<tr>
<td>4=Sedation with movements after intense pain stimuli</td>
<td>60</td>
</tr>
<tr>
<td>5=Light hypnotic state (asleep)</td>
<td>50</td>
</tr>
<tr>
<td>6=Moderate hypnotic state</td>
<td>40</td>
</tr>
<tr>
<td>7=Deep hypnotic state</td>
<td>20</td>
</tr>
<tr>
<td>8=Very deep hypnotic state</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 2 Score of anaesthesia depth and correspondent assigned arbitrary value. Heart rate and arterial pressure variations are expressed vs basal value

Anaesthesia procedure

Patients received ringer-lactate 700 ml and atropine 0.01 mg kg⁻¹ as pre-medication. Anaesthesia was induced using a total i.v. technique: remifentanil 15 µg kg⁻¹ h⁻¹ (no initial bolus) and propofol 1.5 mg kg⁻¹ as a bolus, followed by an infusion at 10 mg kg⁻¹ h⁻³, pancuronium 0.1 mg kg⁻¹. The trachea was intubated and the lungs mechanically ventilated with oxygen/air (FIO₂=0.4). The amounts of remifentanil and propofol administered were monitored during anaesthesia and varied as needed according to the BIS values and the anaesthetist's clinical evaluation. Patients who needed an adjustment of more than 10% from the initial infusion rates were not included in the trial.

EEG monitoring and data acquisition

Fourteen EEG variables were extracted from the EEG (Table 1) and monitored with an Aspect A-1000 EEG monitor (Aspect Medical Systems, Natick, USA) using four Zipprep self-sticking frontal surface electrodes placed on both sides of the outer malar bone (At1 and At2) with Fpz as reference and Fp1 as the ground. Every 5 s the Aspect A-1000 calculated the BIS and the other EEG variables based on a running average of the last 120 artefact-free epochs of data (each epoch represents 0.5 s). Data collected during the whole anaesthesia procedure at 5 s intervals, starting 3 min before the induction and ending when the patient was completely awake, were automatically sent through a serial RS232 interface to a computer (Toshiba, Satellite 4000 CDT) using in-home software.

An assessment score ranging from 1 to 8 indicating the depth of anaesthesia was made by the anaesthetist at each

Patients and methods

Patients

After obtaining institutional ethics committee approval and written informed consent, 200 patients were recruited for our trial. They underwent general abdominal surgery lasting up to 3 h, were of either sex, aged between 20 and 55 yr, weighed between 60 and 75 kg, were of normal size, and categorized as ASA I. The patients’ heart rate, non-invasive arterial pressure, and arterial oxygen saturation were recorded during anaesthesia. Patients were not included in the ANN training database if, at any time, they had during intervention a heart rate less than 55 or greater than 120 beats min⁻¹, systolic arterial pressure less than 65 or greater than 170 mm Hg, or arterial oxygen saturation less than 90%.

In this study we aimed to use an artificial neural network (ANN) to integrate different processed EEG variables in order to estimate the depth of anaesthesia. ANNs can provide models of non-linear or complex systems such as EEG signals where the informative content and data associations are too complex to be extracted by traditional algorithms. The index of anaesthetic depth we derived was compared with the BIS.

the level of consciousness and the probability of intraoperative recall³⁻⁵

In this study we aimed to use an artificial neural network (ANN) to integrate different processed EEG variables in order to estimate the depth of anaesthesia. ANNs can provide models of non-linear or complex systems such as EEG signals where the informative content and data associations are too complex to be extracted by traditional algorithms. The index of anaesthetic depth we derived was compared with the BIS.
time point. During induction and arousal, this score was based on the observer’s assessment and evaluation of alertness and sedation score: OAA/S, ranging from 1 (alert) to 5 (asleep). During surgery the depth of anaesthesia score (from 6 to 8) was based on clinical anaesthesia indices (heart rate, mean arterial pressure, perspiration, lacrimation, eyelid reflex) to distinguish between ‘moderate’, ‘deep’, and ‘very deep hypnotic’ states. Arbitrary discrete values from 100 to 10 were assigned to each state as indicated in Table 2.

**ANN training and testing**

Thirteen EEG variables (Table 1) were used as the input for the ANN (Neuralware NeuralSIM obtained by Aspen Technology, Pittsburgh, PA, USA). The output target was the depth of anaesthesia score assigned by the anaesthesiologist at the same time. Data from 150 patients were collected in order to train the ANN. Data from the other 50 patients were used to test the ANN and for statistical analysis. The Aspect A-1000 gives a signal quality index (SQI) with each record. The SQI is the percentage of good epochs in the last 120 epochs. Data from records with a SQI less than 50% were discarded from the training database. Data recorded during the use of surgical diathermy were also excluded.

The analysis of the ANN performance on the training set led us to choose a final multi-layer perceptron (MLP) network with 13 input nodes, one hidden layer with 18 nodes, and one output node. The network was trained using the standard back-propagation method. The network learning rate was 0.38 and the momentum was 0.8. The NeuralSIM software took 10 h to optimize the network, and the ANN was then trained for 17 h. Training was stopped when the average absolute error was less than 4 and the root mean square error was less than 5. The resulting neural network evaluated depth (NED) ranged from 0 to 100. During training the ANN calculated the correlation coefficient between the target outputs and the corresponding predicted values (NED) produced by the network. The correlation coefficient of the final network (Table 4) shows that the ANN was satisfactorily trained.

**Statistical analysis**

Multiple regression analysis was performed between the EEG processed variables and the BIS in order to evaluate the linear correlations. Pearson’s correlation coefficient was calculated for the BIS and NED data of the test group of patients. The data were then plotted according to the method of Bland and Altman; in order to evaluate the agreement between the methods and the bias, the difference between the methods was plotted against their mean. Statistical analyses were performed using the Graphpad InStat utility, v. 3.00 for Windows (GraphPad Software, Inc., San Diego, CA, USA) and the Analyse-it Software for Microsoft Excel, v. 1.62 (Leeds, UK).

**Results**

Data from nine patients who showed unstable EEG signals (too much artefact) were not included in the training of the ANN. The anthropometric data of the patients included in the training and test groups are listed in Table 3. Figure 1 is a patient recruitment flow chart. A total of 195 000 records from 141 patients (approximately 279 h of anaesthesia) were used to train the ANN. Another 74 500 records from 50 patients (approximately 95 h of anaesthesia) were used to test the ANN. The correlation coefficient calculated by multiple regression analysis between the 13 EEG processed variables and the BIS for the training data was 0.88. The ANN calculated a correlation coefficient between the target assigned values and the corresponding predicted NED of 0.98 (Table 4). The root mean square error between target

---

**Table 3** Antropometric data of the patients included in the training and in the testing group

<table>
<thead>
<tr>
<th></th>
<th>Training group</th>
<th>Testing group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>76</td>
<td>26</td>
</tr>
<tr>
<td>Females</td>
<td>65</td>
<td>24</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>68.2 (6.4)</td>
<td>64.6 (7.9)</td>
</tr>
<tr>
<td>Age (yr) (range)</td>
<td>39.0 (20–58)</td>
<td>37.8 (19–56)</td>
</tr>
</tbody>
</table>
and predicted outputs during the test was 5.9, with an average absolute error of 4 and a maximum absolute error of 29. The percentage of predicted outputs that fell within the tolerance of 20% of the corresponding target outputs was 0.9 and the confidence interval (95%) was 8.9 for the data used in the training. This statistical evaluation is summarized in Table 5. The Pearson's correlation coefficient \( r \) between BIS and NED in the test set was 0.9411 (Fig. 2). The bias calculated by the Bland and Altman method was ±0.199. The limits of agreement were ±10.19 (lower) and 9.79 (upper), indicating very little bias and a very good agreement (Fig. 3).

**Discussion**

The BIS gives a simple and quantitative indication of the depth of anaesthesia on a 0 to 100 normalized scale. It may not prove, however, to be the optimal system because part of the information present in the EEG signals is not utilized. Our experiments show that neural networks can be used to analyse processed EEG data to provide a depth of anaesthesia index as informative as the BIS, during total i.v. anaesthesia with propofol.

The BIS may also prove not robust enough when artefactual signals are present. BIS alterations have been found with pacers during cardiac surgery, with some electrical blankets and, in eye surgery, a sudden increase of the BIS may be observed when some vitrectomy coagulators are used. Automated recognition systems based on autoregressive modelling and neural network analysis of the EEG signals can, in theory, be trained to recognize and filter noise artefacts. Experiments confirm that neural network analysis of the EEG achieves good discrimination between awake and anaesthetized states both in i.v. anaesthesia and in anaesthesia with volatile agents with a good rejection of artefactual signals. The authors conclude that the flexibility and non-linearity of an ANN approach are important factors providing reliability to a
monitoring device for depth of anaesthesia. Naturally a good performance of an ANN highly depends on the quality of training.

Statistical evaluation of our data demonstrates that the Pearson’s correlation coefficient $r$ calculated between NED and BIS is higher than the $r$ coefficient obtained with the multiple linear regression model by correlating the 13 EEG variables (used to calculate the NED) with the BIS. This is an obvious indication that the performance of the neural model is better because the correlation between the EEG processed data and the BIS are not completely linear.

In conclusion, we have developed a neural network based system that can evaluate the depth of anaesthesia from 13 processed EEG variables (excluding BIS). Comparison with correspondent BIS values confirms that neural networks can deal with the data from processed fronto-parietal EEG, where phase information is lost, producing evaluations on the depth of anaesthesia, such as the NED, that can perform as well as the BIS in accuracy and reliability during total i.v. anaesthesia with propofol. Neural network performance may be improved further by using bispectral analysis data.

Finally, we have to consider that the information included in the analysis of the processed EEG is contained in the raw EEG signal. A neural network could, therefore, be trained to predict anaesthesia depth directly from the raw EEG. Successful experiments along these line have been performed by using neural network technology associated with autoregressive models and stochastic complexity measures.14 15

We regard our initial results as extremely encouraging for the evaluation and control of the depth of anaesthesia, both in open and closed loop systems, via neural network based EEG analysis. Trials are in progress to extend the NED evaluation to other anaesthesia regimens.

References
2 Rampil I. A primer for EEG signal processing. Anesthesiology 1998; 89: 980–1002
3 Gan TJ, Glass PS, Windsor MD, et al. Bispectral Index monitoring allows faster emergence and improved recovery from propofol, alfentanil and nitrous oxide anaesthesia. Anesthesiology 1997; 87: 808–15
4 Olofsen E, Dahan A. The dynamic relationship between end-tidal sevoflurane and isoflurane concentrations and bispectral index and spectral edge frequency of the electroencephalogram. Anesthesiology 1999; 90: 1345–53
9 Gallagher JD. Pacer-induced artifact in the Bispectral Index during cardiac surgery Anesthesiology 1999; 90: 636