

Estimation of the Bispectral Index by Anesthesiologists

An Inverse Turing Test

Michael M. Bottros, M.D.,* Ben Julian A. Palanca, M.D., Ph.D.,† George A. Mashour, M.D., Ph.D.,‡
Ami Patel, B.A.,§ Catherine Butler, B.A.,§ Amanda Taylor, B.A.,|| Nan Lin, Ph.D.,#
Michael S. Avidan, M.B.B.Ch., F.C.A.S.A.**

ABSTRACT

Background: Processed electroencephalographic indices, such as the bispectral index (BIS), are potential adjuncts for assessing anesthetic depth. While BIS® monitors might aid anesthetic management, unprocessed or nonproprietary electroencephalographic data may be a rich source of information for clinicians. We hypothesized that anesthesiologists, after training in electroencephalography interpretation, could estimate the index of a reference BIS as accurately as a second BIS® monitor (twin BIS®) (Covidien Medical, Boulder, CO) when provided with clinical and electroencephalographic data.

Methods: Two sets of electrodes connected to two separate BIS® monitors were placed on the foreheads of 10 surgical patients undergoing general anesthesia. Electroencephalographic parameters, vital signs, and end-tidal anesthetic gas concentrations were recorded at prespecified time points, and were provided to two sets of anesthesiologists. Ten anesthesiologists received brief structured training in electroencephalography interpretation and 10 were untrained. Al-

What We Already Know about This Topic

- Calculated electroencephalographic indices in combination with the raw electroencephalographic waveforms may be helpful for assessment of depth of sedation/hypnosis during general anesthesia.

What This Article Tells Us That Is New

- After brief structured education, with access to data from a frontal electroencephalographic recording and coupled with relevant clinical data, anesthesiologists can estimate processed bispectral index fairly accurately.

though electroencephalographic waveforms and open-source processed electroencephalograph metrics were provided from the reference BIS®, both groups were blinded to BIS values and were asked to estimate BIS.

Results: The trained anesthesiologists averaged as close to or closer to the reference BIS® compared with the twin BIS® monitor for 34% of their BIS estimates *versus* 26% for the untrained anesthesiologists. Using linear mixed effects model analysis, there was a statistically significant difference between the trained and untrained anesthesiologists ($P = 0.02$), but no difference between the twin BIS® monitor and trained anesthesiologists ($P = 0.9$).

Conclusion: With limited electroencephalography training and access to clinical data, anesthesiologists can estimate the BIS almost as well as a second BIS® monitor. These results reinforce the potential utility of training anesthesia practitioners in unprocessed electroencephalogram interpretation.

THE goals of general anesthesia include ensuring that patients are physiologically stable and oblivious to the noxious stimuli of surgical intervention. Unfortunately, because we remain unable to detect unconsciousness reliably, we cannot guarantee that patients will not be awake and aware during surgery. However, important advances have been made in our ability to ensure that adequate anesthesia is administered, including routine measurement of exhaled anesthetic concentrations. Monitoring brain activity directly, first suggested in the 1930s, has garnered tremendous interest in the past decade. A variety of candidate depth-of-anesthesia monitors are now available for clinical use. Most of these monitors are based on brain electrical activity or electroencephalography.

* Resident, † Assistant Professor, ** Associate Professor, Department of Anesthesiology, Washington University School of Medicine, St. Louis, Missouri. ‡ Assistant Professor, Department of Anesthesiology, University of Michigan, Ann Arbor, Michigan. § Medical Student, Washington University School of Medicine. || Medical Student, University of Arizona School of Medicine, Tucson, Arizona. # Assistant Professor, Department of Mathematics and Division of Biostatistics, Washington University, St. Louis, Missouri.

Received from the Department of Anesthesiology, Washington University School of Medicine, St. Louis, Missouri. Submitted for publication May 22, 2010. Accepted for publication December 7, 2010. Supported by a grant (CFM-08/15/2007; to M.S.A.) from the Foundation for Anesthesia Education and Research (Rochester, Minnesota) and the American Society of Anesthesiologists (Park Ridge, Illinois). Presented in part at the annual meeting of the American Society of Anesthesiologists, Medical Student Foundation for Anesthesia Education and Research (FAER) session, New Orleans, Louisiana, October 21, 2009, and the Midwest Anesthesia Resident Conference, Cincinnati, Ohio, February 20, 2010.

Address correspondence to Dr. Avidan: Cardiothoracic Anesthesiology and Cardiothoracic Intensive Care, Washington University School of Medicine, Department of Anesthesiology, 660 South Euclid Avenue, Campus Box 8054, St. Louis, Missouri 63110. avidanm@wustl.edu. Information on purchasing reprints may be found at www.anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

Copyright © 2011, the American Society of Anesthesiologists, Inc. Lippincott Williams & Wilkins. Anesthesiology 2011; 114:1093-1101

Despite the fact that the brain is the target organ of general anesthesia, we do not have standard intraoperative brain monitors, as we do for other vital organs. If a patient receives insufficient anesthesia, he or she may experience unintended awareness during surgery, a complication that can have serious consequences, including posttraumatic stress disorder.¹ In an attempt to avoid patient awareness, anesthesia practitioners may give high doses of anesthesia, which can lead to increased drug costs, increased time spent in the operating room, increased time spent in the recovery area, and possibly even increased adverse effects, such as nausea and vomiting.

Some practitioners believe that using a brain monitor during general anesthesia may help to optimize anesthetic administration. One particular processed electroencephalography monitor, the bispectral index monitor (BIS® monitor; Covidien Medical, Boulder, CO), has gained widespread acceptance. Part of the motivation for developing a processed electroencephalographic monitor is the assumption that anesthesia practitioners are unable to interpret unprocessed electroencephalographic patterns in real time.² However, it has been shown that after a brief, structured education session, anesthesia practitioners can reliably recognize patterns in the electroencephalogram trace that indicate states such as wakefulness, light anesthesia, deep anesthesia, and brain quiescence.² Taken in a clinical context, these electroencephalogram patterns may help practitioners to appreciate whether patients are likely to be unconscious and appropriately anesthetized. There are several potential advantages of high resolution electroencephalographic assessment over the processed monitors, which include rapid response time, decreased costs, ability to appreciate artifacts, lack of reliance on proprietary algorithms, and the ability to distinguish among specific brain states (*e.g.*, seizures, isoelectricity).^{3–7}

We hypothesized that anesthesiologists, appropriately trained in electroencephalography interpretation, could estimate the index of a reference BIS® as accurately as a second BIS® monitor (twin BIS®) when provided with relevant clinical and electroencephalographic data. To test this hypothesis, we designed an inverse Turing test to compare anesthesiologists to the twin BIS® monitor in ability to predict the BIS values of human patients undergoing general anesthesia.

Materials and Methods

This was a prespecified substudy of the BAG-RECALL clinical trial (NCT00281489)⁸ at Washington University (St. Louis, Missouri). The trial aimed to assess practitioner ability to interpret unprocessed electroencephalogram trace during anesthesia. In brief, the BAG RECALL trial is a 6,000-patient multicenter trial that was designed to compare a BIS-based protocol with a protocol based on end-tidal anesthetic concentration monitoring for efficacy in preventing unintended intraoperative awareness with subsequent explicit recall in high-risk patients having surgery under general anesthesia. Patients provided written informed consent to participate in the BAG RECALL trial according to institutional standards.

The Human Research Protection Office at Washington University approved this substudy. Data from 10 patients who participated in the BAG RECALL clinical trial were included in this substudy. Patients were selected sequentially for this substudy during July 2009 after they consented to participate in the BAG RECALL trial and it was revealed that they had been randomized to the BIS protocol of the trial. This substudy did not impact clinical care and did not interfere with the conduct of the BAG RECALL trial because the blind of the study was not broken and no postoperative outcomes data were assessed. The substudy was not conducted in real time and patient data were deidentified.

Each of the 10 patients had two sets of electrodes connected to independent BIS® monitors placed on both sides of the forehead while undergoing general anesthesia. Identical settings were used for both monitors, including the smoothing rate, which was set at 15 s. Anesthesia practitioners responsible for patient care had access to information provided by only one of the BIS® monitors (reference BIS®). Three-second electroencephalography epochs were presented to anesthesiologists. The duration of the electroencephalography epoch required for calculation of the BIS has not been disclosed. It has been assumed to be between 20 and 40 s, however it may be longer than this because a 63-s epoch is used to calculate the burst-suppression ratio. An epoch of a minute is consistent with research showing time delays of the BIS in responding to a state change (*e.g.*, unresponsive to responsive).³ Electroencephalography epochs that were shown to anesthesiologists would, therefore, have hypothetically represented the last 5–15% of the electroencephalography epoch from which the BIS algorithm was calculated. Practitioners caring for patients were blinded to all data on the other BIS® monitor (twin BIS®) and were not involved in the current study. BIS numbers, electroencephalogram traces, electroencephalogram parameters from both monitors, vital signs, and anesthetic gas measurements were recorded at prespecified time points during the anesthetic period. These time points included: preinduction wakefulness, induction, intubation, incision, maintenance, closure, extubation, and postanesthetic wakefulness. In total, there were 90 time points.

We devised the inverse Turing test to test the hypothesis that an anesthesiologist can obtain similar information using a single frontal electroencephalographic channel and clinical context to that he or she can obtain using a BIS® monitor. The Turing test was proposed by Alan Turing, Ph.D., O.B.E., F.R.S. (1912–1954), as a philosophical experiment to determine whether a machine displayed evidence of real intelligence or even consciousness. Conceptually, a human would be placed in a room and would converse (*e.g.*, keyboard and screen interface) with a machine (*e.g.*, robot, computer) in another room and a second human in yet another room. If, after questioning, the human could not distinguish between machine and human, the machine would have passed the Turing test and would have displayed intelligence.

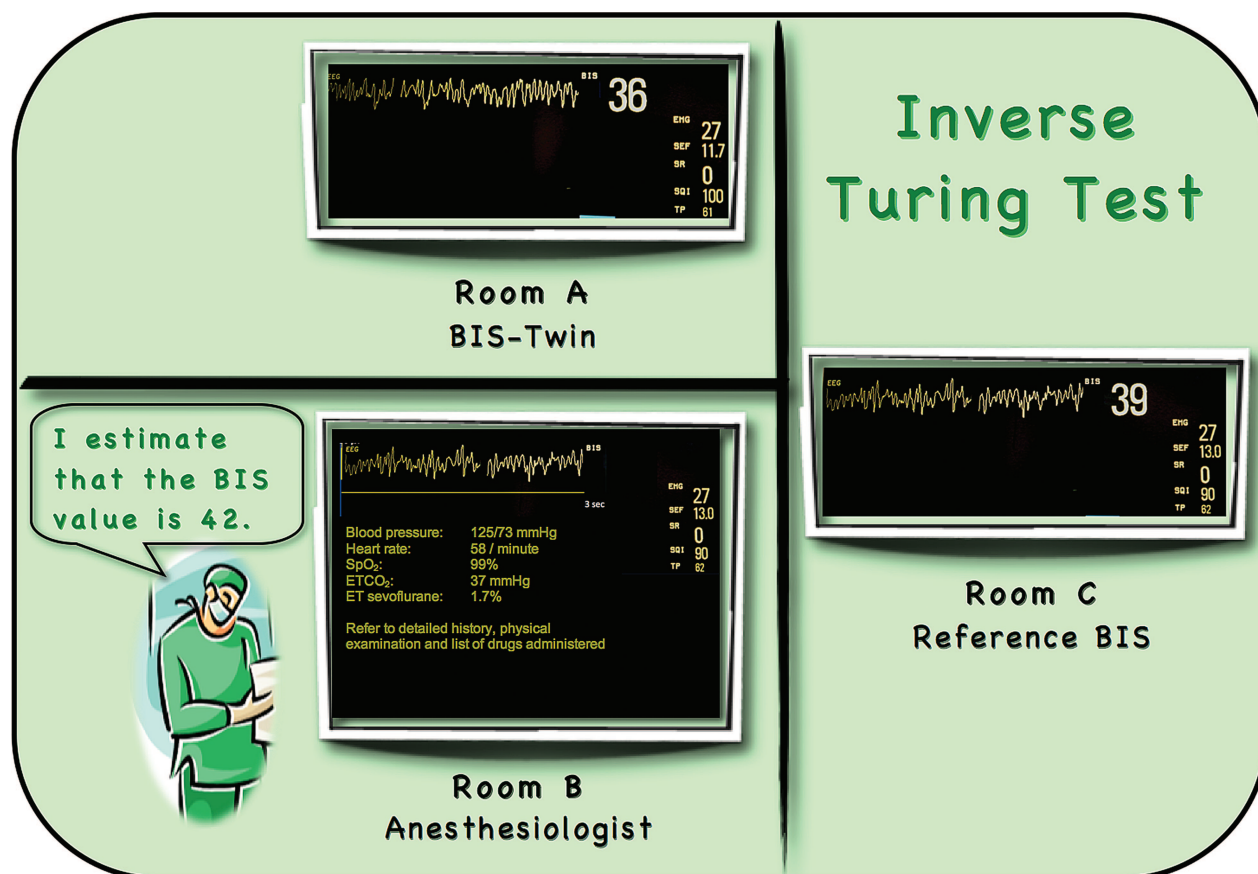


Fig. 1. Inverse Turing test. The reference BIS® (bispectral index) monitor is in metaphorical room C; the twin BIS® monitor is in metaphorical room A; the anesthesiologist is in metaphorical room B. If the reference BIS® monitor cannot “distinguish” between the anesthesiologist and the twin BIS® (i.e., if the anesthesiologist’s BIS estimate is as close to or closer to the reference BIS® than the value displayed by the twin BIS®), the anesthesiologist has passed the conceptual inverse Turing test.

For the purposes of this substudy, we devised an inverse Turing test, which was inspired by the experiment proposed by Turing. The concept behind the inverse Turing test is as follows. A reference BIS® monitor would be in one metaphorical room. The twin BIS® would be in a second metaphorical room. An anesthesiologist would be in yet another metaphorical room and would attempt to estimate the BIS number based on clinical context, electroencephalographic traces, and nonproprietary electroencephalographic parameters (fig. 1). If the anesthesiologist could estimate the reference BIS® value as well as the twin BIS® monitor, the anesthesiologist could be said to have passed the inverse Turing test. An important assumption on which this study depended was that there would be good, but not perfect, agreement between two concurrent BIS® monitors. Based on a previous study where we showed a degree of inpatient variability between two concurrent BIS® monitor readings, we felt that this assumption was reasonable.⁹ An important caveat is that there may be interhemispheric differences in BIS readings attributable to factors such as unilateral brain pathology or decreased regional cerebral blood flow.^{10,11} To avoid such recognized potential confounders, patients with known brain pathology, seizure disorders, or cerebrovascular disease

were not enrolled in this substudy (and were generally excluded from the BAG RECALL clinical trial).

Two groups of attending anesthesiologists (10 anesthesiologists per group) enrolled sequentially and consented to participate in this substudy. In one group, the 10 anesthesiologists received structured training on electroencephalographic interpretation. Training was based on that recommended by Barnard et al² and others and consisted of a 45-min, live presentation that included example electroencephalographic waveforms associated with different depths of anesthesia. We included characteristic changes that occur in the electroencephalogram with increasing concentrations of primarily γ -aminobutyric acid agonist anesthetics, including (1) a decrease in high frequency, low amplitude waves, (2) a concomitant increase in low frequency, high amplitude waves, (3) burst suppression, and (4) isoelectricity (fig. 2).^{4,12} We also instructed anesthesiologists regarding the basic electrophysiologic parameters analyzed by BIS.^{13,14} Nonproprietary parameters displayed by the BIS® monitor were described, including electromyography (power in decibels), spectral edge frequency₉₅ (frequency from 0.5–30 Hz below which 95% of the total electroencephalography power lies), signal quality index (signal quality of the electroencephalog-






Patient State	Device	Features	Reading	Frontal Electroencephalography (EEG) Trace
Wakefulness	EEG	$\uparrow f$, \downarrow Amp, blinks	$\uparrow \gamma, \beta, \alpha \downarrow \theta, \delta$	
	SEF ₉₅	Twenties	26 Hz.	
	BIS	High β power	96	
	ETAC	Age-adjusted MAC	0 MAC	
Sedation	EEG	α oscillations	$\downarrow \gamma, \beta, \uparrow \alpha, \theta, \delta$	
	SEF ₉₅	High teens	19 Hz.	
	BIS	Low β power	78	
	ETAC	Age-adjusted MAC	0.4 MAC	
Unresponsive	EEG	Spindles, K, $\downarrow f$	$\uparrow \alpha, \theta, \delta$	
	SEF ₉₅	Low teens	14 Hz.	
	BIS	Bispectral coherence	52	
	ETAC	Age-adjusted MAC	0.8 MAC	
Surgical anesthesia	EEG	Slow δ waves, $\downarrow f$	δ dominance	
	SEF ₉₅	< 12 Hz.	10 Hz.	
	BIS	Bispectral coherence	42	
	ETAC	Age-adjusted MAC	1.3 MAC	
Deep anesthesia	EEG	BS, isoelectricity	Bursts & flat	
	SEF ₉₅	< 2 Hz. (BS corrected)	1 Hz.	
	BIS	High BSR	9	
	ETAC	Age-adjusted MAC	2 MAC	

Fig. 2. Principles of electroencephalography training for anesthesiologists. The readings shown represent examples of possible readings that may be seen in conjunction with each frontal electroencephalography trace. The electroencephalography traces show 3-s epochs (x-axis) and the scale (y-axis) is 50 μ V. Electroencephalography (EEG) waves in decreasing frequency are as follows: γ , >30 Hz; β , 12–30 Hz; α , 8–12 Hz; θ , 4–8 Hz; δ , 0–4 Hz. The section outlined in the first trace is a blink artifact and the section outlined in the third trace is a sleep spindle. Amp = amplitude; BIS = bispectral index; BS = burst suppression; BSR = burst-suppression ratio; ETAC = end-tidal anesthetic concentration; f = frequency; MAC = minimum alveolar concentration; SEF₉₅ = spectral edge frequency (below which 95% of the EEG frequencies reside).

raphy signal between 0–100%), suppression ratio (percentage over a 63-s epoch that the electroencephalography is in a suppressed state), and total power (total power of the summated electroencephalography component waves in decibels).^{13,14} Based on the work of Morimoto *et al.*,¹⁴ we explained that, at surgical levels of anesthesia, spectral edge frequency and BIS are well correlated. We further suggested that spectral edge frequency could often be a useful surrogate for BIS (fig. 2). At lighter levels of anesthesia, higher β ratio (with increased high-frequency β waves) correlates with BIS (BIS 60–100); at very deep anesthesia, the extent of burst suppression (suppression ratio) correlates with BIS (BIS 0–30).^{13,14} During wakefulness in patients who are not pharmacologically paralyzed, electromyographic values are typically in the 40s and 50s, whereas, during surgical anesthesia, they are typically in the 20s and 30s. Measured artifacts (*e.g.*, blinking and eye movement) were discussed as well as shape and variability of dose-response curves for processed electroencephalographic indices. None of the 20 anesthesiologists had previously received formal training in electroencephalographic interpretation and none of them had routinely used unprocessed electroencephalography as part of their standard practice.

All 20 anesthesiologists were presented with relevant patient histories, clinical information, anesthetic data, hemodynamic parameters, drug doses, anesthetic gas concentrations, and electroencephalographic traces and associated

nonproprietary parameters (spectral edge frequency, electromyography, suppression ratio, signal quality index, and total power). Based on this information, practitioners were asked to estimate what they thought the BIS value was for each

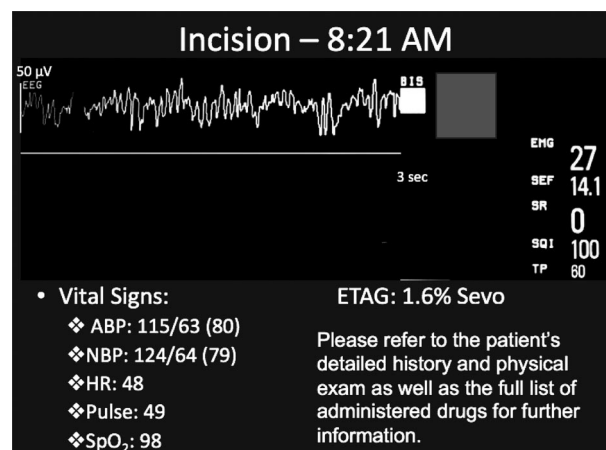


Fig. 3. Clinical scenario screen. An example screen presented to each anesthesiologist at prespecified time points used in estimating bispectral index. ABP = arterial blood pressure; EEG = electroencephalography; EMG = electromyography; ETAG = end-tidal anesthetic gas concentration; HR = heart rate; NBP = noninvasive blood pressure; SEF = spectral edge frequency₉₅; Sevo = sevoflurane; SpO₂ = peripheral oxygen saturation; SQI = signal quality index; SR = suppression ratio; TP = total power.

patient at each of the various prespecified time-points. Figure 3 represents an example screen presented to anesthesiologists when estimating BIS values.

Statistics

Descriptive statistics are presented as text, tables, and graphs. The primary hypothesis of this substudy was that structured electroencephalography training would result in improved estimation of the processed electroencephalograph index. We performed a power analysis using a simulation approach based on a mixed effect model. From pilot data, we assumed that trained anesthesiologists' estimates would be, on average, 5 absolute BIS units closer to the that of the reference BIS® monitor than the untrained anesthesiologists' estimates (*i.e.*, treatment effect size = 5). Other factors considered in power analysis were BIS measurements on 10 patients, patient random effect on the intercept having a variance of 20 (alternatively expressed as SD of the random intercept) and a measurement error variance of 100. These values correspond to a correlation of 0.2 between repeated measures. The result of the power analysis is that, with 10 replicates (that is, for each reference BIS® value, there are 10 repeated estimates from 10 trained and 10 untrained anesthesiologists), the study has 88% power to detect a difference (the specified treatment effect size) between the trained and untrained anesthesiologists at a significance level of 0.05.

We fitted a linear mixed effect model to our results. The following effects were considered:

1. Fixed Effect of the Device. That is, who provided the measurement. This effect has four levels: reference BIS®, twin BIS®, trained anesthesiologist, and untrained anesthesiologist.

2. Random Effect for Experts in the Trained Group and the Untrained Group

3. Random Effect Due to Repeated Measurements for Each Patient. In mathematical terms, the model was conceptualized as follows. Let i be the index for patient; j , the index for device (reference BIS®, twin BIS®, trained anesthesiologist, or untrained anesthesiologist); and y_{ijk} be the k th score given by the j th device for the i th patient. The outcome measure of the model was Y_{ijk} .

Next, let α_i be the patient-specific (random) effect; β_j , the device (fixed) effect; and $\gamma_{(ij)}$, the nested device (random) effect within patient. Our model is then written as $y_{ijk} = \alpha_i + \beta_j + \gamma_{(ij)} + \varepsilon_{ijk}$, where the last term (ε_{ijk}) denotes the random measurement error. By introducing α_i and $\gamma_{(ij)}$ as random effects, we assume that the average response level and the device effect varies randomly across patients.

The linear mixed-effect model was fit using PROC MIXED in SAS (version 9.2; SAS Institute, Inc., Cary, NC). One potential issue in the dataset relates to missing observations. However, when using PROC MIXED, only missing observations were discarded but not the entire dataset. A *post hoc* Tukey test was performed for correction of multiple comparisons. A P

value of less than 0.05 was considered statistically significant. All hypothesis testing was two-tailed. Statistical analyses were performed with SAS or Analyse-it® (Analyse-it Software, Ltd., Leeds, United Kingdom) statistical software.

Results

Characteristics of the patients who were enrolled in this substudy and details pertaining to anesthesia are shown in table 1. There were 90 values for the reference BIS® and the twin BIS®. There were 1,794 values for the 10 trained and 10 untrained anesthesiologists, with six missing BIS estimates. Discordance rates of the twin BIS®, trained anesthesiologists, and untrained anesthesiologists when the reference BIS® was less than 40, 40–60, or more than 60 are shown in table 2. Notably, when the reference BIS® displayed values above 60, about 20% of anesthesiologists' estimates were below 60 and 9% of the twin BIS®'s estimates were below 60.

The effect of training was associated with a significant difference in performance in estimating the value of the reference BIS® (uncorrected $P = 0.004$, corrected $P = 0.019$). Concordance rates in estimating "depth of anesthesia," defined by three prespecified "depths" (BIS >60, BIS 40–60, BIS <60), with the reference BIS® were 86% for the twin BIS®, 72% for the 10 trained anesthesiologists, and 62% for the 10 untrained anesthesiologists. Figure 4 presents box-and-whisker plots showing the median, interquartile ranges, and 90% ranges for the differences among the twin BIS®, trained anesthesiologists, and untrained anesthesiologists compared with the reference BIS®. The median of the twin BIS® estimates was no different from the reference BIS®. The median of trained anesthesiologist BIS estimates deviated from the reference BIS® with a range of 1–4, depending on the observer. Median untrained anesthesiologist BIS estimates deviated from the reference BIS® with a range of 2–11, depending on the observer. On average, the 10 trained anesthesiologists were as close to or closer to the reference BIS® compared with the twin BIS® for 34% of their estimates. On average, the 10 untrained anesthesiologists were as close to or closer to the reference BIS® compared with the twin BIS® for 26% of their estimates. Figure 5 graphically depicts all 90 BIS values for each of the 10 patients as displayed by the reference and twin BIS® monitors. In addition, median estimated BIS values are shown for trained and untrained anesthesiologists.

On aggregate, there was no statistically significant difference between the trained anesthesiologists' and twin BIS®s' performance in predicting the reference BIS® value (uncorrected $P = 0.43$, corrected $P = 0.9$). There was a significant difference in the uncorrected comparison of the untrained anesthesiologists and twin BIS® ($P = 0.04$) that was not significant after Tukey correction ($P = 0.16$).

Discussion

These data demonstrate that basic training is probably associated with improvement in electroencephalographic interpreta-

Table 1. Patient Characteristics

Characteristic	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5
Age, yr	39	53	72	65	25
Sex, M/F	M	M	F	M	F
Diagnosis	Lung nodule	Esophageal cancer	Stomach mass	Left lung lesion	Pulmonary valve stenosis
Procedure	Bronchoscopy and biopsy	Esophagectomy	Gastrectomy and cholecystectomy	Video-assisted thoracoscopic surgery	Pulmonary valve replacement
Functional status	Moderate	Moderate	Marginal	Marginal	Moderate
ASA category	3	3	4	4	3
Height, inches	68	71	63	72	61
Weight, kg	105	101.6	77.2	60.6	41.5
Body mass index	35.2	31.2	30.1	18.1	17.3
Comorbidities	Difficult intubation	Asthma, reflux, sleep apnea, rheumatoid arthritis, daily ethanol	Aortic stenosis, dyslipidemia, hypothyroidism, hiatal hernia, anemia, diabetes, gastric ulcer, depression	Atrial fibrillation, emphysema, steroid use, cardiomyopathy, atherosclerosis, daily ethanol	Hypothyroidism, gastrointestinal bleed, dilated aortic root, impaired right ventricular function
Intubation	Fiber optic	Direct laryngoscopy	Direct laryngoscopy	Direct laryngoscopy	Direct laryngoscopy
Anesthetic drugs	Midazolam 2 mg, fentanyl 150 mcg, propofol 260 mg, phenylephrine 100 mcg, sevoflurane	Midazolam 5 mg, Fentanyl 550 mcg, Propofol 400 mg, Vecuronium 27 mg, Neostigmine 3 mg, Glycopyrrolate 0.6 mg, Sevoflurane	Fentanyl 300 mcg, propofol 150 mg, succinylcholine 100 mg, vecuronium 8 mg, neostigmine 3 mg, glycopyrrolate 0.6 mg, phenylephrine 500 mcg, desflurane	Midazolam 1 mg, fentanyl 250 mcg, propofol 140 mg, rocuronium 70 mg, neostigmine 3 mg, glycopyrrolate 0.4 mg, desflurane	Midazolam 1 mg, fentanyl 275 mcg, methadone 10 mg, propofol 100 mg, vecuronium 4 mg, phenylephrine 350 mcg, desflurane
Duration of anesthesia, min	60	245	210	180	240

ASA = ASA (American Society of Anesthesiologists) Physical Status Classification System.

tion and that a trained anesthesiologist can predict BIS values as well as or better than a second BIS® monitor approximately one third of the time. Furthermore, the finding in this pilot study that there was no significant difference between a trained anesthesiologist and a second BIS® monitor in predicting a reference BIS® index generates the intriguing hypothesis that anesthesiologists are capable of assessing anesthetic depth based on clinical data and basic electroencephalographic parameters as well as a processed electroencephalographic monitor. This hypothesis warrants further investigation. The relatively high discordance rates for both groups of anesthesiologists and the twin BIS® in relation to the reference BIS® for BIS values above 60 is potentially concerning. However, BIS values above 60 have been found to have a poor specificity for return of wakeful responsiveness.¹⁵

There has been a renewed focus on the electroencephalographic assessment of anesthetic effects.^{4,6} It has previously been argued—and is now often assumed—that the use of unprocessed electroencephalography is impractical in the intraoperative setting. The presuppositions of such arguments may no longer be applicable. First, electroencephalographic devices are no longer cumbersome and can be easily accommodated in clinically used modules. Second, digitization of electroencephalographic signals allows for basic processing and more ready interpretation. Finally, our data suggest that trained anesthesiologists can recognize and interpret patterns in the electroencephalogram and can reach clinical conclusions that are generally similar to those they may have reached with the use of a processed algorithm. As with electrocardiography, struc-

Table 2. Rates of Discordance in Depth-of-Anesthesia Determinations between Duplicate BIS® Monitors and Trained *versus* Untrained Anesthesiologists

Reference	BIS® Monitor	Anesthesiologists	
	Twin	Trained	Untrained
<40 (n = 25)	4% of readings 40 – 60	36% of readings 40 – 60	74% of readings 40 – 60; 1% of readings >60
40–60 (n = 31)	23% of readings <40; 7% of readings >60	18% of readings <40; 12% of readings >60	8% of readings <40; 19% of readings >60
>60 (n = 34)	9% of readings ≤60	21% of readings ≤60	20% of readings ≤60

BIS® monitor = bispectral index monitor (Covidien Medical, Boulder, CO).

Patient 6	Patient 7	Patient 8	Patient 9	Patient 10
61	40	52	47	46
M	F	F	F	F
Esophageal cancer	Living kidney donor	Anomalous coronary artery	Ventricular septal defect	Acute respiratory distress
Esophagectomy	Donor nephrectomy	Two-vessel coronary artery bypass surgery	Ventricular septal defect closure	Video-assisted thoracoscopic surgery/ chest tube placement
Marginal	Moderate	Moderate	Moderate	Moderate
4	2	3	3	3
71	65	59	69	72
126	89.6	75.6	68.4	104.3
38.7	32.9	33.7	22.3	31.2
Chronic obstructive lung disease, reflux, sleep apnea, hypertension, difficult intubation, daily ethanol	Hypertension, methamphetamine abuse, regular benzodiazepines	Hypertension, high cholesterol, anxiety, kidney stones, migraine, regular opiates	Reflux, breast cancer, endocarditis, impaired right ventricular function	Chronic obstructive lung disease, hemothorax, rib fracture, regular opiates, difficult intubation
Fiber optic	Direct laryngoscopy	Direct laryngoscopy	Direct laryngoscopy	Fiber optic
Midazolam 3 mg, fentanyl 450 mcg, propofol 180 mg, rocuronium 70 mg, neostigmine 3 mg, glycopyrrolate 0.4 mg, desflurane	Midazolam 5 mg, lidocaine 100 mg, fentanyl 500 mcg, propofol 150 mg, atracurium 30 mg, vecuronium 4 mg, hydromorphone 0.8 mg, ondansetron 4 mg, neostigmine 3 mg, glycopyrrolate 0.4 mg, desflurane	Midazolam 9 mg, fentanyl 1,050 mcg, propofol 200 mg, phenylephrine 400 mcg, sevoflurane	Fentanyl 700 mcg, propofol 150 mg, methadone 20 mg, phenylephrine 500 mcg, sevoflurane	Midazolam 2 mg, fentanyl 1,000 mcg, dexmedetomidine 0.5 mcg/kg/hr, propofol 100 mg, phenylephrine 900 mcg, dexamethasone 8 mg, ondansetron 4 mg, sevoflurane
220	235	270	255	100

tured education and practical experience in electroencephalography would probably instill even greater proficiency with its use and interpretation. We see no fundamental difference in the potential of an anesthesiologist to recognize the slowing frequency of an electroencephalographic waveform *versus* the slowing of a heart rate, a K complex *versus* a premature ventricular contrac-

tion, burst suppression *versus* a run of ventricular tachycardia, or a sleep spindle *versus* torsades de pointes.

It is of interest to note that the Accreditation Council for Graduate Medical Education (Chicago, IL) lists as a requirement for anesthesiology training that “the resident must either personally participate in cases in which electroencephalograph or processed electroencephalographic monitoring is

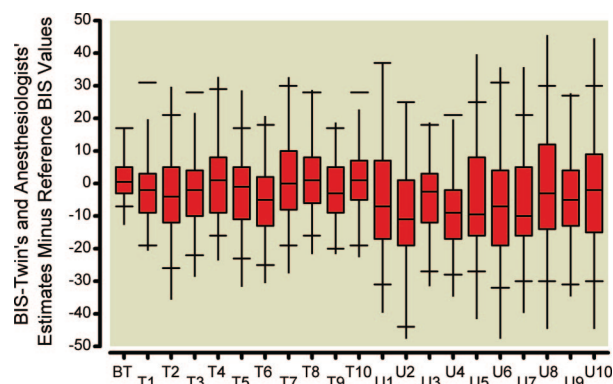


Fig. 4. Differences between twin BIS® (bispectral index) and anesthesiologists compared with reference BIS®. Box-and-whisker plots with the boxes showing the median and interquartile ranges, with whiskers showing 90% of the data for the differences between the twin BIS® (BT), the 10 trained (T1–T10) and 10 untrained (U1–U10) anesthesiologists compared with the reference BIS® monitor. Values of zero indicate no difference compared with the reference BIS®.

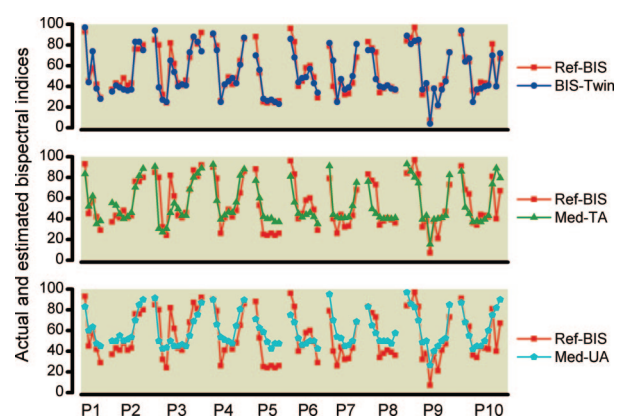


Fig. 5. BIS® (bispectral index) values and estimates. These three graphs depict 90 BIS® values recorded at prespecified time points for each of the 10 patients (P1–P10) as displayed by the reference (Ref-BIS) and twin BIS® (BIS-Twin) monitors. In addition, median estimated BIS values at the same prespecified time points are shown for trained (Med-TA) and untrained (Med-UA) anesthesiologists.

actively used as part of the procedure or have adequate didactic instruction to ensure familiarity with electroencephalographic use and interpretation.”†† The guideline goes on to say that “Bispectral index use and other similar interpolated modalities are not sufficient to satisfy this requirement.”†† As it stands, there are probably few training programs in the United States that implement training in electroencephalography as a formal part of their didactic program, despite its perceived importance. Recent studies have tried to implement formal learning modules in collaboration with neurologists to improve familiarity and success with electroencephalography use and interpretation.^{16,17} Our data suggest that improved skills relevant to clinical practice may also be achieved in the field of anesthesiology.

One of the major challenges in relation to studies dealing with the electrophysiologic monitoring of anesthetic depth is the lack of a “gold standard.” In using a reference BIS® monitor, we were able to judge improvement in performance as objectively as possible. It could be argued that what has been demonstrated with this study is that, with training, anesthesiologists can improve their ability at approximating an index whose value or meaning has not been sufficiently established. However, the BIS® monitor is one of several candidate depth-of-anesthesia monitors that has been approved by the US Food and Drug Administration. In several studies,^{15,18} it has been shown to have a reasonable ability to discriminate between responsiveness and unresponsiveness. No method has been, nor indeed can presently be, validated to assess depth of anesthesia beyond loss of responsiveness. Although the algorithm used to calculate the BIS number is proprietary, the electroencephalographic components of the algorithm have been described.^{13,14} The relative BetaRatio subparameter is the log ratio of power in two empirically derived frequency bands, 30–47 and 11–20 Hz.¹³ The SynchFastSlow subparameter is the contribution from bispectral analysis, and reflects phase coupling between different frequencies. SynchFastSlow is also defined as a log ratio: the log of the ratio of the sum of all bispectrum peaks between 0.5–47 Hz over the sum of the bispectrum in the area 40–47 Hz.¹³ Burst-suppression ratio is also incorporated in the proprietary BIS® algorithm.^{13,14} These features were discussed in the training sessions provided to the trained group of anesthesiologists. The difficulty with establishing the reference BIS® as the gold standard in this study is evidenced by the fact that not all deviations of the anesthesiologist from the reference BIS® were actual errors. In one example, a patient had aroused, opened her eyes and was moving her arms, yet both the BIS and twin BIS® still read 40 (fig. 6). It is noteworthy that, for this time point, all trained anesthesiologists estimated a BIS value above 78, and all untrained

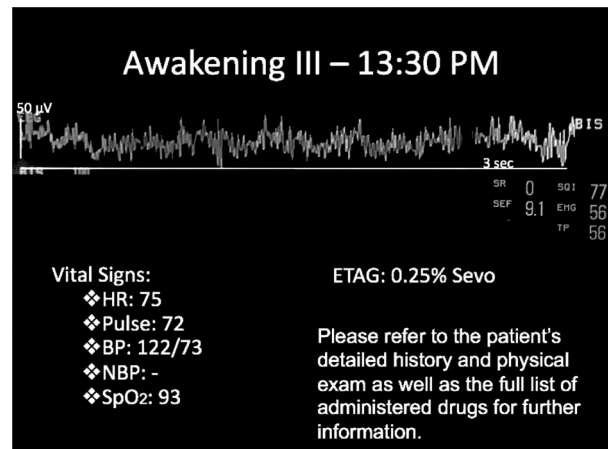


Fig. 6. Bispectral index (BIS) time delay. At this time point, the patient roused suddenly, opened her eyes, and moved her arms. However, both the reference and twin BIS® monitor still read 40. BP = blood pressure; EMG = electromyography; ETAG = end-tidal anesthetic gas concentration; HR = heart rate; NBP = noninvasive blood pressure; SEF = spectral edge frequency₉₅; Sevo = sevoflurane; SpO₂ = peripheral oxygen saturation; SQI = signal quality index; SR = suppression ratio; TP = total power.

anesthesiologists estimated a BIS value above 70. This particular discrepancy between the BIS® monitors and anesthesiologists may have, in part, been attributable to the time delay that occurs with the BIS and other processed electroencephalograph monitors when there is an abrupt change in state, such as an arousal.³ In this case, the clinical context and the unprocessed electroencephalographic trace changed more rapidly than the depth-of-anesthesia index.

This study has several noteworthy limitations that should be emphasized. First, practitioners were not randomly allocated to the electroencephalography training group and no pretesting was done. Therefore, we cannot be certain that the two groups of anesthesiologists were well matched and that differences in performance were attributable to electroencephalography training. However, such differences would potentially be accounted for in the mixed-effect model. Second, the anesthesiologists in this study were not assessing the electroencephalogram trace during clinical administration of anesthesia. We should not infer that the results obtained in the comfort of the classroom are necessarily transferable to the pressurized clinical setting. Third, a major potential advantage of a processed index over an unprocessed electroencephalography trace is that an alert may be linked to a threshold value of an index (*e.g.*, BIS > 60). The results of this study should therefore not be interpreted to suggest that trained anesthesiologists are equivalent or similar in terms of efficacy to a processed electroencephalography index. Fourth, all of practitioners had previously used BIS® monitors. We cannot exclude that previous experience in the use of processed electroencephalography might have helped practitioners estimate BIS on the basis of clinical data and the unprocessed electroencephalography, which is generally displayed on the

†† Armbruster J: Program requirements for graduate medical education in anesthesiology. Residency Review Committee for Anesthesiology, August 4, 2005. Available at: http://www.acgme.org/acWebsite/reviewComment/rev_040pr08_04_05.asp. Accessed February 13, 2011.

BIS® monitor and may have led to “informal training” in the past. Finally, parameters such as blood pressure, heart rate, and end-tidal anesthetic concentration could have influenced practitioner estimates of BIS.

In conclusion, we demonstrate that relatively brief training in electroencephalography improves the performance of anesthesiologists in assessing anesthetic depth, with reference to a target BIS value. These data suggest that anesthesiology practitioners would probably benefit from standardized, formal education in electroencephalography interpretation. Finally, if validated by future studies, our data also suggest that nonproprietary electroencephalographic measures in conjunction with clinical context may be as informative as commercially available depth-of-anesthesia monitors.

The authors dedicated this study to Alan Turing Ph.D., O.B.E., cryptographer, mathematician, logician, philosopher, and founder of computer science. He contributed much to science and to humanity. The authors thank the members of the BAG RECALL research team and the anesthesiologists who agreed to participate in this substudy.

References

1. Leslie K, Chan MT, Myles PS, Forbes A, McCulloch TJ: Post-traumatic stress disorder in aware patients from the B-awake trial. *Anesth Analg* 2010; 110:823–8
2. Barnard JP, Bennett C, Voss LJ, Sleight JW: Can anaesthetists be taught to interpret the effects of general anaesthesia on the electroencephalogram? Comparison of performance with the BIS and spectral entropy. *Br J Anaesth* 2007; 99: 532–7
3. Zanner R, Pilge S, Kochs EF, Kreuzer M, Schneider G: Time delay of electroencephalogram index calculation: Analysis of cerebral state, bispectral, and Narcotrend indices using perioperatively recorded electroencephalographic signals. *Br J Anaesth* 2009; 103:394–9
4. Voss L, Sleight J: Monitoring consciousness: The current status of EEG-based depth of anaesthesia monitors. *Best Pract Res Clin Anaesthesiol* 2007; 21:313–25
5. Dahaba AA: Different conditions that could result in the bispectral index indicating an incorrect hypnotic state. *Anesth Analg* 2005; 101:765–73
6. Bennett C, Voss LJ, Barnard JP, Sleight JW: Practical use of the raw electroencephalogram waveform during general anesthesia: The art and science. *Anesth Analg* 2009; 109:539–50
7. Palanca BJ, Mashour GA, Avidan MS: Processed electroencephalogram in depth of anesthesia monitoring. *Curr Opin Anaesthesiol* 2009; 22:553–9
8. Avidan MS, Palanca BJ, Glick D, Jacobsohn E, Villafranca A, O'Connor M, Mashour GA, BAG-RECALL Study Group: Protocol for the BAG-RECALL clinical trial: A prospective, multi-center, randomized, controlled trial to determine whether a bispectral index-guided protocol is superior to an anesthesia gas-guided protocol in reducing intraoperative awareness with explicit recall in high risk surgical patients. *BMC Anesthesiol* 2009; 9:8
9. Niedhart DJ, Kaiser HA, Jacobsohn E, Hantler CB, Evers AS, Avidan MS: Inpatient reproducibility of the BISxp monitor. *ANESTHESIOLOGY* 2006; 104:242–8
10. Kodaka M, Nishikawa Y, Suzuki T, Asano K, Maeyama A, Miyao H: Does bilateral bispectral index monitoring (BIS) detect the discrepancy of cerebral reperfusion during carotid endarterectomy? *J Clin Anesth* 2009; 21:431–4
11. Estruch-Pérez MJ, Ausina-Aguilar A, Barberá-Alacreu M, Sánchez-Morillo J, Solaz-Roldán C, Morales-Suárez-Varela MM: Bispectral index changes in carotid surgery. *Ann Vasc Surg* 2010; 24:393–9
12. Kortelainen J, Koskinen M, Mustola S, Seppänen T: EEG frequency progression during induction of anesthesia: From start of infusion to onset of burst suppression pattern. *Conf Proc IEEE Eng Med Biol Soc* 2007; 2007:1570–3
13. Rampil IJ: A primer for EEG signal processing in anesthesia. *ANESTHESIOLOGY* 1998; 89:980–1002
14. Morimoto Y, Hagiwara S, Koizumi Y, Ishida K, Matsumoto M, Sakabe T: The relationship between bispectral index and electroencephalographic parameters during isoflurane anesthesia. *Anesth Analg* 2004; 98:1336–40
15. Schneider G, Gelb AW, Schmeller B, Tschakert R, Kochs E: Detection of awareness in surgical patients with EEG-based indices—bispectral index and patient state index. *Br J Anaesth* 2003; 91:329–35
16. Fahy BG, Chau DF, Bensalem-Owen M: Evaluating the requirements of electroencephalograph instruction for anesthesiology residents. *Anesth Analg* 2009; 109:535–8
17. Fahy BG, Chau DF, Owen MB: The effectiveness of a simple novel approach on electroencephalograph instruction for anesthesiology residents. *Anesth Analg* 2008; 106:210–4
18. Schneider G, Wagner K, Reeker W, Hänel F, Werner C, Kochs E: Bispectral Index (BIS) may not predict awareness reaction to intubation in surgical patients. *J Neurosurg Anesthesiol* 2002; 14:7–11