Bispectral Index Monitoring during Sedation with Sevoflurane, Midazolam, and Propofol

Andra E. Ibrahim, M.D.,* Julie K. Taraday, B.A.,† Evan D. Kharasch, M.D., Ph.D.‡

Background: Bispectral Index (BIS) has been used to measure sedation depth. Ideally, to guide anesthetic management, range of BIS scores at different sedation levels should not overlap, and BIS should be independent of drug used. This study assessed ability of BIS to predict sedation depth between sevoflurane, propofol, and midazolam. Quality of recovery was also compared.

Methods: Patients undergoing surgery with local or regional anesthesia and sedation were randomized to sevoflurane (n = 23), midazolam (n = 21), or propofol (n = 22). Sedation was titrated to Observer's Assessment of Alertness–Sedation score of 3 (responds slowly to voice). BIS and Observer's Assessment of Alertness–Sedation were measured every 5 min. BIS prediction probability (Pp) was compared between drugs. Recovery was assessed by BIS and Digit Symbol Substitution and memory tests.

Results: Bispectral Index of responders to voice was significantly different from nonresponders (86 ± 10 vs. 74 ± 14, mean ± SD; P < 0.001) However, wide variability and overlap in BIS were observed (25th–75th percentile, responders vs. nonresponders: 79–96 vs. 65–83). BIS of responders was different for sevoflurane versus propofol and midazolam. BIS was a better predictor of propofol sedation than sevoflurane or midazolam (Pp = 0.87 ± 0.11, 0.76 ± 0.01, and 0.69 ± 0.02, respectively; P < 0.05). At 10 min after the procedure, 76, 48, and 24% of sevoflurane, propofol, and midazolam patients, respectively, returned to baseline Digit Symbol Substitution scores (P < 0.05). Excitement–disinhibition occurred in 70, 36, and 5% of sevoflurane, propofol, and midazolam patients, respectively (P < 0.05).

Conclusion: Individual BIS scores demonstrate significant variability, making it difficult to predict sedation depth. The relation between BIS and sedation depth may not be independent of anesthetic agent. Quality of recovery was similar between drugs, but excitement occurred frequently with sevoflurane.

THE Bispectral Index (BIS) has been used as a measure of the hypnotic effects of anesthetic agents. Although more commonly used as a measure of hypnosis during general anesthesia, BIS has also been used during sedation. BIS has been used to evaluate depth of sedation with propofol,1–2 midazolam,3 and sevoflurane.3,5 Specifically, BIS has been evaluated for its ability to predict response to command,5,6 memory impairment,7 learning during anesthesia,8 and movement to skin incision.9–11 These reports have attempted to correlate BIS values to clinical measurements to predict depth of sedation–anesthesia. Drummond12 suggested that if a monitor of sedation depth is intended for use as a reliable and accurate guide of anesthetic management, two conditions need to be fulfilled: (1) not only must the average values measured by the device in two different depths of anesthesia be statistically different, but also no overlap should exist between the range of values seen in those two states; and (2) the choice of anesthetic agent should have no effect on measured values that differentiate various depths of anesthesia. Recent evidence has questioned the ability of the BIS® monitor (Aspect Medical Systems, Nantick, MA) to fulfill these conditions. Large variability and overlap in BIS scores at distinct depths of anesthesia have been observed, which would make differentiation of these anesthetic depths difficult.13,14 Furthermore, several reports have suggested that BIS measurements may not be independent of drug used. For example, BIS values during propofol at loss of response to voice command were different when fentanyl, nitrous oxide, or alfentanil was added to the anesthetic.6,14,15 When ketamine was used for sedation or added to a propofol infusion, accuracy was significantly affected.16,17 Thus, BIS needs further evaluation to determine if it can be used to guide anesthetic management. The first purpose of this study was to evaluate BIS measurements at various depths of sedation and to compare the ability of BIS to predict depth of sedation between three sedation regimens: sevoflurane, propofol, and midazolam.

Sedation with sevoflurane is a relatively new technique. Sevoflurane sedation has been reported to result in more rapid recovery of cognitive function than midazolam, but with a higher incidence of excitement–disinhibition.18 Propofol and midazolam are more commonly used for sedation; however, the speed of recovery between propofol, midazolam, and sevoflurane has not been compared. Therefore, the second purpose of this study was to compare the quality of sedation and speed of recovery between sevoflurane, propofol, and midazolam sedation.

Methods

Participants

Sixty-six patients (American Society of Anesthesiologists status I–III; age, 19–76 yr) were enrolled in an open-label, randomized investigation approved by the
institutional review board at the University of Washington (Seattle, WA) and provided written informed consent. A subset of 35 patients (23 sevoflurane, 12 midazolam) was the University of Washington cohort of a separate multicenter investigation assessing the quality of recovery with sevoflurane compared with midazolam. Because part of this investigation was a subset of a multicenter investigation comparing midazolam and sevoflurane, patients underwent a dual, nested randomization first to propofol versus midazolam or sevoflurane, then to midazolam or sevoflurane. Patients were scheduled for elective surgery of anticipated duration 0.5–2 h and requested either local or regional anesthesia with sedation. Individuals were excluded if they were pregnant, taking opioids or sedatives within 24 h before their enrollment, or at risk for aspiration. All patients fasted for a minimum of 6 h before surgery. No preoperative opioids or prophylactic antiemetics were given. Remifentanil infusion was permitted for analgesia during regional block placement. No other preoperative sedation was allowed.

**Anesthetic Technique**

All patients received an air–oxygen mixture (fraction of inspired oxygen, 0.4 ± 0.1) at 2 l/min via circle system (facemask connected to a semiclosed anesthetic circuit) and were randomized after enrollment to either inhaled sevoflurane (n = 23), intravenous midazolam (n = 21), or intravenous propofol (n = 22). All patients were monitored with an electrocardiograph, noninvasive blood pressure, pulse oximeter, capnograph, oxygen analyzer, and BIS® monitor. A facemask was applied before administration of study drug and was held manually or fixed with a rubber headstrap to achieve a tight seal. Blood pressure, heart rate, oxygen saturation, end-expired carbon dioxide, end-expired sevoflurane concentration, and BIS score were assessed every 1–2 min for the first 10 min of study drug administration or until maintenance was reached, whichever was later, then every 5 min. An independent observer who was not blinded to the study drug assessed sedation level every minute using the Observer’s Assessment of Alertness–Sedation (OAAS) scale until maintenance, then every 5 min. BIS scores were recorded before the OAAS measurement to assure that verbal or tactile stimulation used to assess OAAS level did not affect the BIS score. Sedation was titrated to an OAAS score of 3 (table 1). Maintenance level was defined as three consecutive OAAS scores of 3; subsequent assessments were made every 5 min. Deviation from maintenance level did occur, which was caused by changing surgical stimulation or difficulty titrating or maintaining sedation at the target level. Sevoflurane was introduced slowly to achieve an OAAS score of 3. Inadequate (or excessive) sedation was treated by increasing (or decreasing) the concentration by 0.2–0.6% until the desired effect was reached. Midazolam was titrated slowly to the desired effect. Within a 2-min period, no more than 2.5 mg was administered to patients younger than 60 yr, and no more than 1.5 mg in patients aged 60 yr or older. Inadequate sedation was treated by further administration of drug given by slow titration in increments judged by the investigator to reach the desired effect. Excessive sedation was treated by holding midazolam doses until the patient returned to an OAAS score of 3. Every midazolam dose administered was recorded. Propofol was administered by infusion, ranging from 25 to 175 μg · kg⁻¹ · min⁻¹, supplemented with bolus doses ranging from 10 to 60 mg, to achieve the desired level of sedation. The level of sedation was targeted to an OAAS score of 3 throughout the procedure until the last suture or procedure equivalent.

Local anesthesia and regional anesthesia were administered to 23 and 43 patients, respectively. Details of anesthesia technique and surgical procedures are shown in table 2. Block placement occurred before the administration of study drug. If analgesia during block placement was necessary, a remifentanil infusion (n = 5) was used and discontinued after block placement was complete. Initiation of sedation was not begun until the patient had returned to an OAAS score of 5. Intraoperative fentanyl administration was allowed in the event of pain, as a treatment for excitation and excessive patient movement, or prophylaxis of painful surgical stimulation.

The incidence of side effects (e.g., apnea, airway obstruction, excitation) was recorded by the independent observer. Severe excitation–disinhibition was defined as agitation and uncontrollable patient movement that, in the investigator’s opinion, resulted in difficult or unsafe operating conditions, and conversion to general anesthesia was considered necessary. Moderate excitation–disinhibition was defined as agitation and uncontrollable movement that was not believed to compromise surgical

---

**Table 1. Observer’s Assessment of Alertness/Sedation (OAAS)**

<table>
<thead>
<tr>
<th>Responsiveness</th>
<th>Speech</th>
<th>Facial Expression</th>
<th>Eyes</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responds readily to name</td>
<td>Normal</td>
<td>Normal</td>
<td>Clear, no ptosis</td>
<td>5</td>
</tr>
<tr>
<td>Lethargic response to name</td>
<td>Mild slowing</td>
<td>Mild relaxation</td>
<td>Glazed/mild ptosis</td>
<td>4</td>
</tr>
<tr>
<td>Responds to name only if called repeatedly</td>
<td>Slurring</td>
<td>Marked relaxation</td>
<td>Glazed/marked ptosis</td>
<td>3</td>
</tr>
<tr>
<td>Responds only after mild prodding</td>
<td>Not recognizable</td>
<td>—</td>
<td>—</td>
<td>2</td>
</tr>
<tr>
<td>No response to prodding or shaking</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
</tbody>
</table>

OAAS is the lowest score in any of the four categories.
continuation. The OAAS score was measured every ory tests. A second independent observer, who was score, Digit Symbol Substitution Test (DSST), and mem- baseline cognitive function were assessed by OAAS dent observer recorded BIS data.

on the aforementioned clinical guidelines. The indepen- cause the anesthesiologist was not aware of the BIS study drug was discontinued, and again at 10 min. Be-

was measured every minute during the first 5 min after administration of study drug and was recorded every minute during initiation of sedation, then every 5 min during recovery. At the conclusion of surgery, BIS was measured every minute during the first 5 min of recovery, then at 10 and 30 min. The DSST requires a subject to substitute symbols for numbers, and the DSST score represents the number of correct substitutions completed in 90 s. The DSST was given at baseline (before entering the operating room) and at 5, 10, and 30 min of recovery.

The memory tests included immediate recall and de- layed recall. The immediate recall test used tape record- ings of 16 words, balanced on word frequency and normative free recall, at a rate of one word every 5 s. One of three different word lists was used for the administra-
tion of each immediate recall test. Immediately after the list was completed, patients were asked to recall in any order as many of the words as possible from the list. The score of the immediate recall test was the number of correct words recalled immediately after listen- ing to one of three different word lists. The immedi- ate recall test was administered at baseline and 10 and 60 min of recovery. The delayed recall test involved asking the patients to recall the words from any previously presented lists and was administered 60 min and 24 h after the end of the procedure. The 60-min delayed recall test was given before the presentation of the 60-min immediate recall test and tested the recall of words from the baseline test and the 10-min word list. The 24-h delayed recall test tested recall from all three word lists. Separate word lists were not used for the delayed recall test.

Subjective self-assessment of quality of recovery was measured by visual analog scales determined at baseline and 5, 10, and 30 min of recovery. Attributes assessed (and scored from 0 to 100) included level of alertness—sedation (almost asleep to wide awake), energy level (no energy to full of energy), clear-headedness—confusion (confused to clear-headed), coordination—clumsiness (extremely clumsy to well coordinated), anxiety (calm—relaxed to extremely nervous), and nausea (none to severe nausea).

Statistical Analysis

All enrolled patients were included in analysis of intraoperative BIS, OAAS scores, complications, and side ef- fect data. Analysis of study drug administration and re-
covery data (OAAS, DSST, and memory scores) included only patients who successfully completed the study. One-way analysis of variance (ANOVA) was used to compare age, height, weight, American Society of Anesthesiologists status, fentanyl, and study drug administration among the groups. BIS scores for responders versus nonresponders, local versus regional, and opioid use was compared with t test or Mann–Whitney rank sum test. The BIS data for responsiveness to voice was com-
pared among the three drugs using Kruskal-Wallis one-
way ANOVA on ranks or one-way ANOVA, and post hoc analysis was performed with the Dunn method or Stu-
dent-Newman-Keuls method for multiple comparisons. The relation between BIS and probability of a positive

Table 2. Categories of Surgical Procedures and Anesthesia Technique

<table>
<thead>
<tr>
<th>Surgical procedures</th>
<th>Sevoflurane</th>
<th>Propofol</th>
<th>Midazolam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthopedic surgery</td>
<td>15</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Breast biopsy</td>
<td>5</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Inguinal hernia repair</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Miscellaneous (Hickman catheter placement, transurethral resection of bladder tumor, etc.)</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Anesthesiology, V 95, No 5, Nov 2001
response to verbal command was determined using logistic regression analysis of a quantal end point (response to voice defined as OAAS ≥ 3 and nonresponsive to voice defined as OAAS ≤ 2). DSST test scores were compared with a two-way repeated measures ANOVA model using treatment groups and time as factors and their interactions. Visual analog scale scores, postprocedure OAAS and BIS scores, memory scores, and time to postprocedure events were compared with one-way ANOVA or Kruskal-Wallis one-way ANOVA on ranks. Post hoc analysis was performed with Student-Neuman-Keuls method or the Dunn method for multiple comparisons, or t tests with a Bonferroni correction. Gender, race, return to preoperative OAAS and DSST scores, frequency of types of anesthesia, and side effects were compared with the chi-square test. Post hoc analysis was performed with Tukey-type multiple comparison tests. P values < 0.05 were considered significant. Results are designated as mean ± SD or median (range).

Ability of BIS and end-tidal sevoflurane concentration to predict the depth of sedation was evaluated using prediction probability (\( P_k \)). \( P_k \) is the probability that an indicator correctly predicts the depth of sedation. The mathematical basis of \( P_k \) was described by Smith et al.\textsuperscript{26} To compute the \( P_k \), the BIS score was analyzed as the predicting variable, and the true observed sedation depth was the value of the variable to be predicted. The \( P_k \) computed on this case is the estimate of the probability that the BIS will correctly predict the depth of sedation. An indicator that predicts perfectly the depth of sedation has a \( P_k \) value of 1.0, whereas an indicator that performs no better than chance has a \( P_k \) value of 0.5. \( P_k \) values were calculated for BIS (for sevoflurane, midazolam, and propofol separately) and end-tidal sevoflurane. The \( P_k \) values in each of the aforementioned categories were calculated in two different ways. First, the \( P_{k_{OAAS}} \) was calculated to indicate the probability of correctly predicting the OAAS score. The BIS score was analyzed as the predicting variable, and the OAAS score was the value of the variable to be predicted. The \( P_k \) computed on this case is the estimate of the probability that the BIS will correctly predict the OAAS score. Second, the \( P_{k_{VC}} \) was calculated to indicate the probability of correctly predicting whether a subject is responsive to voice command. The BIS score was analyzed as the predicting variable, and the ability of the subject to respond to voice command (responsiveness to voice defined as OAAS ≥ 3, nonresponsiveness to voice defined as OAAS ≤ 2) was the value of the variable to be predicted. The \( P_k \) computed on this case is the estimate of the probability that the BIS will correctly predict the responsiveness to voice. \( P_{k_{OAAS}} \) and \( P_{k_{VC}} \) values of sevoflurane BIS, midazolam BIS, propofol BIS, and end-tidal sevoflurane were compared using one-way ANOVA. Post hoc analysis was performed with the Student-Newman-Keuls method. The jackknife method was used to compute the SE of the estimate.\textsuperscript{26}

Results

Patient Characteristics

Sixty-six patients were assessable for intraoperative BIS, OAAS, and end-tidal sevoflurane data (23 sevoflurane, 22 propofol, 21 midazolam). Three patients (2 sevoflurane, 1 propofol; \( P = 0.384 \)) showed severe excitation–disinhibition and were prematurely withdrawn intraoperatorily. Sixty-three remaining patients (21 sevoflurane, 21 propofol, and 21 midazolam) were evaluable for recovery data. No significant difference in gender, race, height, weight, or American Society of Anesthesiologists status was found among groups.

Intraoperative Variables

Sevoflurane dose and end-tidal concentration were 0.4 ± 0.5 minimum alveolar concentration–h and 0.89 ± 0.25% (0.43 ± 0.12 minimum alveolar concentration). Average total and individual midazolam doses were 9.8 ± 3.7 and 0.9 ± 0.4 mg, respectively. Average propofol infusion rate and incremental bolus dose were 79 ± 32 \( \mu \)g · kg\(^{-1} \) · min\(^{-1} \) and 17 ± 11 mg, respectively. Sevoflurane, propofol, and midazolam were administered for 68 ± 37, 66 ± 35, and 76 ± 36 min, respectively \( (P = 0.642) \). There was no difference between groups in the number of patients who received fentanyl or the total dose administered (sevoflurane: \( n = 12 \), 81 ± 70 \( \mu \)g; propofol: \( n = 10 \), 70 ± 43 \( \mu \)g; midazolam: \( n = 10 \), 60 ± 29 \( \mu \)g). The type of surgical procedures and anesthetic technique were not different among groups. All patients who received remifentanil had an OAAS score of 5 and baseline BIS of 96 or greater before study drug administration. Time of initiation of sedation (minute from initial study drug administration to time maintenance occurred) was not different among groups (sevoflurane, 17 ± 10 min; propofol, 12 ± 3 min; midazolam, 12 ± 6 min). The time that sedation deviated from desired level (OAAS = 3) was similar among groups (65% sevoflurane, 60% midazolam, 59% propofol). The end-tidal sevoflurane concentration was significantly different between those who responded to voice command and nonresponders (0.75 ± 0.40 vs. 0.93 ± 0.41%; \( P < 0.001 \)).

Bispectral Index

Average BIS was significantly different between patients who responded to voice (OAAS ≥ 3) and nonresponders (OAAS ≤ 2) (table 3). However, mean BIS of sevoflurane responders was significantly different from midazolam and propofol responders, and mean BIS of propofol nonresponders was significantly different from midazolam and sevoflurane nonresponders. Mean BIS scores of patients who had spinal–epidural anesthesia...
were 1, 2, 3*, 6, and 11* higher than patients who had local anesthesia at OAAS scores of 5, 4, 3, 2, and 1, respectively (*P < 0.05). There was no difference in BIS scores between patients who did or did not receive fentanyl. The relation between BIS and probability of response to verbal command for all subjects receiving sevoflurane, propofol, and midazolam was determined using logistic regression analysis (fig. 1). The midazolam curve is incomplete because the lowest BIS recorded for patients who received midazolam was 65. The BIS50 (BIS at which 50% of patients respond to voice) was 77 for sevoflurane and 61 for propofol. BIS50 could not be calculated for midazolam because few patients who received midazolam were nonresponders.

Prediction probability values, including P_{KV} and P_{OAAS}, were calculated for BIS (for sevoflurane, midazolam, and propofol separately) and end-tidal sevoflurane. BIS was a better predictor (as measured by P_{OAAS}) of depth of sedation with propofol than midazolam or sevoflurane. P_{OAAS} for propofol BIS (0.87 ± 0.01, mean ± SEE) was significantly higher than P_{OAAS} for all other comparisons (midazolam, 0.69 ± 0.02; sevoflurane, 0.76 ± 0.01; and end-tidal sevoflurane, 0.40 ± 0.02; P < 0.05). The P_{KV} of sevoflurane BIS, midazolam BIS, and propofol BIS was also compared. No significant differences were found between the P_{KV} values for BIS during sedation with propofol (0.90 ± 0.02), midazolam (0.77 ± 0.15), and sevoflurane (0.80 ± 0.02). For sevoflurane, the P_{KV} BIS (0.80 ± 0.02) was significantly better than P_{KV} for end-tidal sevoflurane (0.41 ± 0.03; P < 0.05).

The relation between BIS, sedation depth, and end-tidal sevoflurane concentration showed large interindividual and intraindividual variability (fig. 2). For example, two subjects were responsive to voice at a BIS score of 41 and 52, whereas many patients were unresponsive at a BIS score of 98. No patient who received midazolam had a BIS of less than 65 (despite several patients achieving OAAS scores of 1 and 2). BIS scores at each OAAS level had significant overlap (fig. 3). For example, propofol patients with an OAAS score of 3 had BIS scores ranging from 41 to 98, whereas patients with an OAAS score of 1 ranged from 33 to 95. In addition, 25th–75th percentile BIS scores of responders versus nonresponders were 79–96 versus 65–83, respectively.

<table>
<thead>
<tr>
<th>Drug</th>
<th>Responders (OAAS = 3, 4, 5)</th>
<th>Nonresponders (OAAS = 1, 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td>86 ± 10*</td>
<td>74 ± 14</td>
</tr>
<tr>
<td>Sevoflurane</td>
<td>91 ± 9†</td>
<td>78 ± 13</td>
</tr>
<tr>
<td>Propofol</td>
<td>83 ± 11*</td>
<td>62 ± 12‡</td>
</tr>
<tr>
<td>Midazolam</td>
<td>85 ± 9*</td>
<td>77 ± 9</td>
</tr>
</tbody>
</table>

Values are mean ± SD. *P < 0.001 versus nonresponders. †P < 0.05 versus midazolam and propofol responders. ‡P < 0.05 versus sevoflurane and midazolam nonresponders.
Recovery Variables

Time from drug discontinuation to an OAAS score of 5 (8 ± 9, 13 ± 12, 9 ± 10 min), eligibility for discharge (39 ± 28, 34 ± 17, 35 ± 17 min), and actual discharge from phase 1 recovery (40 ± 38, 31 ± 35, 40 ± 47 min) were not different among sevoflurane, propofol, and midazolam, respectively.

Postprocedure OAAS scores were measured every minute for the first 5 min, then at 10 and 30 min after drug discontinuation. There was no difference in average OAAS scores at any time point after the procedure. There was also no significant difference in the percentage of patients who reached OAAS score of 5 at any time point after the procedure. All patients returned to an OAAS score of 5 at 30 min after the procedure.

Cognitive recovery was measured with DSST and memory tests. Average DSST scores were not different between groups. However, the percentage of patients whose DSST scores returned to baseline was significantly higher for sevoflurane at 10 min into recovery (fig. 4). Immediate recall scores and delayed recall tests during recovery were not significantly different among the groups. No difference was observed between sevoflurane, propofol, and midazolam for any visual analog scale score.

Side Effects

All enrolled patients were included in this analysis. Two sevoflurane patients and one propofol patient (P = nonsignificant) experienced severe excitation–disinhibition and were withdrawn from the study and analysis of recovery data, but not analysis of intraoperative and side-effect data. The occurrence of moderate disinhibition–excitement was more frequent with sevoflurane and propofol (16 [70%] sevoflurane, 8 [36%] propofol vs. 1 [5%] midazolam; P < 0.05 for all paired comparisons). Nausea was more frequent with sevoflurane (5 [22%] sevoflurane vs. 0 propofol and midazolam; P < 0.05). No difference was observed in the incidence of coughing, laryngospasm, apnea, and other complications between groups. The incidence of an event-free anesthetic was significantly greater with midazolam (17 [81%] midazolam vs. 7 [30%] sevoflurane and 11 [50%] propofol; P < 0.05 for all comparisons).

Discussion

Bispectral Index

The assessment of depth of sedation has traditionally been performed by observing clinical parameters such as appearance, response to voice, pain, and surgical stimulation. These parameters are qualitative, and assessment of response to voice requires patient stimulation, which may itself alter the depth of sedation. Bispectral Index demonstrates significant variability, and overlap of measurements between various depths of sedation makes it difficult to use as a guide of anesthetic management: “First, not only must the average values yielded by the device in two distinct states be statistically different, but also the range of values seen in those two states should not overlap.” Although Bispectral Index has been found to be an effective measure of depth of sedation with propofol, midazolam, isoflurane, and sevoflurane,1,3,4,6,27 the Bispectral Index demonstrates significant variability, and overlap of measurements between various depths of sedation makes it difficult to use as a guide of anesthetic management. In our investigation, mean Bispectral Index scores of patients who responded to voice were significantly different from those who did not respond. However, individual Bispectral Index scores demonstrated significant variability, making it difficult to rely on a single Bispectral Index value to predict depth of sedation. In addition, the range of values of patients who were unresponsive to voice overlapped with the range of values of patients who did respond to voice. At levels of deep sedation, several patients showed very high Bispectral Index scores (normally interpreted to suggest that the patient is wide awake), and several patients who were wide awake had low Bispectral Index scores. These results are similar to those reported elsewhere. For example, Sleigh and Donovan28 found no difference between average Bispectral Index scores in awake patients compared with average Bispectral Index scores of the same patients at time of unresponsiveness (measured by time of “syringe drop”) after propofol induction. Gajraj et al.13 found a mean Bispectral Index value during consciousness of 85 (range, 56–98), but Bispectral Index values during unconsciousness overlapped and ranged from 40 to 94. Iselin-Chaves et al.14 reported four patients who were responsive to voice with a Bispectral Index score of approximately 50, and one patient with an OAAS score of 1 (unresponsive to mild prodding) and a Bispectral Index of 95. Glass et al.27 reported one patient who was responsive to voice at a Bispectral Index of approximately 40. Mychaskiw et al.29 reported a case of explicit recall of intraoperative events with a Bispectral Index of 47.
The second condition is that “the critical threshold values that distinguish depth of anesthesia states of interest should not be influenced by choice of anesthetic agent.”

In our investigation, BIS was influenced by choice of anesthetic agent. BIS of the sevoflurane and propofol patients frequently decreased to less than 60, whereas the BIS of midazolam patients was never less than 65 despite several patients achieving an OAAS score of 1 (unresponsive). In addition, mean BIS scores of sevoflurane responders were significantly different from BIS scores of midazolam and propofol responders (BIS50 for sevoflurane was 77, but BIS50 for propofol was 61), and BIS scores of propofol nonresponders were significantly different from those of sevoflurane and midazolam nonresponders. The predictive ability (measured by the PK_OAAS) revealed that BIS was a slightly more accurate predictor of depth of sedation with propofol than sevoflurane or midazolam. Other investigators have also found that BIS is influenced by anesthetic agent. For example, Mi et al. found BIS values at unresponsiveness to voice command were 66 with propofol alone compared with 75 when propofol was administered with fentanyl. Kearse et al. found a BIS50 of 65 for propofol alone but 76 for propofol with nitrous oxide. The pooled BIS50 value for propofol, sevoflurane, and midazolam measured by Glass et al. was 67. Iselin-Chaves et al. reported a BIS50 of 64 for propofol alone and 72 for propofol plus alfentanil. Loss of response to voice occurred at a mean BIS score of 65 with propofol infusion and 94 when ketamine was added to the sedation regimen. Another study found that BIS and levels of sedation with ketamine do not correlate. Addition of opioids to sevoflurane or propofol anesthesia has also been found to decrease the correlation to patient movement. Vernon et al. found that BIS values of patients who received isoflurane-alfentanil anesthesia and did not respond to incision could not be differentiated from BIS values of patients who received propofol-alfentanil and moved in response to incision.

The explanation of these variations in BIS is not known. One interpretation is that regional anesthesia may increase interpatient variability. It has been speculated that spinal and epidural anesthesia may have an effect on BIS. Morley et al. found that spinal anesthesia increased the BIS, primarily because of an increase in power in the β wave band. In contrast, patients with combined lumbar epidural and general anesthesia had a BIS approximately 10 units lower than those with a general anesthetic alone, at the same end-tidal sevoflurane concentration. In the current investigation, there was no consistent difference in BIS between patients with or without regional anesthesia. Opioids may affect BIS values. One study found a significant correlation between remifentanil dose and BIS score, whereas others found that BIS cannot distinguish wakefulness from loss of response to command during fentanyl induction. Furthermore, BIS varies considerably during midazolam and fentanyl anesthesia and was not an accurate measure of depth of anesthesia when using this combination of agents. The present investigation found no difference in BIS scores among patients who received or did not receive fentanyl. In this study, remifentanil administration preoperatively may have affected intraoperative BIS. However, the number of patients who received remifentanil was small, remifentanil is rapidly eliminated, and all of these patients had a BIS greater than or equal to 96 and an OAAS score of 5 before initiating sedation; hence, this is unlikely. Another possibility may be that different components of anesthesia (loss of consciousness, obtunding motor or hemodynamic response) have different mechanisms; thus, a single index may be unable to successfully measure depth of anesthesia. Variability may be caused by nonelectroencephalograph electrical signals, which cause a common artifact and corrupt the BIS calculations and may be more prominent during sedation than general anesthesia. This includes the appearance of high-frequency facial electromyograph activity, which causes the BIS to err higher. BIS variability may also result from variation in electrode montage. Finally, the BIS score represents a 30-s moving average of previously collected data, plus an additional lag time for signal processing; thus, frequent changes in anesthetic depth do not reflect a steady state and increase intrapatient variability in BIS data. In summary, based on the data found in this investigation, the BIS as a monitor for depth of sedation did not fulfill either of the two criteria suggested for guiding anesthetic management.

Other investigations have used PK to assess the ability of BIS to measure depth of sedation. Although these other results showed a greater accuracy of BIS as an anesthetic depth indicator than found in this investigation, this may be a result, in part, of differences in methods. Katoh et al. administered sevoflurane at depths ranging from sedation to general anesthesia, and did so before the start of surgery, and patients were maintained at a constant level of sedation for 15 min without stimulation. Iselin-Chaves et al. studied healthy volunteers maintained with target-controlled infusions of propofol and alfentanil at steady state for 10 min. In contrast, our subjects were undergoing surgery with conditions and levels of stimulation that were constantly changing, causing alterations in level of consciousness and BIS. Electroencephalograph effect has been found to lag behind changes in anesthesia (specifically, end-tidal sevoflurane), possibly because of neuronal or receptor-related dynamics or the time required for anesthetic wash-in and wash-out in the brain (sevoflurane t1/2keo of BIS has been reported to be 3.5 ± 2.0 min).
End-tidal Sevoflurane Measurements

In contrast to Katoh et al.,1 we found a poor correlation between end-tidal sevoflurane concentration and depth of sedation. Katoh et al. maintained constant end-tidal sevoflurane concentrations for 15 min before each BIS measurement, to establish equilibrium and accurately reflect brain concentration.4 Our investigation, in contrast, involved rapid alterations of inspired sevoflurane concentration in response to changing surgical conditions. Variable stimulation, which is typical of surgery performed with local anesthesia, leads to greater intersubject variability in depth of sedation because arousal of the patient results in a fluctuating clinical state despite maintenance of constant drug concentrations. Furthermore, our end-tidal measurement may not accurately reflect effect site (brain) concentration because of rapid alterations of inspired sevoflurane concentrations. However, our results may be more clinically applicable, because the operating room is a dynamic environment requiring frequent alterations in drug concentration in response to changing conditions. Because end-tidal sevoflurane ED50 values for loss of response to command are significantly different among the groups of young and older patients,5 age-related differences in our patient population may also have contributed to our lower predictive probability. Therefore, our results do not support the use of end-tidal sevoflurane measurement as a sedation depth indicator when inhaled sevoflurane is administered by facemask and requiring frequent alterations in drug concentration. BIS is a better predictor of sedation depth in this situation than end-tidal sevoflurane measurements.

Sedation

Sevoflurane, midazolam, and propofol all provided effective sedation during surgery with regional or local anesthesia, with similar quality of recovery. Conversely, other studies have found significant differences. Propofol59–62 and sevoflurane6 sevoflurane and sevoflurane18 sedation is associated with more rapid recovery and faster return of cognitive function than midazolam. Although sevoflurane sedation has never previously been compared with propofol, general anesthesia with sevoflurane results in faster emergence and improved cognitive performance than propofol.43,44 The small number of patients in this investigation may provide an explanation for the lack of a significant difference between drugs.

The largest difference between sedation techniques was the incidence of excitation–disinhibition. Excitation occurred most frequently with sevoflurane followed by propofol. Sevoflurane excitation has been rarely reported in adults undergoing mask induction.44,45 Propofol excitation has been reported in 14–33% of patients.46–48 The explanation for the development of excitation–disinhibition is unclear. The incidence of excitation with sevoflurane and propofol may seem higher than observed in typical clinical use possibly because each study group received a “pure” anesthetic. A combination (e.g., midazolam, fentanyl) may reduce the incidence of excitation–disinhibition. Furthermore, this investigation targeted an OAAAS score of 3, whereas many clinical sedation and monitored anesthetic care cases target a deeper level of sedation. Because of the high incidence of excitation with both sevoflurane and propofol, midazolam was believed to provide a less eventful, smoother sedation technique, if only one drug is to be used.

In summary, individual BIS scores demonstrate wide variability, making it difficult to assess depth of sedation. BIS appears to have drug-specific characteristics. BIS was a slightly more accurate predictor of depth of sedation with propofol than sevoflurane or midazolam. All drugs provided similar quality of recovery, but excitation occurred more frequently with sevoflurane and propofol.

The authors thank Warren D. Smith, Ph.D. (Professor, Biomedical Engineering Program, California State University, Sacramento, CA), for his contribution and assistance with analysis of the P5 statistic; and Meghan Creger, B.A. (Clinical Research Assistant, University of Washington, Seattle, WA), for assistance with the performance of this investigation.

References

5. Liu J, Singh H, Bito H, Sata S: Influence of age on hypnotic requirement, bispectral index, and 95% spectral edge frequency associated with sedation induced by sevoflurane. ANESTHESIOLOGY 2000; 92:55–61
15. Mi WD, Sakai T, Singh H, Kudo T, Kudo M, Matsuki A: Hypnotic endpoints vs. the bispectral index. 95% spectral edge frequency and median frequency