Ability of the Bispectral Index, Autoregressive Modelling with Exogenous Input–derived Auditory Evoked Potentials, and Predicted Propofol Concentrations to Measure Patient Responsiveness during Anaesthesia with Propofol and Remifentanil


Background: This study was conducted to compare the performance accuracy of the independent variables Bispectral Index (BIS), A-Line ARX index (AAI), and predicted propofol effect-site concentration (CePROP) to measure the dependent variables of loss of responses to different stimulation defined as loss of response to verbal command (LOR\text{verbal}), eyelash reflex (LOR\text{lash}), and noxious stimulus (LOR\text{noxious}) during stepwise increased levels of propofol infusion with and without remifentanil.

Methods: Forty-five patients were randomly allocated to one of three groups (0, 2, and 4 ng/ml remifentanil) to receive graded CePROP and predicted effect compartment controlled remifentanil (CeREMI). At every step, the ability to respond to verbal command using the Observer’s Assessment of Alertness/Sedation Scale (OAA/S), eyelash reflex, and electrical tectonic noxious stimulus were compared against BIS, AAI, and CePROP. Prediction probability and sensitivity/specificity were calculated.

Results: Increasing CeREMI increased BIS and AAI values at LOR\text{verbal} and LOR\text{lash} and increasing CePROP. Similar findings were found for LOR\text{noxious}. The overall prediction probability to measure the hypnotic component of anesthesia remained accurate in the three groups for BIS, AAI, and CePROP. Combined information from CePROP, CeREMI, and BIS or AAI increased the overall prediction probability for predicting the OAA/S scale and LOR\text{lash}. Less accuracy to LOR\text{noxious} was found in all independent variables.

Conclusions: Although BIS, AAI, and CePROP were influenced by remifentanil during propofol administration, their ability to detect OAA/S and LOR\text{lash} remained accurate. Improved performance is obtained when BIS and AAI are measured in conjunction with drug targeted effect-site concentrations. Remifentanil decreases the ability of these independent variables to detect LOR\text{noxious}.

Both electroencephalography- and midlatency auditory evoked potential (MLAEP)-derived variables have been proposed as measures of the hypnotic state during anaesthesia. For the electroencephalography, the Bispectral Index (BIS) incorporated in the A-2000 BIS® monitor (Aspect Medical Systems, Inc., Newton, MA) has been proven to have a high sensitivity and specificity to measure anesthetic drug effect, compared with other processed electroencephalographic variables. Previously, MLAEP has been used by various investigators to study anaesthetic depth. Recently, Struys et al. developed a new method for extracting the MLAEP from the electroencephalographic signal by using an autoregressive model with an exogenous input (ARX) adaptive model. This method allows extraction of the MLAEP signal within 15–25 sweeps of 110 milliseconds’ duration each, resulting in only a 6- to 15-s response delay time. A new monitoring variable, called the A-Line ARX index (AAI), is then calculated from this fast extracted MLAEP wave. This new technology is incorporated in a recently commercialized system called A-Line® (Danmeter A/S, Odense, Denmark). Recently, Struys et al. compared the accuracy of both BIS and AAI for measuring loss of responses to different stimulation defined as loss of verbal command (LOR\text{verbal}), loss of eyelash reflex (LOR\text{lash}), and loss of response to noxious stimulus (LOR\text{noxious}) during steady state propofol administration. We found that BIS, AAI, and predicted effect-site concentration of propofol (CePROP) revealed a similar level of information on LOR\text{verbal} and LOR\text{lash} but did not predict LOR\text{noxious}. Also, AAI has recently been shown to reliably assess the level of consciousness during propofol, sevoflurane, and midazolam anesthesia. The BIS correlated better to propofol plasma concentrations than AAI, but AAI correlated better than BIS to the clinical signs during recovery from propofol anesthesia.

In previous work done with the AAI, propofol was administered solely. Although a clear pharmacodynamic interaction between remifentanil and propofol has been described, controversy still exists on the influence of opiates on the accuracy of both BIS and MLAEP to measure loss of response to different stimulation. During remifentanil monoinfusion, AAI measured changes in patients’ level of arousal better than BIS. However, in this study, a different electrode position was used, where the AEP was recorded between a frontal (+) and an occipital electrode (−). It is possible that a change in
montage such as this could lead to significantly different results compared with those from the current study.

A fundamental question is whether the electroencephalograph- or AEP-derived variables can be used to optimize drug delivery. A first step to answer this question is to examine the performance and accuracy of these monitors under various conditions and to compare them with the accuracy and usefulness of on-line calculated drug effect-site concentration. A secondary question is how opiates affect these performance parameters. This study was conducted to assess the performance accuracy of the AAI to reflect the hypnotic component of anesthesia and to measure loss of responses to different stimulation defined as LORverbal, LORlash, and LORnoxious during stepwise increased levels of propofol infusion with and without remifentanil. The performance of AAI was compared with that of BIS and CePROP.

Materials and Methods

After Institutional Ethics Committee (Ghent University Hospital, Gent, Belgium) approval, informed consent was obtained from 45 female patients with American Society of Anesthesiologists class I, aged 18–60 yr, who were scheduled to undergo ambulatory gynecologic surgery. Exclusion criteria included weight less than 70% or more than 130% of ideal body weight, neurologic disorder, and recent use of psychoactive medication, including alcohol. They were randomly allocated to one of the three groups. In all groups, patients received a “staircase” computer-controlled infusion of propofol targeting the effect compartment. Initially, an effect-site concentration of 1.5 μg/ml was targeted in the group without remifentanil, and 1 μg/ml was targeted in the two other groups; this was increased every 4 min by 0.5 μg/ml until loss of response to all relevant clinical measures of anesthetic depth was observed. In the 0 ng/ml remifentanil group, no remifentanil was given. In the 2 ng/ml and 4 ng/ml remifentanil groups, an effect compartment controlled infusion of remifentanil was started 4 min before the start of propofol. Calculated remifentanil effect-site concentrations (CeREMI) of 2 and 4 ng/ml were targeted, respectively.

Propofol and remifentanil were administered via a computer-assisted continuous infusion device to a target effect-site concentration (RUGLOOP) using a three-compartment model enlarged with an effect-site compartment. For propofol, the pharmacokinetic–dynamic model previously published by Schnider et al.18,19 was used. For remifentanil, the pharmacokinetic–dynamic model previously published by Minto et al.20,21 was used. CePROP was computed to yield a time-to-peak effect of 1.6 min after bolus injection, as also published by Schnider et al.18,19 and clinically confirmed by Struys et al.25 For remifentanil, a t1/2 of 1.020619 min was applied as published by Minto et al.20,21 Propofol and remifentanil infusion were administered using a Fresenius Modular DPS Infusion Pump connected to a Fresenius Base A (Fresenius Vial Infusion Systems, Brézins, France). RUGLOOP steers the pump at infusion rates between 0 and 1,200 ml/h via an RS-232 interface. By using this infusion technique, we were able to obtain a steady state condition for both propofol and remifentanil at every target level after 4 min of infusion. Hereby, steady state is defined as the equilibrium between the calculated plasma and effect-site concentration of the drug. Remifentanil and propofol were infused via a large left forearcm vein. Every patient received approximatelv 200 ml crystalloid fluid during the study period. No fluid load was given before induction. No patient received preanesthetic medication. No other drugs were given. All patients maintained spontaneous ventilation via a facemask delivering 6 l/min O2.

Heart rate and noninvasive blood pressure, oxygen saturation measured by pulse oximetry (SpO2), and capnography were recorded at 1-min intervals using an A53 monitor (Datex, Helsinki, Finland). BIS (version 3.4) was derived from the frontal electroencephalogram (At-Fpz1) and calculated by the A-2000 BIS® monitor using a BIS-Sensor® (Aspect Medical Systems, Inc.). The standardization time of the BIS® monitor was set at 15 s. The AAI from the MLAEP was calculated using the A-Line® monitor. The MLAEPs were elicited with a bilateral click stimulus of 70 dB in intensity and 2 ms in duration. Three electrodes (A-Line AEP electrodes; Danmeter A/S) were positioned at the mid forehead (+), left forehead (reference), and left mastoid (−). The extraction of the MLAEP using a short moving time average technique together with an ARX model and the calculations of the AAI are described elsewhere.8

Ten seconds before each increase in CePROP (after 4 min of infusion at the specific target effect-site concentration), the independent variables of response (BIS, AAI, and CePROP) were recorded. Immediately after that, patient responsiveness to different stimuli was tested using the dependent variables eyelash reflex, Observer’s

<table>
<thead>
<tr>
<th>Score</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Responds readily to name spoken in normal tone</td>
</tr>
<tr>
<td>4</td>
<td>Lethargic response to name spoken in normal tone</td>
</tr>
<tr>
<td>3</td>
<td>Responds only after name is called loudly and/or repeatedly</td>
</tr>
<tr>
<td>2</td>
<td>Responds only after mild prodding or shaking</td>
</tr>
<tr>
<td>1</td>
<td>Responds only after painful trapezius squeeze</td>
</tr>
<tr>
<td>0</td>
<td>No response after painful trapezius squeeze</td>
</tr>
</tbody>
</table>

Table 1. Responsiveness Scores of the Modified Observer’s Assessment of Alertness/Sedation Scale

---

44 RUGLOOP, written by Tom De Smet, M.Sc. (Medical Engineer, DEMED Engineering, Temse, Belgium), and Michel M. R. F. Struys, M.D., Ph.D. (Professor of Anesthesia, Ghent University, Gent, Belgium). More information available at http://www.anesthesia-uzgent.be.

Anesthesiology, V 99, No 4, Oct 2005
Assessment of Alertness/Sedation Scale (OAA/S²⁴, table 1), and response to electrical tetanus (at 100 Hz and 50 mA for 2 s; applied to the volar forearm level) assessments, in that order. Verbal stimulation proceeded noxious stimulus. Patients were considered responsive to verbal stimulus if their OAA/S value was 3, 4, or 5 and were considered unresponsive to verbal command at OAA/S values of 0, 1, and 2. Transition between a responsive and an unresponsive state was defined as loss of response to verbal command (LOR
\textsubscript{verbal}). Loss of response to eyelash reflex was defined as LOR
\textsubscript{lash}. Patients were considered responsive to noxious stimulus if they responded to the tetanic electrical stimulation, regardless of whether they responded to trapezius squeeze during the OAA/S assessment. Loss of response to tetanic stimulus was defined as LOR
\textsubscript{noxious}.

Both BIS and AAI indices were also logged automatically. RUGLOOP digitally recorded the BIS index every 10 s, and the A-Line monitor recorded AAI index values nominally every 6 s. The time marks of both systems were synchronized with the manual timing for stimulus and manually recorded events to within ±1 s.

Statistical Analysis

The significance level was set at 5% unless otherwise reported. The ability of the independent variables (BIS, AAI, and CePROP) to detect the level of OAA/S, LOR
\textsubscript{lash}, and LOR
\textsubscript{noxious} was evaluated using prediction probability (P
\textsubscript{k}), which compares the performance of independent variables with different units of measure, as developed by Smith et al.²⁵,²⁶ Consider a variable such as BIS or AAI and a “definitive standard” measure of anesthetic depth such as the multilevel OAA/S score or the two-level responsiveness (yes/no) to eyelash reflex or noxious stimulus. Then, a P
\textsubscript{k} of 1 for the BIS or AAI variable would mean that BIS or AAI always increases (decreases) as the anesthesia gets lighter (deeper) according to the definitive standard depth measure. Such an independent variable can perfectly measure anesthetic depth. Alternatively, a P
\textsubscript{k} value of 0.5 would mean that the independent variable is useless for measuring anesthetic depth. For the OAA/S score, a P
\textsubscript{k} was computed for all OAA/S levels combined. Similarly, P
\textsubscript{k} values for LOR
\textsubscript{lash} and LOR
\textsubscript{noxious} were determined. The jackknife method was used to compute the SE of the estimate, based on the assumption that all assessments were independent.²⁵,²⁶ A Student t test with Bonferroni correction was used to evaluate whether the P
\textsubscript{k} for one variable was different from another one. Significance level was set at 0.0167. Prediction probability was calculated using a custom spreadsheet macro, P
\textsubscript{k}MACRO, developed by Smith et al.²⁵,²⁶ The P
\textsubscript{k} value was calculated for each independent variable in each group. A P
\textsubscript{k} analysis for each independent variable with the three groups pooled was performed.

To study whether the combined information from BIS and AAI together with the drug effect-site concentrations offers more accurate information than the independent variables alone, two new composite variables have been designed based on the combined information from BIS + CePROP + CeREMI and AAI + CePROP + CeREMI, and a P
\textsubscript{k} analysis was performed on both composite variables. The OAA/S score was used as clinical comparator. The P
\textsubscript{k} on the combined information was calculated using ordinal logistic regression, as described in the Appendix.

The power on the P
\textsubscript{k} values was calculated using a t statistic defined as the difference considered of clinical importance divided by the SE of the difference between two independent variables. Assuming a P
\textsubscript{k} difference of 0.05 as being of significant with an SE of 0.02, 15 patients should be included to find significant differences with P < 0.025 (Bonferroni correction for two t tests). Previous assumption were based on previous results of the AAI and other AEP-independent variables.²⁷

By applying Probit analyses, the effective concentration or index at which 50% (ED
\textsubscript{50}) and 95% (ED
\textsubscript{95}) of the patients reached LOR
\textsubscript{verbal}, LOR
\textsubscript{lash}, and LOR
\textsubscript{noxious} were calculated for all independent variables. For all independent variables, ED
\textsubscript{50} and ED
\textsubscript{95} values were compared between groups using one-way analysis of variance statistics. If significant, an unpaired two-sided Student t test with Bonferroni correction was used (P < 0.0167).

We calculated cutoff (threshold) values for the ability of the BIS, AAI, and CePROP to detect LOR
\textsubscript{verbal}, LOR
\textsubscript{lash}, and LOR
\textsubscript{noxious} in each group. For these calculations, we used “positive” to denote a test result that suggested responsiveness and “negative” to denote a test result that suggested nonresponsiveness. We assumed that increases in the BIS and AAI and a decrease in CePROP corresponded to an increased likelihood of responsiveness. We computed sensitivity as the proportion of responsive patients with positive test results (value higher than cutoff value for BIS and AAI and lower than cutoff value for CePROP). Similarly, we computed specificity as the proportion of nonresponsive patients with negative test results (value lower than cutoff value for BIS and AAI and higher than cutoff value for CePROP). We computed the cutoff values for each independent variable and specificity at a level of 100% sensitivity and at which the sum of sensitivity and specificity were highest. The same sensitivity/specificity analyses were performed for composite variables.

Results

The demographics (mean ± SD) of the 45 female patients in the three groups are shown in table 2. No significant demographic differences were found between groups.

Figures 1A and B show the behavior of BIS and AAI
versus CePROP and CeREMI. The correlation coefficient between BIS and the drug effect-site concentrations ($r = 0.92$) was significantly higher than for AAI ($r = 0.82$). In all groups, a stepwise increase in CePROP resulted in a monotonic decrease in the OAA/S score. BIS and AAI decreased in all groups with decreasing OAA/S scores, as shown in figure 2.

Figures 3A–C show the behavior of BIS, AAI, and CePROP at LOR lash. For all independent variables, a significant decrease of the variable values between the responsive and unresponsive state was observed in all groups. The BIS, AAI, and CePROP values at LORnoxious are plotted in figures 3D–F. In all groups, all independent variables changed significantly.

By applying Probit analyses, the effective BIS, AAI, or CePROP at which 50% ($ED_{50}$) and 95% ($ED_{95}$) of the patients produced a LOR verbal, LOR lash, and LOR noxious were calculated and are shown in table 3. For all three groups, the probability curves for the observations at LORverbal, LORlash, and LORnoxious are shown in figures 4A–C for BIS, in figures 4D–F for AAI, and in figures 5A–C for CePROP.

The ability of the BIS, AAI, and CePROP to predict the level of the OAA/S, LORlash, and LORnoxious as presented by the $PK$ values is shown in table 4. Overall, similar performances were found for BIS, AAI, and CePROP. A $PK$ analysis for each independent variable with the three groups pooled revealed similar ability to detect both the

---

Table 2. Anthropometry

<table>
<thead>
<tr>
<th></th>
<th>0 ng/ml Remifentanil</th>
<th>2 ng/ml Remifentanil</th>
<th>4 ng/ml Remifentanil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>33 ± 5</td>
<td>33 ± 5</td>
<td>34 ± 4</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>63 ± 10</td>
<td>66 ± 11</td>
<td>65 ± 9</td>
</tr>
<tr>
<td>Height, cm</td>
<td>167 ± 6</td>
<td>168 ± 6</td>
<td>167 ± 6</td>
</tr>
</tbody>
</table>

---

Fig. 1. Nonlinear regression analysis from the raw data from all groups (0, 2, and 4 ng/ml remifentanil effect-site concentration [Ce]) at different propofol effect-site concentrations. ($A$) Correlation with the Bispectral Index (BIS); ($B$) correlation with the A-Line ARX index (AAI).

Fig. 2. Raw data at every Observer’s Assessment of Alertness/Sedation Scale (OAA/S) score for Bispectral Index (BIS; $A$), A-Line ARX index (AAI; $B$), and propofol effect-site concentration (Ce; $C$). Data from the 0 ng/ml remifentanil (Remi) group are presented as black circles, data from the 2 ng/ml remifentanil group as white circles, and data from the 4 ng/ml remifentanil group as black triangles.
OAA/S levels and LORlash. In contrast, the overall ability for detection of LORnoxious was lower when pooling the groups together.

By calculating the $t$ statistic, we found that this study including 15 patients for each group had the power to determine significant differences between independent variables to predict OAA/S larger than 0.052, which is in accordance with our initial assumption that only differences larger than 0.05 would be considered significantly different.

The $P_K$ values to detect the OAA/S level using the combinations of input variables were better compared to the $P_K$ values of pooled data for BIS, AAI, and CePROP alone. For the composite index using the combination BIS + CePROP + CeREMI, the $P_K$ value was 0.91 (SE = 0.01), and for the composite index using the combination AAI + CePROP + CeREMI, it was 0.93 (SE = 0.01) to detect the OAA/S level.

We also performed more in-depth sensitivity/specificity analysis. Within the three groups, table 5 shows the cutoff values for each independent variable at which the sum of the sensitivity and specificity was the highest, representing the independent variable value where the

---

![Fig. 3. (A–C) Raw data for (no) response to eyelash reflex for Bispectral Index (BIS), A-Line ARX index (AAI), and propofol effect-site concentration (Ce). (D–F) Raw data for (no) response to noxious stimulus for BIS, AAI, and propofol effect-site concentration. $P < 0.05$ for each group between the responsive and nonresponsive states. Data from the 0 ng/ml remifentanil (Remi) group are presented as black circles, data from the 2 ng/ml remifentanil group as white circles, and data from the 4 ng/ml remifentanil group as black triangles.

Table 3. ED$_{50}$ (95% CI)/ED$_{95}$ Values of BIS, AAI, and EC$_{50}$ (95% CI)/EC$_{95}$ Propofol for All Groups at LORverbal, LORlash, and LORnoxious

<table>
<thead>
<tr>
<th>Group</th>
<th>Remifentanil</th>
<th>Remifentanil</th>
<th>Remifentanil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I: 0 ng/ml Remifentanil</td>
<td>26 (25–28)/20*§</td>
<td>33 (31–34)/18*§</td>
<td>40 (39–42)/25*§</td>
</tr>
<tr>
<td>Group II: 2 ng/ml Remifentanil</td>
<td>61 (60–62)/55*§</td>
<td>66 (65–68)/58*§</td>
<td>77 (75–78)/65*§</td>
</tr>
<tr>
<td>Group III: 4 ng/ml Remifentanil</td>
<td>2.9 (2.7–3.1)/3.8*#</td>
<td>2.4 (2.2–2.6)/3.1*#</td>
<td>2.0 (1.8–2.2)/2.7*#</td>
</tr>
</tbody>
</table>

* $P < 0.0167$ for ED$_{50}$ levels between all groups. † $P < 0.0167$ for ED$_{50}$ levels between group I vs. II and I vs. III. ‡ $P < 0.0167$ for ED$_{95}$ levels between all groups. § $P < 0.0167$ for ED$_{50}$ levels between group I vs. III and II vs. III. ¶ $P < 0.0167$ for ED$_{95}$ levels between group I vs. III and I vs. III.

AAI = A-Line ARX index; BIS = Bispectral Index; CePROP = calculated propofol effect-site concentration; LORverbal = loss of response to verbal command; LORMlash = loss of eyelash reflex; LORnoxious = loss of response to electrical tetanic stimulus.

---

Anesthesiology, V 99, No 4, Oct 2003

Downloaded From: http://anesthesiology.pubs.asahq.org/pdfaccess.ashx?url=/data/journals/jasa/931203/ on 10/12/2018
overall "errors" are minimized.28 Table 6 shows the different cutoff values and their specificity at level of 100% sensitivity. When using the combined information of BIS or AAI together with CePROP and CeREMI, the sensitivity/specificity profiles were more accurate. For both combinations, the best sensitivity-versus-specificity combinations were found at 100% sensitivity for 82% specificity.

Table 7 shows at which average OAA/S score at LORlash and LORnoxious occurred in each of the three groups. The OAA/S is an ordinal scale, but it was considered continuous to be able calculate an average of OAA/S score with a decimal value.

Discussion

This study was conducted to compare the performance accuracy of the independent variables BIS, AAI, and CePROP to measure anesthetic depth and how opiates quantitatively influence this information. Therefore, we observed the influence of remifentanil on the accuracy of the AAI, a new index calculated from the MLAEP, to reflect the hypnotic component of anesthesia (measured by the different levels of the OAA/S) and to measure loss of responses to different stimulation defined as LORverbal, LORlash, and LORnoxious during stepwise increased levels of propofol infusion with and without remifentanil. The performance of AAI was compared with that of BIS and CePROP.

In previous work, our group concluded that, during propofol anesthesia (without opiates) under steady state conditions, BIS, AAI, and CePROP were accurate independent variables to measure the hypnotic component of anesthesia and loss of responsiveness to different stimuli. Hemodynamic variables did not perform accurately enough to measure the hypnotic-anesthetic status of the patient. BIS correlated best with CePROP, followed by the AAI. Hemodynamics did not correlate well.8 To study the hypnotic component of anesthesia, two clinical measures (also called dependent variables) were used. We selected the OAA/S score because it provides a good correlation with a clinical reflection of the hypnotic component of anesthesia and has been

Fig. 4. Probability of loss of response to verbal commend (LORverbal), eyelash reflex (LORlash), and a tetanic electrical stimulus (LORnoxious) as a function of Bispectral Index (BIS) and A-Line ARX index (AAI). Data from the 0 ng/ml remifentanil (Remi) group are presented as a solid line, data from the 2 ng/ml remifentanil group as a dotted line, and data from the 4 ng/ml remifentanil group as a dashed line.

Fig. 5. Probability of loss of response to verbal commend (LORverbal), eyelash reflex (LORlash), and a tetanic electrical stimulus (LORnoxious) as a function of the propofol effect-site concentration (Ce). Data from the 0 ng/ml remifentanil (Remi) group are presented as a solid line, data from the 2 ng/ml remifentanil group as a dotted line, and data from the 4 ng/ml remifentanil group as a dashed line.
tested prospectively. Also, loss of eyelash reflex was used because it is a simple binary variable and commonly applied by anesthesiologists in clinical practice to detect loss of consciousness.

In all groups, a stepwise increase in CePROP resulted in a monotonic decrease in BIS and AAI as shown in figure 2. Similar to our previous work using propofol alone, the correlation between BIS and the combined drug effect-site concentrations was higher than for the AAI. In all groups, a monotonic decrease in both BIS and AAI was observed at decreasing OAA/S levels as shown in figure 2. Similar findings were found for LORlash as seen in figures 3A–C.

LORverbal and LORlash were reached at higher independent variable values in a dose-dependent manner when adding remifentanil. This can be concluded from the Probit analyses shown in figure 4 and the extracted ED50 and EC95 as seen in table 3. Although some differences between groups for the ED50 and ED95 did not reach the level of significance, a clear trend could be observed. Our results are in contrast with the findings from Iselin-Chaves et al. who reported no significant influence of opiates on the BIS and MLAEP to detect the probability of loss of consciousness and lack of recall by using the OAA/S score. However, they used lower concentrations of opiates (50 and 100 ng/ml alfentanil; potency ratio alfentanil:remifentanil around 40:1). Recently, Mi et al. also found that when applying fentanyl pretreatment during propofol administration, the hypnotic endpoints were achieved at higher BIS values. The addition of remifentanil potentiates the effect of propofol. LORverbal and LORlash were detected at lower CePROP in the 2 and 4 ng/ml remifentanil groups compared to the 0 ng/ml remifentanil group, as shown in figure 5. Also, EC50 and EC95 propofol decreased significantly or tended to decrease in a remifentanil dose-dependent manner, as also described previously by others.

A possible bias might have been caused by the influence of opiates on the clinical dependent variables. Both

### Table 4. Prediction Probability (Pp) Described as Mean (SE) for Each Independent Variable (BIS, AAI, and CePROP) Using OAA/S, LORlash, and LORnoxious

<table>
<thead>
<tr>
<th>OAA/S</th>
<th>0 ng/ml Remifentanil</th>
<th>2 ng/ml Remifentanil</th>
<th>4 ng/ml Remifentanil</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAI</td>
<td>0.88 (0.02)</td>
<td>0.87 (0.02)</td>
<td>0.88 (0.02)</td>
</tr>
<tr>
<td>BIS</td>
<td>0.93 (0.01)</td>
<td>0.90 (0.02)</td>
<td>0.88 (0.02)</td>
</tr>
<tr>
<td>CePROP</td>
<td>0.92 (0.02)</td>
<td>0.92 (0.03)</td>
<td>0.92 (0.02)</td>
</tr>
<tr>
<td>LORlash</td>
<td>0.94 (0.02)</td>
<td>0.96 (0.02)</td>
<td>0.87 (0.04)</td>
</tr>
<tr>
<td>BIS</td>
<td>0.96 (0.02)</td>
<td>0.92 (0.03)</td>
<td>0.94 (0.04)</td>
</tr>
<tr>
<td>CePROP</td>
<td>0.94 (0.02)</td>
<td>0.95 (0.02)</td>
<td>0.94 (0.02)</td>
</tr>
<tr>
<td>LORnoxious</td>
<td>0.87 (0.04)*</td>
<td>0.86 (0.04)*</td>
<td>0.81 (0.05)*</td>
</tr>
<tr>
<td>BIS</td>
<td>0.88 (0.05)*</td>
<td>0.86 (0.05)</td>
<td>0.86 (0.04)*</td>
</tr>
<tr>
<td>CePROP</td>
<td>0.83 (0.06)</td>
<td>0.86 (0.04)*</td>
<td>0.87 (0.04)*</td>
</tr>
</tbody>
</table>

* P < 0.05 compared to pooled data P0:

AIAI = A-Line ARX index; BIS = Bispectral Index; CePROP = calculated propofol effect-site concentration; OAA/S = Observer’s Assessment of Alertness/Sedation Scale; LORlash = loss of eyelash reflex; LORnoxious = loss of response to electrical tetanic stimulus.

### Table 5. Cutoff Values (Specificity) at Level of 100% Sensitivity

<table>
<thead>
<tr>
<th>0 ng/ml Remifentanil</th>
<th>2 ng/ml Remifentanil</th>
<th>4 ng/ml Remifentanil</th>
</tr>
</thead>
<tbody>
<tr>
<td>LORverbal</td>
<td>29 (91–91; 182)</td>
<td>42 (82–81; 163)</td>
</tr>
<tr>
<td>AAI</td>
<td>61 (99–94; 193)</td>
<td>64 (95–76; 171)</td>
</tr>
<tr>
<td>BIS</td>
<td>3.5 (99–76; 176)</td>
<td>3.0 (98–57; 155)</td>
</tr>
<tr>
<td>CePROP</td>
<td>34 (90–88; 178)</td>
<td>47 (96–82; 171)</td>
</tr>
<tr>
<td>LORlash</td>
<td>67 (90–80; 178)</td>
<td>75 (89–65; 174)</td>
</tr>
<tr>
<td>AAI</td>
<td>23 (63%)</td>
<td>38 (59%)</td>
</tr>
<tr>
<td>BIS</td>
<td>53 (46%)</td>
<td>68 (65%)</td>
</tr>
<tr>
<td>CePROP</td>
<td>4.0 (39%)</td>
<td>3.5 (14%)</td>
</tr>
<tr>
<td>LORnoxious</td>
<td>20 (76–64; 140)</td>
<td>40 (92–71; 163)</td>
</tr>
<tr>
<td>AAI</td>
<td>44 (84–71; 155)</td>
<td>75 (76–71; 147)</td>
</tr>
<tr>
<td>BIS</td>
<td>4.0 (80–71; 151)</td>
<td>2.5 (88–50; 138)</td>
</tr>
</tbody>
</table>

AIAI = A-Line ARX index; BIS = Bispectral Index; CePROP = calculated propofol effect-site concentration; LORlash = loss of eyelash reflex; LORnoxious = loss of response to electrical tetanic stimulus; LORverbal = loss of response to verbal command.

At the 98% sensitivity level.

AIAI = A-Line ARX index; BIS = Bispectral Index; CePROP = calculated propofol effect-site concentration; LORlash = loss of eyelash reflex; LORnoxious = loss of response to electrical tetanic stimulus; LORverbal = loss of response to verbal command.
OAA/S (from level 2 to 0) and eyelash reflex use tactile stimuli, which might have been lost earlier when adding remifentanil. To demonstrate this, the averaged OAA/S scores at LORlash and LORnoxious were calculated and are shown in table 7. LORlash results showed some increase. A clear increase in the OAA/S value is observed when remifentanil is added. In the 2 and 4 ng/ml remifentanil groups, several patients were still responsive at LORnoxious. This resulted in larger transitions when the transition between the response and the unresponsive state, for example, from OAA/S 3 to OAA/S 0 without passing through OAA/S levels 2 or 1, causing an increase in averages. This finding obviously confirms that remifentanil blocks the response to pain, without causing LORverbal and LORlash but proves also that it might be difficult to quantify the dependence of the electronic independent variables such as BIS and AAI to opiates because the definitive standard to which they were compared changed its behavior through the groups. This fact highlights that care should be taken when validating loss of response to different stimuli by using OAA/S score or eyelash reflex when opioids are administered. However, because no better validated clinical scoring techniques exist at this moment, we have to accept this possible bias.

Skepticism still exists in the literature regarding the accuracy of cerebrally derived parameters to measure anesthetic depth when adding opiates. Several articles suggested a weaker correlation between BIS and CePROP in the presence of opiates. Although they state that this might reveal the importance of an analgesic component on the efficacy of depth-of-anesthesia electronic monitors, these studies did not apply specific statistical techniques such as prediction probability or specificity/sensitivity calculations. Also, one should clearly differentiate the phenomenon of the influence of opiates on the changes in cutoff values for LORverbal, LORlash and LORnoxious and the overall change in accuracy of these anesthetic depth independent variables when adding opiates.

The prediction probability, Pk, provides a good alternative to investigate the overall relative performance of the different independent variables to measure the hypnotic component of anesthesia and loss of responsiveness to different stimuli. For BIS, Lysakowski et al. found no decrease in Pk when adding clinical dosages of fentanyl, alfentanil, remifentanil, or sufentanil versus placebo during propofol administration. Iselin-Chaves et al. studied the influence of different dosages of alfentanil on the accuracy of both BIS and nonprocessed MLAEP during propofol administration and found no differences in Pk between groups. In their study, BIS performed better than MLAEP. In our study, as indicated in table 4, the performance results indicate that both BIS and AAI are reliable independent variables for assessing the level of OAA/S and LORlash and did not decrease by the addition of remifentanil. Before being able to conclude that the ability to measure the hypnotic component remained intact, a Pk analysis with the three groups pooled was needed. As seen in table 4, the Pk remained similar to the group-based Pk for both OAA/S and LORlash, indicating an overall equal accuracy of the independent variables with and without the addition of opiates.

At the level of significance used in our study, BIS and AAI were found to be comparable in performance to the estimated steady state propofol concentration, CePROP. When drug effect-site concentrations are known, one might argue that it only make sense to measure BIS or AAI if the combined information offers more accuracy in measuring depth of anesthesia than a single measure alone. Vice versa, it might be asked whether drug effect-site concentrations offer additional information for the clinician when BIS or AAI is used. However, it must be stated that electroencephalographic and MLAEP monitors can be attached to the patient at any random time, whereas calculation of drug effect-site concentration requires knowledge of the complete administration history. As none of the independent variables in the present study gave a Pk larger than 0.9, when pooling the data from the three groups into one, it could be interesting to explore whether a new independent variable defined as a composite of the electronic independent variables and the anesthetics concentrations would produce a larger Pk when predicting the OAA/S. Therefore, we calculated the Pk value for the combined information from both BIS + CePROP + CeREMI and AAI + CePROP + CeREMI using ordinal logistic regression to develop a predicted OAA/S score (see Appendix). For both combinations, the overall prediction probability for predicting the levels of OAA/S increased (or tended to increase) compared to the pooled Pk data.

Because the Pk concept was developed to generalize nonparametric receiver operating characteristic curves area to polytomous ordinal patient state, we thought it was interesting to observe some specific sensitivity/specificity characteristics for BIS, AAI, and CePROP. Statistically speaking, the most frequently applied combined sensitivity/specificity lies at the elbow of the receiver operating characteristic curve, where the sum of both sensitivity and specificity is the highest. However, as defined by J. Drummond recently, a depth of anesthe-

Table 7. Averaged OAA/S Score at LORlash and LORnoxious

<table>
<thead>
<tr>
<th>Group</th>
<th>OAA/S Level at LORlash</th>
<th>OAA/S Level at LORnoxious</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ng/ml Remifentanil</td>
<td>3.6</td>
<td>0.87</td>
</tr>
<tr>
<td>2 ng/ml Remifentanil</td>
<td>4.2</td>
<td>3.4</td>
</tr>
<tr>
<td>4 ng/ml Remifentanil</td>
<td>4.1</td>
<td>4.0</td>
</tr>
</tbody>
</table>

* Average of 15 patients.

LORlash = loss of eyelash reflex; LORnoxious = loss of response to electrical tetanic stimulus; OAA/S = Observer's Assessment of Alertness/Sedation scale.
A specific numeric threshold (cutoff value) that can be interpreted to mean “not aware.” Therefore, we observed the influence of remifentanil on the cutoff value for LORverbal and LORlash and the corresponding sensitivity/specificity profiles for BIS, AAI, and CePROP at the two important points of the receiver operating characteristic curves: (1) highest sum of sensitivity and specificity and (2) level of 100% sensitivity. For both BIS and AAI, the cutoff values at LORverbal and LORlash revealed the highest sum of sensitivity/specificity clearly increased in a dose-dependent manner during remifentanil addition. However, the changes in the values of the highest sum were minimal and similar between the different anesthetic depth–independent variables, as observed in table 5. This means that the addition of remifentanil did not decrease the overall sensitivity/specificity profiles of the tested independent variables when observing the cutoff value at the level of maximum combined sensitivity and specificity. As shown in table 6, the cutoff values detecting LORverbal and LORlash at the 100% sensitivity level were influenced by the addition of remifentanil. For BIS, a trend toward higher cutoff values at the 100% sensitivity level was observed at LORverbal and LORlash. For AAI, the cutoff value only increased in the 4 ng/ml remifentanil group when observing LORverbal and slightly increased between the 0 and 2 ng/ml remifentanil groups, between the 0 and 4 ng/ml remifentanil groups, and when observing LORlash. The CePROP cutoff values decreased at this 100% sensitivity level when adding remifentanil. For all independent variables, the influence of remifentanil on the specificity at a level of 100% sensitivity was calculated. For the AAI, a clear decrease in specificity was observed in the 2 ng/ml remifentanil group when observing LORverbal and a decrease between the 2 and 4 ng/ml remifentanil groups was seen when observing LORlash. This might be caused by the severity of the used statistical criterion, where one outlier might cause a dramatic decrease in specificity at the 100% sensitivity level. No decrease in specificity was observed between groups for BIS when observing LORverbal and LORlash. Inconsistencies are found in the data for CePROP when comparing the specificity results obtained by LORverbal with these obtained by using LORlash. Better sensitivity/specificity profiles were obtained when the information from drug concentrations and BIS or AAI were used together.

This study also tested the influence of opiates on the performance of the independent variables to predict LORnoxious. The supramaximal tetanic stimulus used in this study was previously used by others as a substitute for conventional forms of stimulation in humans. In previous work, when only using propofol without opiates, we observed that measures from the cerebral cortex such as BIS and AAI were poor predictors for LORnoxious. The overall accuracy for LORnoxious tended to be lower for all independent variables compared with their hypnotic prediction accuracy. The Pk values were lower, however, because of the larger SE, not statistically significant at the level of significance used in our study. Sensitivity/specificity profiles for all independent variables in the three groups were calculated at the two classic points. As seen in table 5, the values of the highest sum of sensitivity and specificity to LORnoxious were lower compared with the values for the detection of LORverbal and LORlash. Also, the cutoff values at the point of highest sensitivity/specificity were clearly influenced by the addition of remifentanil in a dose-dependent manner. At the level of 100% sensitivity to detect LORnoxious, a very low specificity and a clear influence of remifentanil on the cutoff values was observed (table 6).

A strong interaction between propofol and remifentanil is observed at the ED50 and ED95 levels when studying the analgesic component of this interaction at LORnoxious, as plotted in figure 5 and table 3. This interaction is more pronounced than the hypnotic interaction between remifentanil and propofol, resulting in the fact that in this “staircase study,” patients in the 2 and 4 ng/ml remifentanil groups might reach LORnoxious before reaching LORverbal and LORlash, as shown in table 7. Similar findings were observed for fentanyl and propofol by Smith et al. and by Katoh et al. for fentanyl and sevoflurane.

In conclusion, we found that although BIS, AAI, and CePROP were increased by remifentanil during propofol administration, their ability to detect OAA/S and LORlash remained accurate. Improved performance is obtained when BIS and AAI are measured in conjunction with drug effect-site concentrations. Remifentanil decreases the ability of these independent variables to detect LORnoxious.

The authors thank Paul Manberg, Ph.D. (Vice-President, Aspect Medical Systems, Newton, Illinois), Bob Butterfield (Research Fellow, Alaris Medical Systems, San Diego, California), Bernardo Rodriguez and Pablo Martinez (Research Fellows, Dammeter A/S, Odense, Denmark), and Geert Byttebier and Mieke De Bosschere (Statisticians, General Biometric Services and Consulting, Gent, Belgium) for their help during the preparation of this manuscript.

References

2. Sleigh JW, Donovan J: Comparison of Bispectral Index, 95% spectral edge frequency, and approximate entropy of the EEG with changes in heart rate variability during induction of general anesthesia. Br J Anaesth 1999; 82:666–71


14. 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 during propofol infusion with or without fentanyl. Eur J Anaesthesiol 1999; 16:47–52


16. Inselin-Chaves IA, El Moalem HE, Gan Tj, Ginsberg B, Glass PS. Changes in the auditory evoked potentials and the Bispectral Index following propofol or propofol and alfentanil. ANESTHESIOLOGY 2000; 92:1300 –10


21. Minto CF, Schneider TW, Shafer SL. Pharmacokinetics and pharmacodynamics of remifentanil. II. Model application. ANESTHESIOLOGY 1997; 86:24–35


Appendix: Definition of an OAA/S Predictor Based on Combinations of AAI/BIS, Propofol, and Remifentanil Effect Concentrations

Methods

Two composite indexes were defined to study whether the combined information BIS + CePROP + CeREMI or AAI + CePROP + CeREMI offers more accuracy in loss of response to hypnotic stimuli than a single measure alone (BIS, AAI, CeREMI, and CePROP). The composite index was defined as an OAA/S predictor based on combined information from the data set BIS + CePROP + CeREMI or AAI + CePROP + CeREMI. The method of choice for designing a composite index was the ordinal logistic regression. The ordinal logistic regression has the advantage, as compared with multiple linear regression, that the output variable is ordinal. An ordinal variable is a categorical variable that has two or more levels of natural ordering, such as awake, drowsy, asleep, and deep asleep. This is suitable for the kind of data from this study because the OAA/S scale is ordinal and not linear. Two models were proposed: model A, where the input variables are BIS, CePROP, and CeREMI; and model B, where the input variables are AAI, CePROP, and CeREMI. The composite index is the output variable that in both models is a prediction of the OAA/S as shown in figures 6 and 7. The composite index is therefore ordinal in the 0–5 range as the OAA/S.

An ordinal logistic regression model was calculated by using SPSS software (SPSS Inc., Chicago, IL), which also provides a more detailed description of the method. The model is composed of constants, $\theta_i$, factors, $\beta_n$, and a link function, here the “logit.” The constants are associated with the ith event, which in the current study refers to the level of OAA/S. The factors are obtained from each variable (input), which here are BIS, AAI, CePROP, and CeREMI at every jth sample in time. Hence, the two models are defined in equations 1 and 2

$$\text{Input variables}$$

\begin{align*}
\text{BIS} & \\
\text{CePROP} & \\
\text{CeREMI} & \\
\text{Output variable} & \\
\text{OAAS predictor} & \\
\end{align*}

Fig. 6. Model A for calculating the Observer’s Assessment of Alertness/Sedation Scale (OAA/S) predictor from the combined information Bispectral Index (BIS) + propofol effect-site concentration (CePROP) + remifentanil effect-site concentration (CeREMI).
Fig. 7. Model B for calculating the Observer’s Assessment of Alertness/Sedation Scale (OAA/S) predictor from the combined information A-Line ARX index (AAI) + propofol effect-site concentration (CePROP) + remifentanil effect-site concentration (CeREMI).

\[
g(OAA_{ij}) = \theta_i + \beta_{remi}CeRemi_i + \beta_{CePROP}CePROP_i + \beta_{BIS_i} \tag{1}
\]

The accumulated probability of being at or lower than the ith level of OAA/S at the jth sample in time, \( g_{ij} \), is now calculated as

\[
g_{ij} = \frac{e^{g(OAA_{ij})}}{1 + e^{g(OAA_{ij})}} \tag{3}
\]

Example: The following state has been recorded: remifentanil = 0 ng/ml, propofol = 4 µg/ml, AAI = 25. When using these data as input to model B, the corresponding probabilities for OAA/S 0, 1, 2, 3, 4, and 5 are 0.7944, 0.8867, 0.9620, 0.9935, 0.9998, and 1. The model output is always the OAA/S level, where the accumulated probability exceeds 0.67. In this example, the predicted OAA/S is 0 because the first probability (0.7944) is larger than 0.67.

**Results**

Tables 8 and 9 show the parameters of the ordinal logistic regression for models A and B, with the corresponding levels of significance for each constant and factor as well as the P value for the whole model. A P value less than 0.05 means that the parameter is significant in the model. Both models were statistically significant. The OAA/S values predicted by the two models were compared with the measured OAA/S and evaluated using the \( P_k \) analysis, shown in table 10.

### Table 8. Ordinal Logistic Regression of Model A (BIS)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Value</th>
<th>SD</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>-3.718</td>
<td>1.851</td>
<td>-2.01</td>
<td>0.045</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>-2.995</td>
<td>1.850</td>
<td>-1.62</td>
<td>0.105</td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>-1.832</td>
<td>1.853</td>
<td>-0.99</td>
<td>0.323</td>
</tr>
<tr>
<td>( \theta_4 )</td>
<td>-0.109</td>
<td>1.868</td>
<td>-0.06</td>
<td>0.954</td>
</tr>
<tr>
<td>( \theta_5 )</td>
<td>3.228</td>
<td>1.896</td>
<td>1.70</td>
<td>0.089</td>
</tr>
<tr>
<td>( \beta_{remi} )</td>
<td>0.9548</td>
<td>0.1163</td>
<td>8.21</td>
<td>0.000</td>
</tr>
<tr>
<td>( \beta_{prop} )</td>
<td>2.6277</td>
<td>0.3395</td>
<td>7.74</td>
<td>0.000</td>
</tr>
<tr>
<td>( \beta_{bis} )</td>
<td>-0.10034</td>
<td>0.01779</td>
<td>-5.64</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Test that all slopes are zero; P value < 0.001.

\( \beta_{remi} = \) constants of the ordinal logistic regression model used in the prediction of the OAA/S levels 1 to 5; \( \beta_{prop} = \) coefficient of the ordinal logistic regression model multiplied with the propofol (prop) effect-site concentration in the prediction equation; \( \beta_{bis} = \) coefficient of the ordinal logistic regression model multiplied with the BIS in the prediction equation; SD = standard deviation.

### Table 9. Ordinal Logistic Regression of Model B (AAI)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Value</th>
<th>SD</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>-9.265</td>
<td>1.117</td>
<td>-8.29</td>
<td>0.000</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>-8.559</td>
<td>1.094</td>
<td>-7.83</td>
<td>0.000</td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>-7.386</td>
<td>1.057</td>
<td>-6.99</td>
<td>0.000</td>
</tr>
<tr>
<td>( \theta_4 )</td>
<td>-5.583</td>
<td>1.009</td>
<td>-5.53</td>
<td>0.000</td>
</tr>
<tr>
<td>( \theta_5 )</td>
<td>-1.9793</td>
<td>0.9563</td>
<td>-2.07</td>
<td>0.038</td>
</tr>
<tr>
<td>( \beta_{remi} )</td>
<td>1.0143</td>
<td>0.1194</td>
<td>8.50</td>
<td>0.000</td>
</tr>
<tr>
<td>( \beta_{prop} )</td>
<td>3.0479</td>
<td>0.3091</td>
<td>9.86</td>
<td>0.000</td>
</tr>
<tr>
<td>( \beta_{AAI} )</td>
<td>-0.064</td>
<td>0.0097</td>
<td>-6.61</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Test that all slopes are zero; P value < 0.001.

\( \beta_{remi} = \) constants of the ordinal logistic regression model used in the prediction of the OAA/S levels 1 to 5; \( \beta_{prop} = \) coefficient of the ordinal logistic regression model multiplied with the remifentanil (remi) effect-site concentration in the prediction equation; \( \beta_{prop} = \) coefficient of the ordinal logistic regression model multiplied with the propofol (prop) effect-site concentration in the prediction equation; \( \beta_{AAI} = \) coefficient of the ordinal logistic regression model multiplied with the AAI in the prediction equation; SD = standard deviation.

### Table 10. Prediction Probability (\( P_k \) of the Indices)

<table>
<thead>
<tr>
<th>Model</th>
<th>( P_k ) (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A (BIS + CePROP + CeREMI)</td>
<td>0.91 (0.01)</td>
</tr>
<tr>
<td>Model B (AAI + CePROP + CeREMI)</td>
<td>0.93 (0.01)</td>
</tr>
</tbody>
</table>

AAI = A-Line ARX index; BIS = Bispectral Index; CePROP = calculated propofol effect-site concentration; CeREMI = calculated remifentanil effect-site concentration.