An improved model for predicting postoperative nausea and vomiting in ambulatory surgery patients using physician-modifiable risk factors

Pankaj Sarin,[1] Richard D Urman,[1] Lucila Ohno-Machado[2]

[1]Department of Anesthesiology, Perioperative and Pain Medicine, Brigham and Women's Hospital, Harvard Medial School, Boston, Massachusetts, USA
[2]Division of Biomedical Informatics, University of California, San Diego, La Jolla, California, USA

Correspondence to
Dr Pankaj Sarin, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA;
psarin1@partners.org

PS and RDU contributed equally to this paper

Received 27 January 2012
Accepted 10 April 2012

Published Online First
12 May 2012

ABSTRACT

Objective Postoperative nausea and vomiting (PONV) is a frequent complication in patients undergoing ambulatory surgery, with an incidence of 20%–65%. A predictive model can be utilized for decision support and feedback for practitioner practice improvement. The goal of this study was to develop a better model to predict the patient's risk for PONV by incorporating both non-modifiable patient characteristics and modifiable practitioner-specific anesthetic practices.

Materials and methods Data on 2505 ambulatory surgery cases were prospectively collected at an academic center. Sixteen patient-related, surgical, and anesthetic predictors were used to develop a logistic regression model. The experimental model (EM) was compared against the original Apfel model (OAM), refitted Apfel model (RAM), simplified Apfel risk score (SARS), and refitted Sinclair model (RSM) by examining the discriminating power calculated using area under the curve (AUC) and by examining calibration curves.

Results The EM contained 11 input variables. The AUC was 0.738 for the EM, 0.620 for the OAM, 0.629 for the RAM, 0.626 for the SARS, and 0.711 for the RSM. Pair-wise discrimination comparison of models showed statistically significant differences ($p<0.05$) in AUC between the EM and all other models, OAM and RSM, RAM and RSM, and SARS and RSM.

Discussion All models except the OAM appeared to have good calibration for our institution's ambulatory surgery data. Ours is the first model to break down risk by anesthetic technique and incorporate risk reduction due to PONV prophylaxis.

Conclusion The EM showed statistically significant improved discrimination over existing models and good calibration. However, the EM should be validated at another institution.

OBJECTIVE

The goal of this study was to develop a better model to predict the patient's risk for postoperative nausea and vomiting (PONV) by incorporating both non-modifiable patient characteristics and modifiable practitioner-specific anesthetic practices.

BACKGROUND AND SIGNIFICANCE

PONV and postoperative pain are the most frequent complications in day surgery patients, affecting patient recovery, discharge, and overall satisfaction. Although both can significantly delay discharge, patients are more concerned about having PONV and are willing to pay up $100 out of pocket to prevent it.[1] Furthermore, having PONV greatly decreases patient satisfaction and increases three- to fourfold the risk of a costly overnight hospitalization.[2] With the number of ambulatory surgery cases increasing each year, PONV is considered the 'big little problem' and much research into this complication has been conducted over the last 30 years. Unfortunately, the incidence has remained at around 25%–30% despite changes in anesthetic practices,[3 4] possibly due to the multifactorial nature of PONV. At our institution, the Brigham and Women's Hospital, a preliminary study indicated that the overall PONV rate was 25%, with a rate as high as 65% recorded in high risk patient groups.[5]

Several regimens have been shown to reduce the incidence of PONV in the ambulatory surgery population. These regimens include avoiding inhalational agents for general anesthesia (GA), providing regional anesthesia to limit the use of opioids, and/or providing prophylactic anti-emetics in an attempt to decrease the risk of PONV.[6 7] Each of these regimens has been shown to be effective[8] but, given the multi-factorial nature of PONV, not guaranteed to eliminate PONV.[2] However, blindly applying these regimens to all ambulatory surgery patients is neither practical nor in the patient's best interest. Providing regional anesthesia may be associated with increased time getting the patient ready for surgery as well as serious complications associated with the technique. Avoiding inhalational agents for GA may raise the incidence of intraoperative awareness. Prophylactic use of anti-emetics increases the risk of having medication related side effects[9] and is not 100% effective in preventing PONV.[10] All three factors increase the cost of anesthesia to the patient and the healthcare system as a whole.[4]

The current strategy for reducing PONV relies on risk-stratifying the patient and then tailoring treatment based on the classification. Many models have been developed to classify patients into risk groups.[11 12] One of the first models for risk assessment of PONV was developed by Palazzo and Evans.[13] This model was created using logistic regression (LR) to identify variables associated with increased PONV in patients undergoing orthopedic surgery. Since then, models have been created using data from patients undergoing a greater variety of surgical procedures. The best known of these are the scoring systems of Apfel *et al*[11] and Koivuranta *et al.*[14] Both used LR on data from inpatients undergoing a number of different procedures and recorded the outcome of PONV up to 24 h after surgery. Although the models considered patient

history (eg, sex, history of PONV, history of motion sickness), intraoperative factors (eg, medications given, fluids administered), and surgical factors (eg, type of surgery, duration of surgery), both Apfel and Koivuranta decided to use the top few variables in order to simplify the final models. In both of these models, patients considered to be at extremely high risk are recommended to receive multiple anti-emetics as well as avoidance of GA and opioids if possible, whereas patients considered to be at low risk are recommended to receive no prophylaxis and treatment only if PONV develops.[6 15]

Unfortunately, the current models cannot be effectively used for any decision support purposes. Even the 'best' models do not have great discriminatory ability in predicting PONV.[10 12 16] They cannot be used for decision support beyond risk stratifying patients based on patient history, as they include only non-modifiable factors rather than factors that the physician can control. Furthermore, they cannot be used to educate anesthesiologists about practices that can potentially reduce a patient's risk of PONV.[17] Although the reasons for the simplification of the original models were valid[11 12] when the models were first developed, it should now be possible to apply improved yet more complex models[18] to predict PONV risk for decision support in anesthesia management systems that are routinely utilized in the care of patients.

We aim to design the first model to-date that can be used to better predict the patient's risk for PONV by taking into consideration not only non-modifiable patient characteristics (eg, gender, smoking history, prior history of PONV) but also modifiable physician-specific anesthetic practices.

MATERIALS AND METHODS

Based on an extensive literature search[19] and consultation with a local domain expert, 16 of the strongest predictors were selected for consideration in the model. From these clinical data, the following variables were used for model development: age (closest decade), female sex, smoking history, history of PONV, history of motion sickness, type of surgery, duration of surgery (in minutes), type of anesthesia provided, intraoperative dose of fentanyl (in micrograms), use of ondansetron for prophylaxis, use of dexamethasone for prophylaxis, use of scopolamine patch for prophylaxis, use of metoclopramide for prophylaxis, use of intramuscular ephedrine for prophylaxis, postoperative dose of opioids in analgesic equivalents, and maximum postoperative pain score.

After the study was approved by the hospital IRB, we prospectively collected data on ambulatory surgery cases from September 1, 2005 to July 1, 2007 at a tertiary care academic center. Surgeries included in the study were breast biopsy and lumpectomy, breast plastics, hysteroscopy/D&C, pelviscopy, tubal ligation, laparoscopic cholecystectomy, and herniorrhaphy. The anesthesia providers documented paper anesthesia charts in routine fashion and a research assistant transferred data from the chart to an electronic database immediately after surgery. The assistant followed each patient through their recovery and recorded data on recovery room parameters, including nausea/vomiting scores, anti-emetics received, and recovery room stay times. A complete list of variables for data collected is noted in table 1 (see Results section). Missing or unclear data were corrected by consulting the anesthesiologist or nurse responsible for documenting the data as soon as possible.

Clinical care by anesthesiologists and nurses was not altered. Each anesthesiologist was free to develop and execute an anesthetic plan, including administration of any prophylactic anti-emetics they wished. Recovery room nurses followed standard postoperative templates for administration of anti-emetics and analgesics.

Collected data were combined with length of surgery data from hospital computer systems.

Model building and analysis

The original dataset was examined for invalid or missing values and all patients with any invalid data were eliminated from the model building process. Data were imported into SAS from a CSV file. Categorical variables were reformulated as dummy variables where needed. For example, the categorical variable of surgery type was transformed into dummy variables. Continuous variables were not converted. For example, 'duration of surgery' (rounded to nearest 60 min) and 'age' (rounded to nearest decade) were left as continuous variables. Patient history factors and the outcome variable of PONV were considered binary variables. These transformations resulted in a total of 26 continuous and categorical variables for use in modeling.

Prior to model building, the data were randomly divided into a two-third training set and a one-third test set. These two sets were compared to confirm that data had indeed been split randomly and the frequency of the variables' values was not significantly different (table 1). The PROC LOGISTIC function in SAS was used to build an LR model using a stepwise algorithm on the training set with the 26 variables. Variables were entered into the model if they met a significance level of $p<0.50$ and were allowed to stay in the model if they met a significance level of $p<0.10$.[20] Model building stopped when no additional variables met these criteria.

The experimental model (EM) created was compared against the original Apfel model (OAM), a refitted Apfel model (RAM), the simplified Apfel risk score (SARS), and a refitted Sinclair model (RSM). In order to refit the Apfel and Sinclair models, new β coefficients were calculated using the variables from the original model using the PROC LOGISTIC function. Each of the models was then run against the test data. To estimate the discriminating power of the models, receiver operating characteristic (ROC) curves were created and the areas under the curves (AUC) were calculated using SAS (figure 1). The ROC curve allows visualization of the relationship between sensitivity and specificity at different probability thresholds. The AUC determines how well patients who had PONV could be distinguished from patients who did not have PONV using the model's risk prediction calculation. An AUC of 1.0 represents perfect discrimination, whereas an AUC of 0.5 represents no discrimination.[21]

Pair-wise AUC comparisons were performed to evaluate differences in discrimination between models as described previously.[22 23] Here, a non-parametric comparison of the ROC is carried out.

Calibration curves were also developed to examine the accuracy and goodness of fit of each model. For each model, the predicted number of observations was plotted against the expected number of observations for each of 10 risk percentiles created from the test set. For the simplified Apfel model, it was only possible to categorize patients into four bins. A slope of 1 with an intercept of 0 indicates perfect calibration, whereas a larger slope indicates an over-estimation of occurrence of PONV by the model. Calibration was further evaluated with Hosmer–Lemeshow goodness of fit χ^2 estimates using deciles.[24]

SAS V.9.1 (Windows platform) was used for LR modeling and analysis.

Table 1 Distribution of patient, anesthetic, and surgery characteristics between training and test datasets

Variable	Training set (n=1670)	Test set (n=835)	p Value
Duration (h)	2.8±1.0	2.8±1.0	0.92
Surgery type (%)			0.10
Hysteroscopy	25.7	27.8	
Breast biopsy/lumpectomy	25.5	23.6	
Pelviscopy/myomectomy	18.5	19.7	
Inguinal and other hernia	14.9	14.8	
Breast plastics	5.6	5.2	
Laparoscopic cholecystectomy	5.2	3.7	
Tubal ligation	3.4	3.5	
Other gyn surgery	1.3	1.8	
Age (decade)	4.8±1.4	4.9±1.3	0.26
Sex (% female)	87.8	88.6	0.60
Weight (kg)	69.3±14.4	69.5±14.7	0.73
History of PONV (%)	25.4	23.2	0.26
History of motion sickness (%)	32.3	32.3	1.00
Smoker (%)	10.2	11.6	0.27
Type of anesthesia (%)			0.35
Intravenous induction with inhalational agent and N_2O maintenance	22.3	24.1	
Intravenous induction with inhalational agent maintenance (no N_2O)	34.8	32.9	
Inhalation induction with inhalational agent and N_2O maintenance	11.5	12.3	
Inhalation induction with inhalational agent maintenance (no N_2O)	5.5	6.0	
Intravenous induction and intravenous anesthetic maintenance	2.0	1.0	
Monitored anesthesia care	24.0	23.7	
Prophylactic ondansetron (%)	47.5	47.9	0.87
Prophylactic dexamethasone (%)	41.3	39.0	0.30
Preoperative or intraoperative application of scopolamine patch (%)	17.3	17.3	1.00
Intraoperative metoclopramide =10 mg (%)	54.8	55.2	0.86
Intraoperative metoclopramide >10 mg (%)	2.5	4.7	0.003
Prophylactic IM ephedrine (%)	4.0	5.5	0.10
Intraoperative fentanyl dose (μg)	91.7±80.0	87.8±79.7	0.25
Maximum pain score in phase 1	3.3±2.4	3.3±2.5	0.77
Postoperative opioids (analgesic equivalent)	0.32±0.42	0.29±0.37	0.08
Crude PONV rate (%)	22.6	22.2	0.84
Average number of postoperative rescue anti-emetics given	0.22±0.61	0.21±0.59	0.71

None of the variables had statistically significant differences between these sets ($p>0.05$). p Value indicates the result of a χ^2 test for categorical variables or of a t test for continuous variables.
gyn, gynecological; IM, intramuscular; PONV, postoperative nausea and vomiting.

RESULTS

Data from 2505 patients were used in the model building and testing process. The distribution of patient, anesthetic, and surgery characteristics was not significantly different between the training and test sets and is shown in table 1.

Ultimately, the EM contained 11 input variables, including five patient history and outcome factors, two surgical factors, and four anesthetic factors. Patient history and outcome factors included: age, female sex, history of motion sickness, history of PONV, and maximum pain score postoperatively. Surgical factors included duration of surgery and type of surgery. Types of surgery found to have a significant effect on PONV were laparoscopic cholecystectomy, tubal ligation, and pelviscopy. Anesthetic factors found to have an effect on PONV included: intraoperative fentanyl dose, prophylactic dexamethasone use, prophylactic ondansetron use, and type of anesthetic performed. Inhalation induction with use of N_2O/anesthetic agent for maintenance, use of monitored anesthesia care, and intravenous induction with propofol but no use of N_2O were found to have an effect on risk of PONV. PONV was the binary outcome variable in the model.

The 26 input variables, the final EM, its β coefficients, and ORs are shown in tables 2 and 3.

The strongest patient history and surgical factors increasing risk were laparoscopic cholecystectomy, history of PONV, tubal ligation, pelviscopy, history of motion sickness, and duration of surgery. Anesthetic factors increasing risk included the type of anesthetic utilized, with inhalation induction with concomitant use of N_2O increasing risk the most. Age and prophylactic use of ondansetron or dexamethasone reduced PONV risk.

New β coefficients calculated for the RAM and the RSM are shown in table 4.

Results of the performance of the EM as compared to that of the OAM, the RAM, the SARS, and the RSM on the validation dataset are shown in table 5.

The AUC for the EM was 0.738, 0.620 for the OAM, 0.629 for the RAM, 0.626 for the SARS, and 0.711 for the RSM. ROC curves for each model are shown in figure 1.

Pair-wise discrimination comparison of models is shown in table 5. Statistically significant differences ($p<0.05$) in AUC were noted between the EM and all other models, OAM and RSM, RAM and RSM, and SARS and RSM.

Calibration, as indicated by goodness of fit by the Hosmer–Lemeshow χ^2 test, is shown in figure 2. All models, except the OAM, appeared to have good calibration for our ambulatory surgery data as indicated by a p value of >0.05.

DISCUSSION

In our EM, history of PONV, female sex, history of motion sickness, age, duration of surgery, type of surgery, intraoperative fentanyl dose, prophylactic dexamethasone use, prophylactic

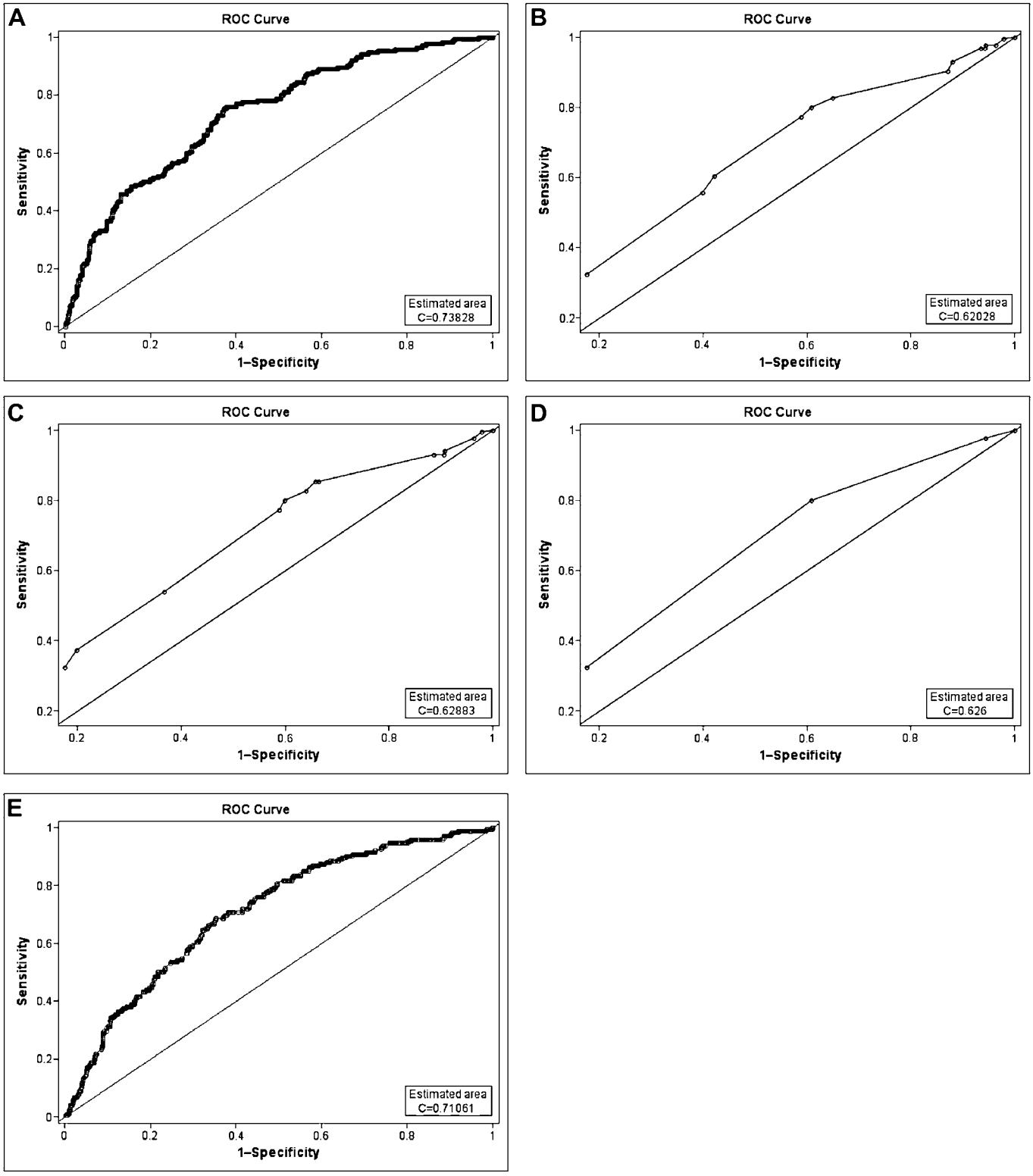

Figure 1 Receiver operating characteristic (ROC) curves for various models evaluated on the test dataset: (A) experimental LR model (EM), (B) original Apfel model (OAM), (C) refitted Apfel model (RAM), (D) simplified Apfel risk score (SARS), and (E) refitted Sinclair model (RSM).

ondansetron use, type of anesthetic performed, and maximum pain score postoperatively were determined to be independent predictors of PONV. Inclusion of these variables is consistent with published studies as well as other models.

One of the strongest predictors of PONV was the patient's history of prior PONV (OR 2.7). Prior history of motion sickness resulted in a 1.5 increased risk of PONV. Female gender increased risk by 2.2, consistent with other models that have demonstrated it to be an important predictor of PONV.[25] An increase in one decade of age resulted in slightly decreased risk of PONV (OR 0.9), whereas every additional hour of surgery increased the risk of PONV (OR 1.2).

Several types of surgery were found to have a significant effect on PONV. Of these, laparoscopic cholecystectomy increased the

Table 2 Twenty-six input variables used in model building

	Variable	Present or absent from final model
1	Duration	Present
2	Surgery type: hysteroscopy	Absent
3	Surgery type: breast biopsy/lumpectomy	Absent
4	Surgery type: pelviscopy/myomectomy	Present
5	Surgery type: inguinal and other hernia	Absent
6	Surgery type: breast plastics	Absent
7	Surgery type: laparoscopic cholecystectomy	Present
8	Surgery type: tubal ligation	Present
9	Age	Present
10	Sex	Present
11	History of postoperative nausea and vomiting (PONV)	Present
12	History of motion sickness	Present
13	Smoker	Absent
14	Type of anesthesia—intravenous induction with inhalational agent and N_2O maintenance	Absent
15	Type of anesthesia - intravenous induction with inhalational agent maintenance (no N_2O)	Present
16	Type of anesthesia—inhalation induction with inhalational agent and N_2O maintenance	Present
17	Type of anesthesia—inhalation induction with inhalational agent maintenance (no N_2O)	Absent
18	Type of anesthesia—monitored anesthesia care	Present
19	Prophylactic ondansetron	Present
20	Prophylactic dexamethasone	Present
21	Preoperative or intraoperative application of scopolamine patch	Absent
22	Intraoperative metoclopramide use	Absent
23	Prophylactic intramuscular ephedrine	Absent
24	Intraoperative fentanyl dose	Present
25	Maximum pain score in phase 1	Present
26	Postoperative opioids (analgesic equiv.)	Absent

risk of PONV the most, with an OR of 2.3. Tubal ligation increased risk twofold. Pelviscopy also increased risk (OR 1.4), a finding that is consistent with the literature.[15]

Anesthetic factors found to have an effect on PONV included: intraoperative fentanyl dose, type of general anesthetic performed, prophylactic use of dexamethasone, and prophylactic use of ondansetron. Previous models have not differentiated between intraoperative and postoperative doses of opioids and most have focused on long term opioids such as morphine. The literature is inconsistent on whether intraoperative opioids actually increase risk, although most current literature indicates that postoperative opioid dose does correlate with increased risk of PONV. Here, although there is a statistically significant increase in PONV risk with intraoperative fentanyl use, an OR of 1.004 indicates that risk is not increased greatly and so this may not be clinically significant. The lack of a greater association of intraoperative fentanyl use with PONV risk may be due to the use of relatively low doses of fentanyl in ambulatory surgery patients at our institution.

Our model did not find that postoperative opioids are associated with increased risk. However, a patient's postoperative pain was shown to be correlated with the PONV risk. The higher the maximum pain score (scale of 0–10, with 10 being the worst pain experienced) that a patient reported, the more at

Table 3 Final experimental model (EM) created using logistic regression

	β Coefficient	p Value	OR estimate	95% CI
Intercept	−2.6592	<0.0001		
Age	−0.1473	0.005	0.863	0.778 to 0.957
Female sex	0.8088	0.003	2.245	1.297 to 3.886
History of postoperative nausea and vomiting	1.0086	<0.0001	2.742	2.072 to 3.629
History of motion sickness	0.4015	0.003	1.494	1.141 to 1.956
Maximum postoperative pain score	0.1200	<0.0001	1.128	1.065 to 1.194
Duration of surgery	0.1796	0.01	1.197	1.044 to 1.372
Surgery type (vs 'other')				
Laparoscopic cholecystectomy	0.8401	0.001	2.317	1.373 to 3.907
Tubal ligation	0.6724	0.03	1.959	1.055 to 3.636
Pelviscopy	0.3143	0.06	1.369	0.983 to 1.908
Intraoperative fentanyl dose	0.00376	<0.0001	1.004	1.002 to 1.006
Intraoperative dexamethasone use	−0.3428	0.01	0.710	0.540 to 0.932
Intraoperative ondansetron use	−0.3458	0.02	0.708	0.530 to 0.946
Type of anesthesia (vs 'other')				
Intravenous induction with inhalational agent maintenance (no N_2O)	−0.3243	0.04	0.723	0.523 to 0.999
Inhalational induction with inhalational agent and N_2O maintenance	0.6192	0.003	1.857	1.231 to 2.804
Monitored anesthesia care	−1.2323	<0.0001	0.292	0.180 to 0.471

Table 4 β Coefficients and ORs for refitted Apfel and Sinclair models

	Original β coefficient	New β coefficient	p Value	Odds ratio estimate	95% CI
Apfel model					
Intercept	−2.28	−3.48	<0.0001		
Female sex	1.27	1.19	<0.0001	3.279	1.972 to 5.452
History of PONV *or* motion sickness	0.65	0.90	<0.0001	2.471	1.937 to 3.153
Non-smoker	0.72	0.14	0.49	1.155	0.767 to 1.739
Postoperative opioid use	0.78	0.13	<0.0001	2.351	1.806 to 3.061
Sinclair model					
Intercept	−5.97	−2.41	<0.0001		
Male sex	−1.03	−0.70	0.01	0.498	0.735 to 0.894
History of PONV	1.14	0.98	<0.0001	2.652	2.039 to 3.449
Smoker	−0.42	−0.12	0.58	0.888	0.585 to 1.348
Age	−0.14	−0.21	<0.0001	0.811	0.735 to 0.894
Duration of surgery	0.46	0.30	<0.0001	1.355	1.192 to 1.541
Primary anesthesia=GA	2.36	1.22	<0.0001	3.388	2.232 to 5.143
Breast plastics	1.90	−0.20	0.46	0.818	0.481 to 1.393
Gyn surgery, but not D&C	1.20	0.07	0.64	1.070	0.808 to 1.416

GA, general anesthesia; gyn, gynecological; IM, intramuscular; PONV, postoperative nausea and vomiting.

risk he/she was for PONV (OR 1.1). Although some experts have hypothesized that postoperative pain is associated with PONV,[15 26] it has been hard to separate out whether postoperative pain or postoperative opioid use is the cause of increased PONV, as a patient complaining of pain postoperatively is treated with opioids.

It is reassuring that the type of anesthetic given appeared to be associated with PONV risk in our model. Recent studies indicate that exposure to an inhalation agent increases PONV risk[7] and many experts advocate avoidance of inhalation agents for high risk groups.[6 15] In our model, we were able to differentiate between different types of general anesthetic techniques that some have postulated, but have not proven, may increase risk. For example, our model indicates that inhalation induction with concomitant use of nitrous oxide increases risk almost twofold (OR 1.9) over other anesthetics such as inhalation induction without use of nitrous oxide. This increased risk is even greater when compared to intravenous induction with inhalational agent maintenance but no N_2O use. Monitored anesthesia care is associated with a greatly reduced risk of PONV over general anesthetic techniques. None of the previously published models break down risk by anesthetic technique, something that can be controlled by the anesthesiologist.

As expected, prophylactic use of dexamethasone and ondansetron independently reduced the risk of PONV. This risk reduction was similar for each medication and is consistent with the 20%–25% risk reduction quoted in the literature.[27 28] Ours is the first model to incorporate these variables and to demonstrate to physicians the risk reduction resulting from their use after controlling for confounding variables. Prophylactic use of scopolamine did not appear in the model. This may be due to the fact that the outcome of PONV was only recorded up to time of discharge, whereas scopolamine is thought to have more of an effect on post-discharge nausea and vomiting. Use of intravenous metoclopramide 10 mg or 20 mg was not shown to affect PONV risk. This result is consistent with the literature which indicates that numbers needed to treat to prevent one case of PONV with metoclopramide are very high.[29] Intramuscular ephedrine use did not appear in the model either, perhaps due to the fact that only 4.5% of all patients received this treatment and its efficacy could not be determined with such a low frequency.

History of smoking did not appear in our model, even though it is well proven to be associated with decreased risk of PONV. This result may be due to the patient population studied. Here, only 10.7% of patients were smokers and, therefore, the risk associated with being a non-smoker may be incorporated into

Table 5 Pair-wise discrimination comparison of models and summary of discrimination and calibration performance for each model

	Experimental model (EM)		Original Apfel model (OAM)		Refitted Apfel model (RAM)		Simplified Apfel risk score (SARS)	
Model	**Diff.**	**p Value**	**Diff.**	**p Value**	**Diff.**	**p Value**	**Diff.**	**p Value**
Experimental model (EM)								
Original Apfel model (OAM)	0.118	<0.0001						
Refitted Apfel model (RAM)	0.110	<0.0001	−0.009	0.48				
Simplified Apfel risk score (SARS)	0.112	<0.0001	−0.006	0.43	0.003	0.73		
Refitted Sinclair model (RSM)	0.028	0.03	−0.090	0.0001	−0.082	<0.0001	−0.085	0.0001

Model	**AUC for test set**	**95% CI for AUC**	**HL χ^2**	**HL (p)**
Experimental model (EM)	0.738	0.698 to 0.779	10.0	0.27
Original Apfel model (OAM)	0.620	0.576 to 0.665	66.9	<0.0001
Refitted Apfel model (RAM)	0.629	0.585 to 0.673	1.8	0.87
Simplified Apfel risk score (SARS)	0.626	0.584 to 0.668	0.5	0.77
Refitted Sinclair model (RSM)	0.711	0.669 to 0.752	7.4	0.50

AUC, area the under receiver operating characteristic curve; Diff, AUC difference; HL χ^2, Hosmer–Lemeshow χ^2; HL (p), Hosmer–Lemeshow probability > χ^2; p Value, p value of difference.

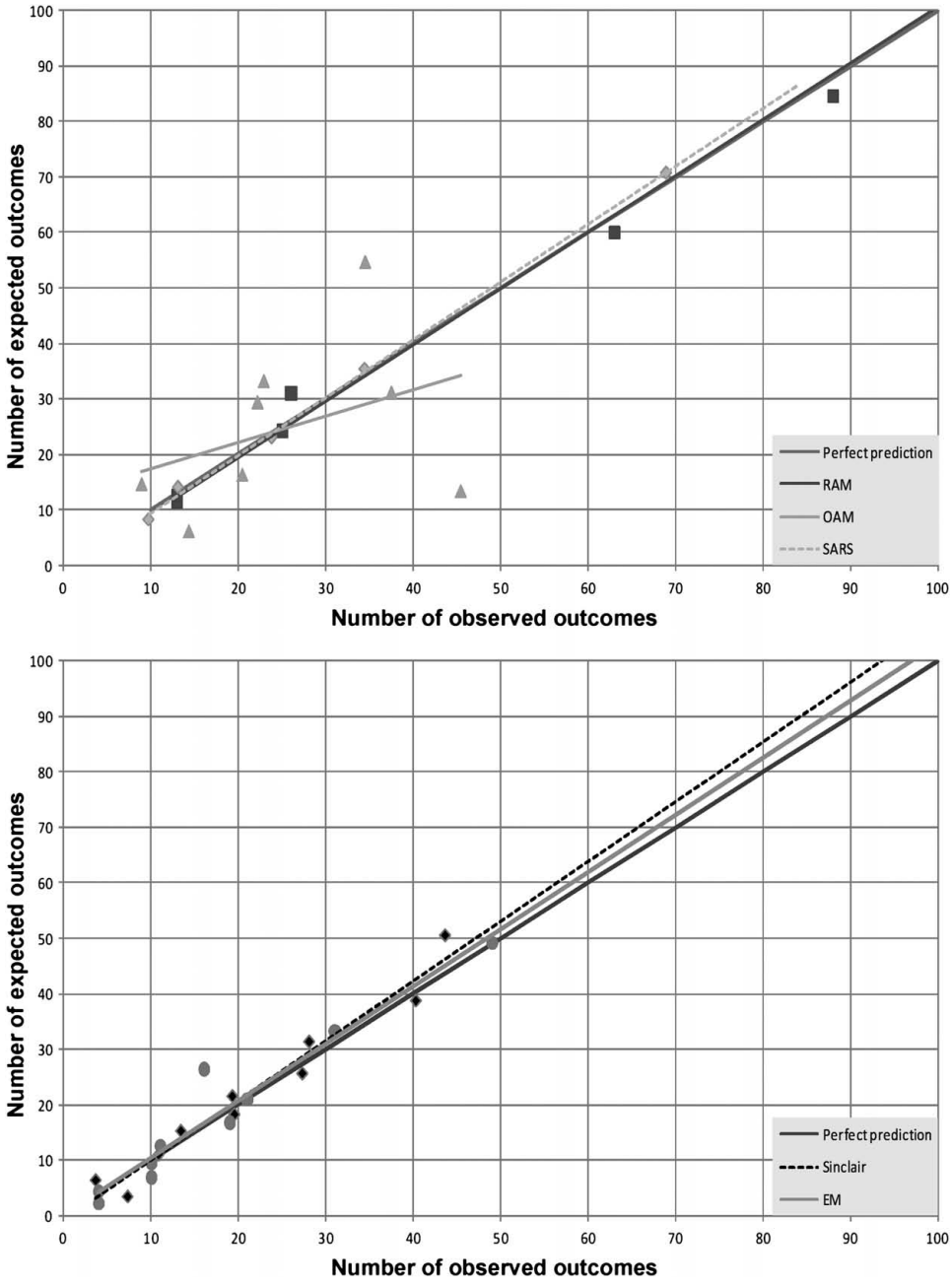

Figure 2 Calibration curves for models evaluated on the test dataset. EM, experimental model; OAM, original Apfel model; RAM, refitted Apfel model; SARS, simplified Apfel risk score.

the baseline risk rather than be an independent predictor. This hypothesis is further supported by the fact that when β coefficients were calculated to adjust models that usually contain non-smoking as a predictor, non-smoking did not show up as a statistically significant variable in this population.

The newly developed model shows statistically significant improved discrimination over the OAM, the refitted Apfel score, the SARS, and the RSM.

The OAM, the refitted Apfel score, and the SARS all performed similarly and there was no statistically significant difference between their performances. The OAM had worse calibration than the other two, as was to be expected. The SARS had good calibration for data at our institution because it classified patients into only four quartiles. The performance results of the Apfel models from this study are consistent with results at other validation centers.[11] The Apfel models, although often advocated for use because of their simplicity, are clearly not applicable to our institution's ambulatory surgery patient population and may not be applicable to other institutions' patients either. Another reason for the Apfel models' poor performance may be that the original models were developed using data from inpatients rather than ambulatory surgery patients.

The RSM had good discrimination and good calibration for our institution's data. This result may be explained by the fact that the model was developed using data from ambulatory surgery patients and is very similar to the EM developed in the types of variables used. However, our EM contained more detailed variables and contained predictors that appeared in an aggregated form in the Sinclair model. For example, the Sinclair model differentiates between patients who had GA and those who did not. Our model also considers the contribution of different techniques used to induce GA to the overall risk. This may explain why our model provided better discrimination.

In this study, we validate and demonstrate that three major models do not have great discriminatory ability in determining PONV risk in our patient population, even after recalibration. Furthermore, we have developed a predictive model that takes into consideration not only non-modifiable patient characteristics but also modifiable physician-specific anesthetic practices. Thus, our model goes beyond risk-stratifying patients based on patient history alone. Because the model contains more precise variables than other published models, it can be used for teaching purposes as well as be incorporated into anesthesia information management systems for decision support purposes. For example, it can be used to demonstrate how physician practice, such as anesthetic technique, affects patient risk of PONV. Therefore, this new model could be used in computer decision support to potentially help improve clinician compliance with best practices.

CONCLUSION

Variables appearing in our model are consistent with those in the literature. We prove that this more inclusive model performs better than existing models in the ambulatory surgery patient population. However, our EM should be validated at an outside institution as it is known that risk prediction models often do not perform as well at institutions other than the one where they were developed.

Contributors PS and RDU contributed equally and were involved in all phases of manuscript preparation. LO-M provided overall supervision and direction for this research.

Funding The study was carried out using internal resources.

Competing interests None.

Ethics approval Ethics approval was provided by Brigham and Women's Hospital IRB.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Gan T,** Sloan F, Dear Gde L, *et al*. How much are patients willing to pay to avoid postoperative nausea and vomiting? *Anesth Analg* 2001;**92**:393—400.
2. **Habib AS,** Gan TJ. Evidence-based management of postoperative nausea and vomiting: a review. *Can J Anaesth* 2004;**51**:326—41.
3. **Palazzo MG,** Strunin L. Anaesthesia and emesis. I: Etiology. *Can Anaesth Soc J* 1984;**31**:178—87.
4. **Watcha MF.** The cost-effective management of postoperative nausea and vomiting. *Anesthesiology* 2000;**92**:931—3.
5. **Sarin P,** Philip BK, Mitani A, *et al*. Specialized ambulatory anesthesia teams contribute to decreased ambulatory surgery recovery room length of stay. *Ochsner J* 2012;**12**:94—100.
6. **Gan TJ,** Meyer T, Apfel CC, *et al*. Consensus guidelines for managing postoperative nausea and vomiting. *Anesth Analg* 2003;**97**:62—71, table of contents.
7. **Apfel CC,** Kranke P, Katz MH, *et al*. Volatile anaesthetics may be the main cause of early but not delayed postoperative vomiting: a randomized controlled trial of factorial design. *Br J Anaesth* 2002;**88**:659—68.
8. **Pierre S,** Corno G, Benais H, *et al*. A risk score-dependent antiemetic approach effectively reduces postoperative nausea and vomiting—a continuous quality improvement initiative. *Can J Anaesth* 2004;**51**:320—5.
9. **Tramer MR.** A rational approach to the control of postoperative nausea and vomiting: evidence from systematic reviews. Part I. Efficacy and harm of antiemetic interventions, and methodological issues. *Acta Anaesthesiol Scand* 2001;**45**:4—13.
10. **Domino KB,** Anderson EA, Polissar NL, *et al*. Comparative efficacy and safety of ondansetron, droperidol, and metoclopramide for preventing postoperative nausea and vomiting: a meta-analysis. *Anesth Analg* 1999;**88**:1370—9.
11. **Apfel CC,** Laara E, Koivuranta M, *et al*. A simplified risk score for predicting postoperative nausea and vomiting: conclusions from cross-validations between two centers. *Anesthesiology* 1999;**91**:693—700.
12. **Apfel CC,** Kranke P, Eberhart LH, *et al*. Comparison of predictive models for postoperative nausea and vomiting. *Br J Anaesth* 2002;**88**:234—40.
13. **Palazzo M,** Evans R. Logistic regression analysis of fixed patient factors for postoperative sickness: a model for risk assessment. *Br J Anaesth* 1993;**70**:135—40.
14. **Koivuranta M,** Läärä E.Snare L**,** *et al*. A survey of postoperative nausea and vomiting. *Anaesthesia* 1997;**52**:443—9.
15. **Gan TJ.** Risk factors for postoperative nausea and vomiting. *Anesth Analg* 2006;**102**:1884—98.
16. **van den Bosch JE,** Kalkman CJ, Vergouwe Y, *et al*. Assessing the applicability of scoring systems for predicting postoperative nausea and vomiting. *Anaesthesia* 2005;**60**:323—31.
17. **Overdyk FJ,** Harvey SC, Baldwin D, *et al*. Individualized outcome feedback produces voluntary antiemetic prescribing practice changes. *J Clin Anesth* 1999;**11**:17—23.
18. **Shiffman RN,** Liaw Y, Brandt CA, *et al*. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. *J Am Med Inform Assoc* 1999;**6**:104—14.
19. **Apfel CC,** Kranke P, Greim CA, *et al*. What can be expected from risk scores for predicting postoperative nausea and vomiting? *Br J Anaesth* 2001;**86**:822—7.
20. **Shtatland ES,** Kleinman K, Cain EM. Stepwise methods in using SAS PROC LOGISTIC and SAS ENTERPRISE MINER for prediction. *SAS Users Group International (SUGI) Meeting*. 2003;**258-28**:1—6.
21. **Hanley JA,** McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29—36.
22. **DeLong ER,** DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837—45.
23. **Hanley JA,** McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;**148**:839—43.
24. **Lemeshow S,** Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;**115**:92—106.
25. **Cohen MM,** Duncan PG, DeBoer DP, *et al*. The postoperative interview: assessing risk factors for nausea and vomiting. *Anesth Analg* 1994;**78**:7—16.
26. **Chia YT,** Kuo MC, Liu K, *et al*. Does postoperative pain induce emesis? *Clin J Pain* 2002;**18**:317—23.
27. **Apfel CC,** Korttila K, Abdalla M, *et al*. A factorial trial of six interventions for the prevention of postoperative nausea and vomiting. *N Engl J Med* 2004;**350**:2441—51.
28. **Thomas R,** Jones N. Prospective, randomized, double-blind comparative study of dexamethasone, ondansetron, and ondansetron plus dexamethasone as a prophylactic antiemetic therapy in patients undergoing day-case gynaecological surgery. *Br J Anaesth* 2001;**87**:588—92.
29. **Henzi I,** Walder B, Tramer MR. Metoclopramide in the prevention of postoperative nausea and vomiting: a quantitative systematic review of randomized, placebo-controlled studies. *Br J Anaesth* 1999;**83**:761—71.