Comparison of bispectral index, 95% spectral edge frequency and approximate entropy of the EEG, with changes in heart rate variability during induction of general anaesthesia†

J. W. Sleigh1 and J. Donovan2

1 Anaesthetics Department, Waikato Hospital, Hamilton, New Zealand. 2 Wellington Clinical School, Wellington Hospital, Wellington, New Zealand

We have compared bispectral index (BIS), 95% spectral edge frequency (SEF) and approximate entropy (ApEn) in 37 patients during induction and recovery from a short general anaesthetic. Heart rate variability (HRV) was also compared during induction only. These indices were noted at the start of induction, when a syringe held between the thumb and fingertips was dropped, at insertion of a laryngeal mask or tracheal tube (tube insertion), at incision, at the end of surgery, on return of the gag reflex and when the patient could follow a verbal command. When indices at the start of induction were compared with those at tube insertion, all four decreased significantly. BIS decreased from a mean of 95.38 (SEM 1.02) to 44.22 (1.05), mean SEF from 20.91 (1.19) to 14.14 (0.70) Hz, mean HRV from 37.1 (7.75) to 17.9 (3.6) bpm² and ApEn from 0.90 (0.06) to 0.65 (0.04). Using logistic regression, the indices were compared both individually and in combination as to the power of distinguishing awake (at pre-induction) from asleep (at tube insertion) states. BIS had the best predictive power, with a sensitivity of 97.3%, specificity 94.4%, positive predictive value 94.7% and negative predictive value 97.1%. A combination of the indices conferred no additional predictive advantage.

Br J Anaesth 1999; 82: 666–71

Keywords: monitoring, electroencephalography; monitoring, depth of anaesthesia; anaesthesia, general; measurement techniques, spectral analysis; anaesthesia, depth; heart, heart rate

Accepted for publication: November 27, 1998

Estimation of ‘depth of anaesthesia’ is an ongoing problem in anaesthesia. In our study, we have compared three EEG indices and heart rate variability as measures quantifying the level of anaesthesia at various stages of general anaesthesia.

Bispectral analysis is a mathematical technique of signal processing that quantifies the degree of phase coupling between different frequency components within a signal. In the present commercially available system (A-1000, Aspect medical systems, Natick, MA, USA), output is in the form of a dimensionless bispectral index (BIS) that varies from 0 to 100. This index is a composite quantity based on both the extent of the coupling between frequencies, and a broader statistical discrimination of the awake and asleep states using a previous EEG database. In the awake state, the frequencies are tightly coupled and BIS is close to 100. One advantage of the BIS is that, because it estimates phase difference, it is largely independent of the amplitude of the EEG.

The 95% spectral edge frequency (SEF) is the frequency below which 95% of the EEG spectral power resides. Therefore, SEF reflects the shift in the EEG signal from high frequencies in the awake state to predominantly low frequencies in the anaesthetized state. Unlike BIS, however, SEF does not take into account phase relationships within the EEG signal.

Approximate entropy (ApEn) is a family of statistics related to the Kolmogorov entropy and is claimed to be a non-linear quantification of the regularity of a signal. A perfectly regular signal would have an entropy of zero. The practical applications of the Kolmogorov entropy are limited by the requirement for almost noiseless data, especially when dealing with systems of high dimensionality. ApEn seeks to overcome this limitation by making data comparisons on a scale larger than that of noise. ApEn has been used with mixed success to quantify changes in fetal heart rate variability. In summary, each EEG index examines...
independent aspects of the complex waveform that is the EEG. SEF extracts frequency information from the EEG, BIS extracts phase information and ApEn regularity.

Spectral analysis of heart rate variability (HRV) is measured from the ECG signal, and is thought to reflect in part the cardiovascular regulation by the autonomic nervous system. General anaesthesia markedly depresses HRV and therefore HRV has been suggested as a measure of depth of anaesthesia. By convention, HRV has been separated into high (>0.1 Hz) and low (0.05–0.1 Hz) frequency variation in RR intervals. High-frequency changes are considered to be almost exclusively a function of parasympathetic regulation and are caused by respiration (respiratory sinus arrhythmia). There is good evidence that the afferent part of the reflex arc is mediated mainly via baroreceptor reflexes in healthy subjects. The low-frequency range is considered to be modulated by both sympathetic and parasympathetic supplies to the heart and is related to both baroreflex and respiratory activity. We did not examine HRV through time frequency analysis, as other studies have, but rather we approximated the frequency domain measure (spectral power) in the time domain using trend removal with different length epochs.

It may be surmised that changes in HRV that occur during anaesthesia could reflect more closely the brainstem effects of anaesthesia. In contrast, EEG indices measure cortical effects. For this reason, we hypothesized that a combined measure of some of these variables may quantify the depth of anaesthesia more reliably than a single index. The aims of this study were to compare the usefulness of the EEG and HRV variables in differentiating between the awake and asleep states in patients undergoing routine anaesthesia. The variables were compared singly and in combination.

**Patients and methods**

**Data collection**

The study was approved by the Waikato Ethics Committee and informed consent was obtained from adult patients before undergoing minor surgery under general anaesthesia. EEG data were collected using the ASPECT A-1000 monitor (Aspect Medical Systems, Natick, MA, USA, software version 3.12). While preparing for anaesthesia, five silver–silver chloride EEG pads (Zipprep, Aspect Medical Systems, USA) were attached to the patient’s forehead according to a standard montage, one as the ground (Fpz) and the other four as two separate bipolar channels (F7-T7, F8-T8) over the left and right prefrontal cortex. The low- and high-frequency filters were set at 0.5 Hz and 50 Hz, respectively. A three-lead ECG (lead V5) was also attached to the patient to measure heart rate variability and respiration (Life Scope 9, Nihon Kohden Corporation, Tokyo, Japan).

After a 5-min period for stabilization, recording was begun at the start of injection of the i.v. anaesthetic induction agent (propofol). Data from the EEG monitor and ECG were transferred continuously to two IBM computers for off-line analysis. The timing of specific events was noted: (1) start of injection (start); (2) dropping a 50-ml syringe which was held between the thumb and fingertips (syringe drop); (3) insertion of a laryngeal mask or tracheal tube (tube); and (4) initial incision (incision). In 28 patients we also recorded the time at which they stopped responding to verbal command (‘open your eyes’). This time was either coincident with or occurred slightly before the syringe drop time (mean time difference 9 s). The recording was then stopped until the end of surgery when it was started again at the time that the anaesthetic had been switched off.

The above variables were recorded continuously and specific events were noted: (5) end of surgery; (6) return of the gag reflex (gag); (7) and ability to follow a verbal command (squeeze hand). After the ability to follow a verbal command had been reached, the recording was stopped. Because of variations in the time to awakening and the format of our R wave detection program, we were unable to record the ECG variables during the recovery period.

Patients were anaesthetized according to the routine clinical judgement of the anaesthetist. The anaesthetic technique involved premedication with midazolam 1–2 mg i.v. and fentanyl 50–100 µg i.v., 5 min before induction of anaesthesia with propofol 80–300 mg given over 13–118 s, titrated to patient response. All patients were sufficiently awake after administration of midazolam to co-operate and hold the syringe. In 12 patients a non-depolarizing neuromuscular blocking agent was given and the trachea was intubated. In the remainder, a laryngeal mask was inserted. Anaesthesia was maintained with nitrous oxide and isoflurane (FIO₂ 0.33–0.40). The anaesthetist was blind to the EEG and HRV data. If skin impedance for the EEG electrodes was greater than 10 kW the results were discarded (n=3). Because we wanted to compare the accuracy of the various indices under routine clinical conditions, no attempt was made to influence the anaesthetist’s standard clinical practice.

**Data analysis**

BIS and SEF were calculated by the ASPECT monitor in 5-s epochs, with no smoothing. The monitor has built-in artefact rejection algorithms that preclude display of the indices if they detect significant artefact (usually because of movement of electrodes or diathermy). ApEn was calculated from 2-s (n=512 points) epochs of raw EEG data using the published algorithm. Because ApEn is a family of statistics, there are two further variables that have to be decided. The length of the compared runs (m) was set at 2, as has been used in all published reports. Second, the level at which noise was filtered was set at 5 µV. This is within the recommended range of 0.1–0.25 of the SD of the EEG data.
Calculation of heart rate variability (see appendix)
The primary method of quantifying HRV was by using a simple time domain method related to the SDNN,\textsuperscript{7} which we have termed the 20SecHR\textsubscript{rms} (if using data from 20-s epochs) and 10SecHR\textsubscript{rms} (if using data from 10-s epochs). In order to compare this method with accepted frequency domain methods of HRV analysis, we also derived power spectral densities and transfer functions from the 20-s epochs of heart rate and respiration data. We used the methods developed by Saul and colleagues\textsuperscript{10} implemented using a commercially available computer package (MatLab 5.2, The Mathworks, USA). Because of the rapid changes during induction, and therefore the short segments of heart rate data used, we have reported only high-frequency spectral power (0.1–0.5 Hz).

A purpose-written computer program was used to detect R wave peaks, measure the time between subsequent RR intervals and to edit artefacts. The instantaneous minute volume was then obtained, together with HRV, using another purpose-written program.

Calculation of 20SecHR\textsubscript{rms}
The time domain analogue of spectral power is the variance of the signal. If each epoch of data is de-trended linearly, the root mean square of the deviations from the trend line gives the square root of the total power in the power spectrum. The lower limit of detectable frequency is given by the length of the data epoch. For example, if the data epochs are 10 s each, we are unable to detect power below 0.1 Hz because it has been removed by the de-trending. Thus by comparing the root mean square deviations for different epoch lengths, we are effectively able to examine variation within different frequency bands. In our study, ‘total’ power (>0.05 Hz) was obtained by calculating the deviations in 20-s data epochs, and high-frequency power using 10-s epochs. The difference gives the low-frequency power. In addition, the simple ratio between minute volume and HRV in the high-frequency range (termed ‘gain’) can be calculated.

Approximate effect-site propofol concentrations in the cerebrum were calculated using the commercially available IVA-SIM program.\textsuperscript{11} Calculations in the simulation program were based on the assumption that the drug follows linear pharmacokinetics and that a multicompartment model is applicable. This program uses a three-compartment exponential to model the pharmacokinetics combined with a standard Hill equation to model pharmacodynamic and biophase hysteresis. We entered both the size of the induction bolus and the time taken for injection of the bolus. The program then calculated the estimated effect-site concentration of propofol at 20-s intervals. From this we were able to obtain a crude estimate of the cerebral concentration of propofol at each stage of induction of anaesthesia. It was found that the maximal cerebral propofol concentration occurred consistently at approximately 3 min, in agreement with measurements suggested by Ludbrook and Upton.\textsuperscript{12}

The paired t test with Bonferroni correction for multiple comparisons was used to compare the difference between baseline indices and the rest of the data (at syringe drop time, tube insertion, first surgical procedure, gag time and awake time). The ability of the various indices to differentiate between the awake (start time) and asleep (tube time) states was compared using single and multiple logistic regression models. Initially the model was developed using each single variable, and finally using a combination of all variables. We used high-frequency power as the HRV indicator. From these models we determined the sensitivity, specificity, positive predictive value and negative predictive value of each of the indices for this data set.

Results
The study involved 37 (26 females) ASA I or II patients, aged 16–80 yr, undergoing surgery requiring general anaesthesia. No patient had intraoperative recall.

HRV
Both the time domain and frequency domain methods gave similar results (Table 1). Values decreased significantly during induction of anaesthesia, being most marked for HF power. This has been thought previously to indicate withdrawal of parasympathetic regulation as the patient loses consciousness. However, when examining HF power as a fraction of total power (HF/Total) (an indicator of sympathovagal balance) it can be seen that parasympathetic regulation decreased significantly at the time of syringe drop only (i.e. early in induction). At the same time the ‘input’ power (depth of respiration) is much diminished or absent because the patient is commonly underventilating or apnoeic. The question therefore arises as to how much of the decrease in HF is caused by lack of input (i.e. apnoea) and how much is attributable to lack of response (principally the baroreceptor reflex arc). Our results showed that the ‘gain’ (or transfer function) at the time of syringe drop did not decrease significantly. This suggests that a significant proportion of the initial decrease in HF HRV may be caused by hypopnoea rather than a direct decrease in reflex responsiveness of heart rate to changes in lung volume. The ‘gain’ does not significantly decrease until the time of incision when anaesthetic drug concentrations are higher.

BIS
BIS decreased during anaesthesia (Fig. 1) from a mean of 95.38 (SEM 1.02) at the start to 44.22 (1.05) at tube insertion. However, it should be noted there was no significant difference between BIS at the start and at syringe drop (92.00 (1.90)). During the recovery period, there was no significant difference between BIS at the end of surgery and at gag time. This suggests that BIS is not a reliable indicator of when the patient first starts to gag. Obviously, the gag reflex is a brainstem reflex and therefore may not reliably mirror cortical activity. There was a large overlap
BIS, SEF and ApEn compared with HRV during induction

Table 1 Mean (SEM) changes in HRV at induction of anaesthesia. High frequency (HF) and Total are the root mean variance (10SecHR rms or 20SecHR rms) of de-trended 10-s and 20-s epochs, respectively, and are expressed in units of bpm^2. HF/Total is a dimensionless ratio and reflects sympathovagal balance. Gain (HF/minute ventilation) reflects the strength by which respiration alters heart rate (respiratory sinus arrhythmia) and has units of bpm litre^-1 min^-1. HF (PSD) is the high-frequency power determined from the power spectrum of the heart rate signal. Units are bpm^2. TF refers to the area under the lung volume-to-heart rate transfer function in the frequency range 0.1–0.5 Hz and units are bpm litre^-1. *P<0.05 vs start.

<table>
<thead>
<tr>
<th></th>
<th>HF</th>
<th>Total</th>
<th>HF/Total</th>
<th>Gain</th>
<th>HF (PSD)</th>
<th>TF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>37.1 (7.7)</td>
<td>39.9 (6.3)</td>
<td>0.90 (0.08)</td>
<td>6.4 (1.9)</td>
<td>51.2 (13.7)</td>
<td>8.8 (1.4)</td>
</tr>
<tr>
<td>Syringe drop</td>
<td>13.4 (2.9)*</td>
<td>29.5 (5.0)*</td>
<td>0.57 (0.06)*</td>
<td>3.6 (1.1)</td>
<td>8.3 (2.6)*</td>
<td>10.2 (4.1)</td>
</tr>
<tr>
<td>Tube insertion</td>
<td>17.9 (3.6)*</td>
<td>30.6 (5.4)</td>
<td>0.63 (0.08)</td>
<td>4.9 (1.4)</td>
<td>12.5 (3.4)*</td>
<td>7.2 (1.3)</td>
</tr>
<tr>
<td>Incision</td>
<td>19.1 (4.6)*</td>
<td>22.9 (4.3)*</td>
<td>0.77 (0.07)</td>
<td>2.1 (0.06)*</td>
<td>9.4 (1.4)*</td>
<td>4.3 (1.4)</td>
</tr>
</tbody>
</table>

between BIS at the end of surgery and at awake (mean 79.65 (4.05)), indicating that BIS is not totally reliable when the patient is waking. Fourteen patients were able to follow a verbal command at a BIS values of less than 75 and six patients at a BIS value of less than 60. This could have been caused in part by a lag in the updating of the displayed values because BIS usually increased rapidly to more than 95 within the next 20–30 s. Conversely, it was also noticed that four patients were unable to follow a verbal command with a BIS value of more than 95.

We also compared BIS with the calculated effect-site propofol concentration and found a significant correlation (Pearson’s linear correlation coefficient, r=–0.74). However, there was a wide range of values. For example, a calculated drug concentration of 1.7 ng ml^-1 was associated with a BIS range of 15–98.

**SEF**

Changes in SEF (Fig. 2) were similar to those of BIS, but with greater overlap between the different stages during induction, decreasing from a mean of 20.91 (1.19) Hz at the start to 14.14 (0.70) Hz at tube insertion. In agreement with previous reports, the correlation coefficient for the association between SEF and the calculated effect-site propofol concentration was less (r=0.55) than that for BIS, although this was not compared statistically. Mean SEF at the point of responding to verbal command was 23.56 (0.085) Hz.

**ApEn**

ApEn decreased significantly with increasing depth of anaesthesia (Fig. 3) from a mean starting value of 0.90 (0.06) to 0.65 (0.04) at tube insertion. It changed less during recovery, reaching a mean value of only 0.77 (0.08) at the point of responding to verbal command.

**Comparisons of indices**

BIS was the best of the logistic regression models in all respects, and HRV the worst (Table 2). The combination of indices did not improve on BIS.
Table 2 Comparison of the power to distinguish the awake state, using logistic regression for each individual variable, and the combination of variables.

<table>
<thead>
<tr>
<th></th>
<th>BIS</th>
<th>SEF</th>
<th>HRV</th>
<th>ApEn</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>97.3%</td>
<td>70.2%</td>
<td>69.3%</td>
<td>75.7%</td>
<td>96.3%</td>
</tr>
<tr>
<td>Specificity</td>
<td>94.4%</td>
<td>69.4%</td>
<td>59.1%</td>
<td>71.4%</td>
<td>95.2%</td>
</tr>
<tr>
<td>PPV</td>
<td>94.7%</td>
<td>70.3%</td>
<td>64.0%</td>
<td>73.7%</td>
<td>96.3%</td>
</tr>
<tr>
<td>NPV</td>
<td>97.1%</td>
<td>69.4%</td>
<td>54.2%</td>
<td>73.5%</td>
<td>95.2%</td>
</tr>
</tbody>
</table>

**Discussion**

As has been suggested from previous studies, we found that BIS was superior to other EEG indices of depth of anaesthesia, but was not ideal in differentiating between awake and asleep states. HRV was less reliable than any EEG index as an indicator of consciousness at induction. Combinations of indices did not seem to offer any advantage.

When observing the changes in HRV during induction of anaesthesia, our results from both time domain (105secHRmean) and traditional frequency domain methods were similar to those obtained by Huang and colleagues. These investigators used a more computationally intense time-frequency spectral analysis method to quantify HRV. Although we were not directly comparing methods, it would seem that our simple algorithm may give comparable results, which could be easily implemented in real time.

Several studies have demonstrated a sharp decrease in HRV at induction of general anaesthesia. Most have used frequency domain techniques to quantify HRV. In contrast, Pomfrett and co-workers developed a unique method of estimating respiratory sinus arrhythmia by detecting statistical clusters of phase-locking between heart rate and breathing. Their work has shown some promise as a method of detecting the level of consciousness during anaesthesia, although there seems to be much inter-patient variation limiting the reliability of the technique. Large scale validation studies are awaited.

The other important aspect to our HRV results was that most of the observed early decrease in HRV seemed to be related, at least in part, to hypopnoea that is common on induction of anaesthesia. Actual depression of the gain of the respiratory sinus arrhythmia reflexes appeared to reach a maximum at higher concentrations of anaesthetic (the time of incision), at which point high-frequency power had somewhat recovered. While this result has not been reported previously, it is not surprising when considering the dramatic apnoeas or hypopnoeas that are commonly encountered at induction of general anaesthesia. The importance of measuring depth of respiration when analysing HRV has long been emphasized in the physiological literature.

In a recent article, Gajraj and co-workers compared the ability of BIS, SEF and auditory evoked potentials to distinguish between periods of consciousness and unconsciousness. Consciousness was defined as response to verbal command, and all patients had loss-of-recall during the ‘conscious’ periods in the study. They found that the auditory evoked potential was the best discriminator of loss of consciousness, followed by BIS, with SEF being the least powerful discriminator. The results of our study are similar to those of Gajraj and colleagues and also to those reported by Flailson and colleagues. All three studies consistently demonstrated that there is a small percentage (2–5%) of people who can respond to verbal command with BIS values as low as 55 during recovery from anaesthesia. However, patients are usually amnesic at these levels. The reasons for the ‘failure’ of the BIS to detect awareness are unclear. Awareness at these low BIS values is almost certainly not a result of excessive surgical stimulation. In our study, response to verbal command occurred after surgery had finished, and in Gajraj’s study, patients had adequate spinal anaesthesia. In all of our cases, the raw EEG signal was of good quality. In many of our cases, BIS increased rapidly to more than 95 soon after the response to verbal command. This hysteresis effect highlights a potential problem with the clinical use of BIS. BIS clearly may not reliably warn of impending return of consciousness, but an abrupt increase in BIS usually indicates that some form of change in awareness has just taken place. Because of the strong amnesic effects of general anaesthetics, this may be useful in warding the anesthetist of impending patient recall. However, philosophically it remains second best to the goal of having a monitor that warns the anaesthetist before arousal has taken place.

The time of syringe drop is obviously a transition period. It was surprising that most patients dropped the syringe at relatively high BIS values. There are three possible explanations for this observation. First, this could be a result of a lag in the updating of the display of the EEG machine. However, we do not believe this accounts entirely for the effect because there was no correlation between BIS and the rate of change of BIS. Second, one study has suggested that propofol has a biphasic effect on the EEG. The initial increase in high-frequency activity that occurs as the propofol effect is increasing may artefactually increase inter-frequency coherence. Third, there is loss of neuromotor intention before there is change in electrocortical activity, as detected by BIS. It is also possible that there is artefactual electromyographic interference with the BIS algorithm from the frontalis muscle. We were unable to exclude this as a confounding variable when BIS was used clinically.

Because we wished to see if the indices could be applied clinically, we tested the indices under routine clinical conditions. The effects of using co-induction doses of i.v. midazolam and fentanyl are difficult to quantify precisely but are probably minor. Certainly the starting BIS values were similar to those that have been obtained in a normal alert population, and all patients were able to co-operate fully.

Interpretation of ApEn is not easy. It decreased with increasing depth of anaesthesia. This indicates less disorder...
(entropy)—a more 'regular' EEG signal when anaesthetized. From the results of Pincus, it may be hypothesized that this indicates less integration of the cortical neural subsystems under anaesthesia compared with the awake state. However, it is also possible that ApEn may simply reflect loss of high-frequency desynchronized EEG waves that is a feature of the transition from the conscious to the unconscious state. In essence, ApEn has become a surrogate measure of the increasing predominance of low-frequency components of the EEG signal that accompanies general anaesthesia. ApEn is a family of statistics that indicates relative, rather than absolute, changes. The actual values of ApEn depend on the three variables chosen (n, m and r). We chose the values of the variables based on a priori reasoning. To test if the results were robust to changes in variable values, we tried a range of variables in a small subsample of EEG records. They appeared to give qualitatively similar results.

ApEn clearly does not offer any clinical advantages over BIS as a measure of awareness. It raises some interesting theoretical questions as to the electrocortical events underlying the generation of the EEG and how they are affected by general anaesthesia.

Although each index theoretically reflected a different dimension of the depth of anaesthesia, the use of a combination of multiple indices of level of anaesthesia did not provide any predictive advantages.

Appendix

The ‘10SecHRrms’ and ‘20SecHR rms’ algorithms

These algorithms calculate the variance of the heart rate within consecutive 10-s or 20-s epochs of heart rate data. They are similar to established time domain indices of HRV such as SDNN, except that we have used much shorter time intervals specifically to eliminate the influence of low-frequency changes in heart rate and to concentrate on the high-frequency, respiratory-related components. They may be viewed as a primitive high-pass filter. Their advantages are their simplicity and ease of implementation. They are determined as follows.

1. Acquisition and editing of respiration and RR data for artefacts. Ectopic beats were removed and replaced with a linear interpolation of the pre- and post-ectopic points in the heart rate signal.
2. Conversion of the RR interval data which has been sampled at irregular intervals to a regularly sampled heart rate signal using linear interpolation, as described by Berger and colleagues.23
3. Division of the data into 10-s or 20-s epochs.
4. For each epoch: (i) linear de-trending heart rate signal; (ii) calculate the root mean sum of the squares of the residual data points. This results in an index of HRV that we have termed the ‘10SecHRrms’ or ‘20SecHR rms’; (iii) calculate the ‘gain’ by dividing the ‘10SecHRrms’ by the moving average of minute ventilation. This is a moving average of tidal volume multiplied by respiratory rate, centred on the mid-point of the particular epoch under consideration, and having a width of 20 s.

Acknowledgement

The study was supported with a summer studentship grant from Health Waikato.

References

22. Pincus SM. Greater signal regularity may indicate increased system isolation. Math Biosci 1994; 122: 161–81