An Evaluation of Remifentanil-Sevoflurane Response Surface Models in Patients Emerging from Anesthesia: Model Improvement Using Effect-Site Sevoflurane Concentrations

Ken B. Johnson, MD*
Noah D. Syroid, MS*
Dhanesh K. Gupta, MD†
Sandeep C. Manyam, PhD‡
Nathan L. Pace, MD*
Cris D. LaPierre, BS*
Talmage D. Egan, MD*
Julia L. White, RN*
Diane Tyler, RN*
Dwayne R. Westenskow, PhD*

INTRODUCTION: We previously reported models that characterized the synergistic interaction between remifentanil and sevoflurane in blunting responses to verbal and painful stimuli. This preliminary study evaluated the ability of these models to predict a return of responsiveness during emergence from anesthesia and a response to tibial pressure when patients required analgesics in the recovery room. We hypothesized that model predictions would be consistent with observed responses. We also hypothesized that under non-steady-state conditions, accounting for the lag time between sevoflurane effect-site concentration (Ce) and end-tidal (ET) concentration would improve predictions.

METHODS: Twenty patients received a sevoflurane, remifentanil, and fentanyl anesthetic. Two model predictions of responsiveness were recorded at emergence: an ET-based and a Ce-based prediction. Similarly, 2 predictions of a response to noxious stimuli were recorded when patients first required analgesics in the recovery room. Model predictions were compared with observations with graphical and temporal analyses.

RESULTS: While patients were anesthetized, model predictions indicated a high likelihood that patients would be unresponsive (>99%). However, after termination of the anesthetic, models exhibited a wide range of predictions at emergence (1%–97%). Although wide, the Ce-based predictions of responsiveness were better distributed over a percentage ranking of observations than the ET-based predictions. For the ET-based model, 45% of the patients awoke within 2 min of the 50% model predicted probability of unresponsiveness and 65% awoke within 4 min. For the Ce-based model, 45% of the patients awoke within 1 min of the 50% model predicted probability of unresponsiveness and 85% awoke within 3.2 min. Predictions of a response to a painful stimulus in the recovery room were similar for the Ce- and ET-based models.

DISCUSSION: Results confirmed, in part, our study hypothesis; accounting for the lag time between Ce and ET sevoflurane concentrations improved model predictions of responsiveness but had no effect on predicting a response to a noxious stimulus in the recovery room. These models may be useful in predicting events of clinical interest but large-scale evaluations with numerous patients are needed to better characterize model performance.

Accepted for publication April 23, 2009.
Supported in part by a research grant from General Electric Healthcare and NIH Health General Medicine SBIR Phase I grant no. 2R44GM06615–02.
Please see supplementary material available at www.anesthesia-analgesia.org.
Dwayne R. Westenskow is the Section Editor of Technology, Computing, and Simulation for the Journal. This article was handled by Tony Gin, Section Editor of Anesthetic Clinical Pharmacology, and Dr. Westenskow was not involved in any way with the editorial process or decision.

Copyright © 2009 International Anesthesia Research Society DOI: 10.1213/ANE.0b013e3181afe31c

Populatation models characterizing the pharmacodynamic interactions between hypnotics and opioids, such as propofol and remifentanil, can be applied clinically to predict drug behavior in individual patients.1 We previously reported a series of response surface models that characterized the synergistic interaction between predicted remifentanil effect-site
concentrations (Ce’s) and measured end-tidal (ET) sevoflurane levels in blunting responses to selected verbal and painful stimuli. The aim of this study was to evaluate the ability of these models to predict the return of responsiveness (ROR) in patients undergoing elective surgery during emergence from anesthesia and a response to tibial pressure when patients required analgesics in the recovery room. We hypothesized that model predictions would be consistent with observations. One limitation to these models was that measured ET sevoflurane concentrations do not reflect brain concentrations during non-steady-state conditions, such as emergence from anesthesia. We also hypothesized that response surface model predictions would improve if we account for the lag time between brain and ET sevoflurane concentrations.

**METHODS**

**Section I: Response Surface Model Development**

We revised previously reported sevoflurane-remifentanil pharmacodynamic interaction models. We modified the models to account for the lag time between changes in ET and brain sevoflurane concentrations during non-steady-state conditions. A physiologic model was used to predict brain concentrations during non-steady-state conditions. In response to a step change in inspired sevoflurane, predicted brain concentrations were slower to change than ET measurements. We also revised the models to account for the effect of altitude (the original model was based on data collected at 5000 feet above sea level).

Two sets of revised response surface models were constructed: one set using predicted remifentanil Ce’s and measured ET sevoflurane levels, and a second set using predicted remifentanil Ce’s and predicted sevoflurane effect-site levels. Each set consisted of 2 models: a probability of unresponsiveness model (i.e., lack of response to loud verbal and nonpainful tactile stimuli based on the Observer’s Assessment of Alertness/Sedation (OAA/S) scale (presented in Table 1) and a lack of response to 30 pounds per square inch (PSI) of anterior tibial pressure (pressure algometry) model. Thirty PSI was used in lieu of 50 PSI as in our prior work because our recent experience suggests that 30 PSI is a better surrogate for the painful stimuli encountered in the recovery room. In our recent work, we found that 50 PSI was too strong a stimulus, when compared with stimuli typically encountered in the recovery room.

A pharmacokinetic model for inhaled anesthetics was used to account for the lag between sevoflurane Ce’s and ET concentrations. Model development is presented in the Appendix (see Supplemental Digital Content 1, http://links.lww.com/AA/A13).

The response surface models were constructed using a Greco model structure (Eq. 1). This structure differs from the Logit model structure we originally used to characterize interactions between remifentanil and sevoflurane. The Greco model structure was used to build a revised set of models to be consistent with our prior work with evaluating propofol-remifentanil response surface models in surgical patients. Models were built using a naïve pooled technique.

\[
\text{Effect} = \frac{E_{\max} \cdot \left[ \frac{C_r}{C_{50r}} + \frac{C_p}{C_{50p}} + \alpha \cdot \left( \frac{C_r}{C_{50r}} \cdot \frac{C_p}{C_{50p}} \right)^n \right]}{\left( \frac{C_r}{C_{50r}} + \frac{C_p}{C_{50p}} + \alpha \cdot \left( \frac{C_r}{C_{50r}} \cdot \frac{C_p}{C_{50p}} \right)^n \right) + 1}
\]

\( E_{\max} \) is the maximal effect (i.e., no response to pressure algometry), \( C_{50r} \) (sevoflurane) and \( C_{50p} \) (remifentanil) are the concentrations that produce 50% of the maximal effect when administered individually, \( n \) is the slope of the pharmacodynamic response curve, and \( \alpha \) is the interaction between sevoflurane and remifentanil. Effect ranged from 0 (0% probability of no response) to 1 (100% probability of no response). Model parameters were identified by an iterative process in which the log likelihood between the observations and the model predictions was maximized (Matlab, The MathWorks, Natick, MA). Coefficients of variation were estimated for each parameter using a bootstrap method with 10,000 iterations.

An assessment of how well the revised models fit observations in our volunteer study was made by calculating the percentage of predictions more than 0.5 (response or no response) that agreed with observations. Responses were compared with model predictions using a Spearman rank correlation (\( \rho \)). A two-tailed unpaired \( t \)-test was used to determine whether the Spearman \( \rho \) was significantly different from 0. The null hypothesis was that the revised models did not correlate with observations.

**Section II: Evaluation of Response Surface Models in Patients Undergoing Elective Surgery**

**Patient Selection**

After IRB approval at the University of Utah, informed consent was obtained from 20 patients presenting for elective surgery. Twelve male and 8 female
subjects participated. Patients with a history of ongoing opioid consumption were excluded. Age, weight, and height were recorded for each patient.

**Patient Monitoring**
Each patient was instrumented with a pulse oximeter, noninvasive arterial blood pressure cuff set to cycle every 5 min, a 5-lead electrocardiogram, and an oral or nasal temperature probe. Inspired and expired sevoflurane concentrations were continuously monitored (AS/3 Anesthesia Monitor, Datex-Ohmeda, Helsinki, Finland) and stored every 5 s using a computerized data acquisition system (S/5 Collect, Datex-Ohmeda).

**Experimental Protocol**
Midazolam 12.5 μg/kg was administered IV in the preoperative holding area 10 min before patients were taken to the operating room. Induction consisted of age-adjusted infusions of remifentanil (range 0.1–0.5 μg/kg⋅min⁻¹) and propofol (range 75–300 μg/kg⋅min⁻¹) designed to achieve a loss of response to laryngoscopy within 7 min (Medfusion 2010I or 3010I, Vol. X, No. X, XXX 2009 © 2009 International Anesthesia Research Society) and remifentanil Ce's. Fentanyl Ce's were converted to equivalent remifentanil Ce's using a relative potency of remifentanil to fentanyl = 1:1.2. Measured ET sevoflurane concentrations were used to predict sevoflurane Ce's. Pharmacokinetic models for propofol and midazolam were used to estimate residual propofol and midazolam Ce's at the time of ROR.

**Evaluation of Response Surface Model Predictions**
Model predictions were evaluated at 2 events: ROR after surgery and the first dose of analgesic (fentanyl) administered in the recovery room. For each model (a model based on ET sevoflurane measurements and a model based on estimated effect-site sevoflurane concentrations), a prediction of OAA/S = 1 ranging from 0% to 100% was made every 5 s from once the anesthetic was terminated to 10 min after the time of each patient's ROR. Model predictions were compared with observations with graphical and temporal analyses.

**Graphical Analysis**
Predictions from each model were compared graphically on 2 plots: a prediction versus time plot and a plot of the remifentanil-sevoflurane concentrations at the time of emergence superimposed on a topographical representation of the response surface model predictions of OAA/S = 1. The topographical plot included the 5%, 50%, and 95% isoboles. Isoboles were defined as the set of remifentanil and sevoflurane concentration pairs that produced the same probability of effect. Mean model predictions recorded at ROR (OAA/S > 1) for the measured ET sevoflurane and estimated sevoflurane effect-site-based models were compared with an unpaired two-tailed t-test.

**Empirical Cumulative Distribution**
Model predictions from each patient were sorted according to increasing probability. A percentage value was assigned to each patient as a percentage of all patients according to increasing probabilities. The patient population percentage was plotted versus the sorted model predictions. A uniform distribution of model predictions across patient percentage values from 0% to 100% was considered a good model fit.

**Temporal Analysis**
The time from the 50% model prediction for unresponsiveness to the time of the observed ROR was calculated for each patient. A negative, zero, or positive time difference indicated that the observed ROR occurred before, exactly at, or after the 50% probability of unresponsiveness. Time differences were reported as mean ± sp. The time differences between the measured ET sevoflurane and estimated sevoflurane effect-site-based models were compared with an unpaired two-tailed t-test.

Similarly, 2 models of lack of response to 30 PSI of anterior tibia pressure algometry were evaluated at the time patients received their first dose of fentanyl in the recovery room: a model based on ET sevoflurane measurements and a model based on estimated effect-site sevoflurane concentrations. For each model, a prediction of lack of response to 30 PSI of anterior tibial pressure ranging from 0% to 100% was made.
when patients received analgesic therapy (fentanyl) in the recovery room.

**RESULTS**

**Section I: Response Surface Model Development for Application in Real Time**

Sevoflurane-remifentanil interaction model parameters fit to responses previously recorded in volunteers are presented in Table 2.\(^2\) Models include the probability of unresponsiveness (\(OAA/S = 1\)) and lack of response to 30 PSI of anterior tibial pressure. Two sets of models are presented: one using sevoflurane effect-site concentrations and the other using end-tidal concentrations. Model parameters were estimated using a Greco model structure. Numbers in parentheses represent coefficients of variation.

### Table 2. Sevoflurane-Remifentanil Interaction Model Parameters Fit to Responses Recorded in Volunteers

<table>
<thead>
<tr>
<th>Probability of unresponsiveness ((OAA/S^3 = 1))</th>
<th>Lack of response to anterior tibial pressure (30 PSI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effect site (%)</strong></td>
<td><strong>End tidal</strong></td>
</tr>
<tr>
<td>C(_{50}) Sevoflurane (vol%)</td>
<td>0.74 (8%)</td>
</tr>
<tr>
<td>C(_{50}) Remifentanil (ng/mL)</td>
<td>50.9 (28%)</td>
</tr>
<tr>
<td>Alpha (interaction)</td>
<td>9.4 (56%)</td>
</tr>
<tr>
<td>N (slope)</td>
<td>5.2 (66%)</td>
</tr>
<tr>
<td><strong>Model fit</strong></td>
<td></td>
</tr>
<tr>
<td>Percentage of predictions &gt;0.5 that matched observed responses</td>
<td>92%</td>
</tr>
<tr>
<td>Spearman (\rho) and (P)</td>
<td>0.83, &lt;0.001</td>
</tr>
</tbody>
</table>

Models include probability of unresponsiveness and lack of response to 30 pounds per square inch (PSI) of anterior tibial pressure. Two sets of models are presented: one using sevoflurane effect-site concentrations and the other using end-tidal concentrations. Model parameters were estimated using a Greco model structure. Numbers in parentheses represent coefficients of variation.

\(OAA/S =\) Observer’s Assessment of Alertness/Sedation.

Evaluation of Response Surface Model Predictions

Figure 1 shows model predictions of \(OAA/S = 1\) over time during emergence from anesthesia for each patient. The rapid up and down changes in ET-based model predictions in selected patients (i.e., Patient 10, 12, 14, etc.) were attributable to variations in respiratory function during emergence from anesthesia. The ET concentration-based model consistently predicted a high probability of ROR before the observed ROR. The Ce-based model, by contrast, predicted high probabilities of ROR both before and after the observed ROR.

Figure 2a shows the sevoflurane-remifentanil concentration pairs at the time of emergence superimposed on a topographical representation of the model prediction a \(OAA/S = 1\). Concentrations are nearly equally distributed above and below the 50% isobole (8 below and 12 above) for the sevoflurane effect-site model and predominantly below the 50% isobole (19 below and 1 above) for the ET model. At the time patients responded to voice command or prodding, the mean ET sevoflurane concentration was 0.13 ± 0.13 vol\%, predicted sevoflurane Ce was 0.44 ± 0.17 vol\%, and remifentanil equivalent Ce was 3.5 ± 1.7 ng/mL. The average predicted probability of no response was 53% ± 37% and 9% ± 19% for the effect-site and ET models, respectively (\(P < 0.001\)).

The empirical cumulative distribution plot illustrates the poor distribution of predictions across the patient percentage values for the ET model (Fig. 3a). For example, at the 20th patient percentile, ideally 4 of the model predictions should be between 0% and 20%, and the remainder of the predictions should be above 20%. At the 20th percentile for the ET model, 16 patients were between 0% and 20%. By contrast, the effect-site model had a more uniform distribution of predictions across patient percentage values, but the distribution was weighted at low and high predictions. For example, at the 50th patient percentile, 7 of
the model predictions were below 30%, and 10 of the model predictions were above 70%.

The mean time from the termination of the anesthetic until patients emerged from anesthesia was 5.3 ± 6.1 min. Patients emerged from anesthesia 0.9 ± 4.3 min after the effect-site-based model predicted a 50% probability of unresponsiveness and 4.1 ± 4.9 min after the ET-based model predicted a 50% probability of unresponsiveness (P = 0.02, Fig. 4).

Figure 1. Model predictions of return of responsiveness during emergence from anesthesia for each patient. The solid and dotted black lines show model predictions of Observer’s Assessment of Alertness/Sedation (OAA/S) = 1 for the effect-site- and the end-tidal-based response surface models, respectively. The gray vertical line represents the time at which each patient became responsive (OAA/S > 1). The x-axis represents time ranging from 10 min before emergence to 10 min after. The solid gray horizontal line represents the 50% model probability of OAA/S = 1. NR = no response.
Figure 2b shows the 5%, 50%, and 95% probability isoboles for the 2 models and the sevoflurane-
remifentanil equivalent concentration pairs when patients received their first dose of fentanyl during the first 30 min after surgery. Seven of the 20 patients did not require fentanyl during this time period. Concentration pairs are distributed predominantly above the 50% isobole for the effect-site-based and ET-based models (3 below and 10 above). At the time of first fentanyl request for the management of postoperative pain, the average ET sevoflurane was 0.00\%\,vol\%, predicted sevoflurane effect site was 0.02\%\,vol\%, and remifentanil equivalent Ce was 1.9 \pm 1.3 \,ng/mL. The mean probabilities of no response to 30 PSI of anterior tibial pressure at the time patients required additional fentanyl were 61\% \pm 30\% and 59% \pm 30\% for the effect-site and ET models, respectively ($P = 0.893$).

The empirical cumulative distribution plot illustrates a similar distribution of predictions across patient percentage values for both the ET and effect-site models (Fig. 3b). At the 50th patient percentile, 6 of the model predictions were below 50%, and 7 of the model predictions were above 50% for both models.

**DISCUSSION**

**Section I: Response Surface Model Development**

Revised models using the Greco model structure fit responses from volunteers reasonably well according to the comparison of prediction probabilities to observed responses. The Spearman rank correlation coefficients for the probability of unresponsiveness models had stronger correlations with observations than the lack of response to pressure algometry models. There was no appreciable difference in model fit between the sevoflurane effect-site- and ET-based models. This suggests that the original experiment was conducted at near steady-state conditions because a parameter to adjust for disequilibrium between ET and effect-site sevoflurane concentrations did not appreciably alter the interaction parameters.

In comparison with prior studies, there is large variability between our work and previously published remifentanil $C_{50}$'s for predictions of OAA/S = 1.\(^2\,17\) One explanation may be that the remifentanil $C_{50}$ Greco model parameter is estimated from a data set where the majority of observations were made below the remifentanil $C_{50}$ of 50.9 ng/mL.
Section II: Evaluation of Response Surface Models in Patients Undergoing Elective Surgery

In this study, we conducted a preliminary evaluation of previously developed models of unresponsiveness and lack of response to a moderately painful stimulus (tibial pressure) in patients undergoing elective surgery. One study aim was to evaluate how well these models predicted emergence from anesthesia and the need for additional opioid in the recovery room.

While patients were anesthetized, model predictions indicated a high likelihood that patients would be unresponsive. With the probability of unresponsiveness models, predictions more than 99% were consistent with observations (no patients demonstrated responsiveness). However, after termination of the anesthetic, model predictions of unresponsiveness rapidly decreased (Fig. 1), and individual patients emerged from anesthesia over a wide spectrum of model predictions ranging from 1% to 97%. It is difficult to comment on the ability of these models to discriminate between awake or asleep during time intervals that are of clinical interest and how well the model predictions are calibrated to observations. The statistical tools that would be used to help answer these questions require a study with substantially larger sample sizes.

A second study aim was to evaluate the value of accounting for the lag time between ET and brain sevoflurane concentrations during non-steady-state conditions on model predictions. Model predictions using effect-site sevoflurane concentrations more adequately predicted ROR in patients emerging from anesthesia than model predictions using ET sevoflurane concentrations. As illustrated in Figures 1 and 4, the model predictions of OAA/S = 1 using ET sevoflurane levels typically predicted ROR before emergence. This model misspecification is likely due to the biophase between ET and effect-site sevoflurane concentrations during emergence. This was confirmed by the differences in mean predictions of emergence and mean time intervals between observed and predicted emergence between the 2 models. The improvement was also visualized in the topographical representation of the model predictions of OAA/S = 1 (Fig. 2a). Ideal model predictions would be equally distributed at about the 50% isobole. The concentration pair predictions are more equally distributed around the 50% isobole for the effect-site model but not the end-site model.

The benefit of using Ce’s, however, was not as useful in model predictions in the recovery room. The ET and effect-site model predictions of a lack of response to tibial pressure at the time when additional analgesic was required were similar (Fig. 2b). Mean model probabilities were skewed above the 50% isobole (61% and 59%) predicting a lack of response to anterior tibial pressure in a majority of patients. This suggests that using 30 PSI may be a slightly less painful surrogate of pain encountered in the recovery room (an ideal distribution would be at about the 50% isobole). Because sevoflurane levels were negligible, both models collapsed to a single drug (remifentanil equivalents) pharmacodynamic model. This is the most likely explanation for similar predictions between models (i.e., the sevoflurane concentrations are so low that they are nearly irrelevant in the recovery room in terms of analgesia).

These results in the recovery room, although promising, are difficult to interpret given the narrow spectrum of patients evaluated. We did not control for duration, type, or extent of surgery, use of local infiltration into the surgical site, and patient age. These factors and others may play a role in how patients respond to assessments of perceived pain in the recovery room.

Limitations

Several assumptions were made in implementing the Lerou model to predict sevoflurane Ce’s. Cardiac output was normalized to weight but assumed to be constant (Appendix, see Supplemental Digital Content 1, http://links.lww.com/AA/A13) and did not account for changes in cardiac output with induction of anesthesia, blood loss, etc. Cerebral blood flow was also assumed constant (16% of total cardiac output) and did not account for changes in Paco₂. This could be especially important during emergence from anesthesia when changes in cardiac output and Paco₂ with hyperventilation or permissive apnea could potentially change the time to ROR.18 Future work directed at continuously updating model estimates of cerebral...
blood flow may improve accuracy of sevoflurane effect-site predictions.

In addition, we used the OAA/S score to build a response surface that predicts the probability of unresponsiveness. In the volunteer study, the OAA/S assessments were done in the absence of ongoing pain and without an endotracheal tube in place. We used this model to predict emergence from general anesthesia in patients whose tracheas were intubated after surgeries associated with mild to moderate surgical pain. No volunteer study can fully emulate the complexities of the clinical environment. The differences between the volunteer study on which the models are based and the clinical environment in which the models are applied likely impacts model performance to some degree.

Premedication with midazolam and induction with propofol may have prolonged emergence beyond model predictions, but their contribution was most likely small. For example, midazolam levels were just below the $C_{50}$ for a Ramsay score of 2 (patient cooperative, oriented, and tranquil—5.7 ng/mL) and well below the $C_{50}$ for a Ramsay score of 3 (patient sedated but responds to commands—71 ng/mL). Propofol levels also were well below the $C_{50}$ for loss of responsiveness (2.2 $\mu$g/mL). making its contribution to prolonging emergence negligible.

In conclusion, predicted Ce’s are frequently used to guide the delivery of IV anesthetics using target-controlled delivery systems. Extension of this concept to inhaled anesthetics and to drug-drug interaction models may have value in guiding the delivery of anesthetics in combination by predicting the likelihood of adequate sedation, analgesia, and time to emergence. In fact, clinical pharmacology display systems that apply these concepts and models in real time are currently under development.20,21

REFERENCES

19. Egan TD. Target-controlled drug delivery: progress toward an intravenous “vaporizer” and automated anesthetic administration. Anesthesiology 2003;99:1214–9