Apfel's simplified score may favourably predict the risk of postoperative nausea and vomiting

[La cotation d'Apfel simplifiée peut prédire favorablement le risque de nausées et de vomissements postopératoires]

Sébastien Pierre MD, Hervé Benais MD, Jacques Pouymayou MD

Purpose: To compare two of the latest published scores for predicting postoperative nausea and vomiting (PONV) in potentially high-risk patients.

Methods: Adult in-patients scheduled for throat, thyroid, breast or gynecological surgery under general inhalational anesthesia were studied prospectively over 24 hr for PONV. The latest published score considers four risk factors: female gender, previous history of PONV or motion sickness, non-smoking status and postoperative use of opioids (Apfel-score). The previously published score includes, in addition to these factors, duration, type of anesthesia and surgery (Sinclair-score). The two scores were compared by calculating the area under a receiver operating characteristic (ROC)-curve and plotting calibration curves of the predicted and the observed incidence of PONV.

Results: Five hundred consecutive patients were studied and patients who received prophylactic antiemetics were excluded. Of the remaining 428 patients 49.5% suffered from PONV. Multivariable analysis revealed that age, gender, previous history of PONV or motion sickness and postoperative use of opioids had an impact on PONV. The area under the ROC-curve was significantly greater for the Apfel-score compared to the Sinclair-score (0.71 *vs* 0.64, $P=0.008$). The correlation between the predicted (x) and the observed (y) incidence for the Apfel-score and for the Sinclair-score was $y=1.08x - 0.07$ and $y=0.93x + 0.27$.

Conclusion: In our hospital, the simplified Apfel-score presented with favourable discriminating and calibration properties for predicting the risk of PONV. Therefore, we have implemented this score in our daily clinical practice as well as in an ongoing antiemetic trial.

Objectif : *Comparer deux des dernières cotations publiées servant à prédire l'incidence de nausées et de vomissements postopératoires (NVPO) chez des sujets virtuellement à haut risque.*

Méthodes : *Des patients adultes, hospitalisés pour une intervention chirurgicale pharyngée, thyroïdienne, gynécologique ou du sein, sous anesthésie générale par inhalation, ont participé à l'étude prospective de 24 h sur les NVPO. Les dernières cotations publiées considèrent quatre facteurs de risque : être de sexe féminin, avoir des antécédents de NVPO ou de mal des transports, être non-fumeur et utiliser des opioïdes postopératoires (cotation d'Apfel). Les cotations publiées antérieurement y ajoutent la durée, le type d'anesthésie et de chirurgie (cotation de Sinclair). Les deux cotations ont été comparées en calculant l'aire sous la courbe d'analyse ROC et les courbes d'étalonnage du tracé de l'incidence prédite et observée de NVPO.*

Résultats : *Cinq cents patients successifs ont été étudiés. Ceux qui ont reçu des antiémétiques prophylactiques ont été exclus. Des 428 patients restants, 49,5 % ont eu des NVPO. L'analyse multivariée a révélé que l'âge, le sexe, les antécédents de NVPO ou la maladie des transports et l'usage postopératoire d'opioïdes agissent sur les NVPO. L'aire sous la courbe ROC a été significativement plus grande pour la cotation d'Apfel, comparée à la cotation de Sinclair (0,71 vs 0,64, $P = 0,008$). La corrélation entre l'incidence prévue (x) et observée (y) a été pour la cotation d'Apfel et pour la cotation de Sinclair $y = 1,08 x - 0,07$ et $y = 0,93 x + 0,27$.*

Conclusion : *La cotation d'Apfel simplifiée a présenté des propriétés discriminantes favorables et des qualités d'étalonnage utiles à la prédiction du risque de NVPO. Nous l'avons donc intégrée à notre pratique clinique quotidienne ainsi qu'à un essai antiémétique en cours.*

From the Department of Anesthesia, Institut Claudius Regaud, Toulouse, France.

Address correspondence to: Dr. Sébastien Pierre, Department of Anesthesia, Institut Claudius Regaud, 20-24 rue du Pont St Pierre, 31052 Toulouse, France. Phone: +33 5 61 42 46 11; Fax: +33 5 61 42 41 17; E-mail: pierre@icr.fnclcc.fr

Funding: Institutional resources.

Accepted for publication July 30, 2001.

Revision accepted November 21, 2001.

THE incidence of postoperative nausea and vomiting (PONV) ranges from 20 to 30% after general anesthesia[1] and can reach up to 80% in patients undergoing breast surgery.[2] Since there is still debate between prophylaxis and treatment,[3] two experts have recently proposed a "decision tree".[3,4] They suggested classifying patients in four groups according to their predicted risk for PONV, low-risk patients (<10%), mild-moderate risk patients (10–30%), high-risk patients (30–60%) and extremely high-risk patients (>60%) and applying prophylactic and/or rescue antiemetics depending on this classification.[3,4] As this model appears quite convincing and straightforward, we intend to validate this decision making process in our clinical setting. However, several PONV risk scores are available. One score by Sinclair, Chung and Mezei[5] considers 12 predictors (Sinclair-score) while another, by Apfel and colleagues,[6] considers just four risk factors (Apfel-score). Therefore, we compared both scores to decide which one appears to be most suitable for clinical routine as well as for an intended future study.

Methods

Patient selection

After approval of our Institutional Ethics Committee, we prospectively studied 528 consecutive surgical inpatients between January 20th 2000 and August 5th 2000 in the cancer referral hospital of Toulouse (except 28 patients, who were not studied because of incomplete intraoperative documentation). All adult patients, ASA I to III, scheduled for elective throat, thyroid, breast, and gynecological surgery under general anesthesia with volatile agents were initially included.

Anesthesia

The anesthetic regimen remained open and unchanged in order to represent our daily practice. In short, 10 mg midazolam was given orally for premedication; patients with a history of allergies received 100 mg hydroxyzine. Induction of anesthesia was performed with 10–15 μg sufentanil and 2–3 mg·kg^{-1} propofol (or more if clinically necessary). Intubation was facilitated with 0.15 mg·kg^{-1} cis-atracurium or mivacurium. Anesthesia was maintained with volatile anesthetics (sevoflurane or desflurane) and 5 μg boluses of sufentanil were given as dictated by clinical needs. Neuromuscular blockade was reversed with a combination of 40 μg·kg^{-1} neostigmine and 15 μg·kg^{-1} atropine.

Data collection

In addition to the anesthetic protocol, the anesthesiologist completed a form with variables necessary to calculate the probability of PONV by the two models (Table I). Thyroid surgery was classified as "ENT" and breast surgery as "plastic". In the postanesthetic care unit and in the surgical unit, trained nurses recorded on the same form any episode of retching or vomiting. Nausea was assessed hourly during the first two hours, every two hours for the following four hours and every four hours until the 24th hour. Nausea was evaluated on a three-point scale from 0 (no nausea), 1 (mild nausea) to 2 (severe nausea). A patient was classified to have had PONV if any nausea and/or vomiting occurred within the first 24 postoperative hours. On the following day, the first and the second author consulted nurses, reviewed records and anesthetic protocol, and interviewed patients to ensure high quality data collection. The patients stayed at least 24 hr in the surgical unit.

Statistical analysis

As the anesthetic protocol differed from those studied by Apfel *et al.* and by Sinclair, Chung and Mezei (premedication with hydroxyzine, co-induction with midazolam), we performed a multivariable analysis including these two categorical variables: premedication with hydroxyzine (no=0; yes=1) and co-induction with midazolam (no=0; yes=1). PONV was the dependent variable. As previously suggested,[7] premedication with a antihistaminic drug was inversely correlated with PONV (odds ratio; OR 0.45, 95% confidence interval; CI 0.25–0.82, P=0.009), whereas co-induction with benzodiazepine did not reach statistical significance. Therefore, patients who received antihistaminic drugs were excluded so that we repeated multivariable analyses with the remaining variables of 428 patients (Table I), in order to characterize our specific risk factors of PONV.

The score of Sinclair, Chung and Mezei (Sinclair-score) adapted for our surgical specialities requires the consideration of the coefficients in a logit model (Table I) to calculate the probability of PONV,[5]

$$P = 1/(1 + e^{(\text{logit } P)})$$

where

logit (P)= -5.97 + -0.14 *Age + -1.03 *Sex + -0.42 *Smoking + 1.14 *PONVHistory + 0.46 *Duration + 2.36 *GA + 1.48 *ENT + 1.9 *Plastic + 1.2 *Gyn.

The score from Apfel and co-workers (Apfel-score) – although originally based on logistic regression too – was simplified in a way so that just the number of the four predictors female gender, history of PONV or motion sickness, non-smoking status and the use of postoperative opioids needs to be considered.[6] In their cross-validation they could demonstrate that the inci-

TABLE I Variables used to calculate risk of PONV

Age	Age in yr/10 (e.g., 52 yr= 5.2 decades)
Sex	Female=0; male= 1
Previous PONV (postoperative nausea and vomiting) or motion sickness	no=0; yes=1
Smoking	no= 0; yes= 1
Duration (duration of surgery)	in 30-min increments
GA (general anesthesia)	no=0; yes=1
ENT (ear, nose and throat surgery)	no=0; yes=1
Plastic (plastic surgery)	no=0; yes=1
Gynecological (gynecological surgery)	no=0; yes=1
POUO (postoperative use of opioids)	no=0; yes=1
Mdz (co-induction with midazolam)	no=0; yes=1

TABLE II Patient characteristics

Age	53 (45–63)
Female gender	386 (90)
Previous postoperative nausea and vomiting or motion sickness	97 (23)
Smoking	102 (24)
Duration (hr)	1 (1–1.5)
Throat and thyroid surgery	105 (24)
Breast surgery	290 (68)
Gynecological surgery	33 (8)
Postoperative use of opioids	295 (69)
Co-induction with midazolam	266 (62)

Results are presented as median (lower-upper quartiles) or number of patients (%).

dences of PONV were 10%, 21%, 39%, 61% and 79% if zero, one, two, three or four of the mentioned risk factors were present, respectively.

For each patient, the two theoretical risk scores were calculated. For example, the risk score of PONV for a 40- yr-old, non-smoking woman, with no history of PONV or motion sickness, undergoing breast surgery and with no expected use of postoperative opioids is 21% and 39%, respectively, for the Sinclair- and the Apfel-score. These risks were calculated for every patient. The patients were ranked according to the predicted risks. Each risk was used as a decision criterion, i.e., all patients with a lower risk were predicted not to suffer PONV while patients above that risk were predicted to suffer from PONV. This results in a high number of corresponding sensitivities and specificities which leads to the construction of the receiver operating characteristic (ROC)-curve.

The area under the ROC-curve,[8] was used to estimate the discriminating power of the scores (AUC). An AUC of 0.5 means that the score cannot discriminate patients with PONV and patients without PONV. Conversely, an AUC of 1 represents a perfect discrimination. To demonstrate the usefulness of the score for different risk groups, the patients were classified by their calculated probability of PONV into five risk percentiles for the Sinclair-score and into five groups (10%, 21%, 39%, 61% and 79%) for the Apfel-score.

The actual incidence of PONV was plotted against the mean of the predicted incidence and compared using weighted linear regression analysis. The slope and the intercept of the fitted regression line illustrate whether the score tested under- or overestimates PONV. A slope of 1 with an intercept of 0 represents a perfect calibration. Finally, as the group allocation suggested by White and Watcha in their "decision tree"[3,4] contained only four groups, 5% (<10% risk of PONV), 20% (10–30% risk of PONV), 45% (30–60% risk of PONV) and 80% (>60% risk of PONV), patients were categorized again by their calculated probability of PONV and new weighted linear regression analyses were executed. Most of the calculations were performed using SPSS 6.13 (SPSS Inc., USA). The comparison of the ROC curves was calculated with MedCalc (Version 4.20 for Windows).

Results

Amongst the 428 analysed patients, 49.5% experienced PONV. Nausea occurred in 47% of patients and vomiting in 26%, predominantly during the first six hours (72% of all patients with nausea and 75% of all patients who vomited). Most of our patients were non-smoking women without previous history of PONV, undergoing breast surgery (Table II).

Multivariable analysis revealed that age, gender, previous history of PONV or motion sickness and postoperative use of opioids had an impact on PONV in our setting (Table III). When logistic regression analysis was restricted to these four, strongest factors, age inversely correlated with PONV (OR 0.83 per decade, 95% CI 0.71–0.98) which corresponded to a 17% decrease of the probability of PONV for a ten-year increase in age. Male sex was also associated with a lower incidence of PONV (OR 0.41, 95% CI 0.20–0.88), i.e., the OR for females was similar to other studies 2.44 (1.2–5.0). A previous history of PONV or motion sickness and the postoperative use of opioids increased the risk of PONV more than fourfold (OR 4.30, 95% CI 2.47–7.44 and OR 4.46, 95% CI 2.75–7.21, respectively). Smoking did not reach statistical significance (P=0.068) and was rejected by the model, as other non-significant variables such as the type of surgery.

The areas under the ROC-curve (Figure 1), were significantly different: 0.64 for the Sinclair-score and 0.71 for the Apfel-score (P=0.008).

TABLE III Results of multivariable analysis with all variables in the equation for the 428 remaining patients

Independent variables	*Beta*	*S.E.*	*P-value*	*OR*	*95% CI*
Gender	-0.89	0.45	0.048	0.41	0.17; 0.99
Age	-0.22	0.09	0.013	0.81	0.68; 0.96
Previous PONV	1.50	0.29	0.000	4.46	2.52; 7.88
POUO	1.56	0.26	0.000	4.78	2.87; 7.95
Smoking	-0.50	0.27	0.068	0.61	0.36; 1.04
Midazolam	-0.31	0.23	0.170	0.73	0.47; 1.14
Plastic surgery	-0.97	0.90	0.284	0.38	0.07; 2.23
Gynecological surgery	-0.94	0.96	0.325	0.39	0.06; 2.55
ENT surgery	-0.68	0.91	0.457	0.51	0.09; 3.03
Duration (30-min intervals)	-0.01	0.08	0.854	0.99	0.85; 1.15
Constant	1.06	0.98	0.283	2.88	

S.E.=standard errors; PONV=postoperative nausea and vomiting; POUO=postoperative use of opioids; ENT=ears, nose and throat; OR=odds ratio; CI=confidence interval.

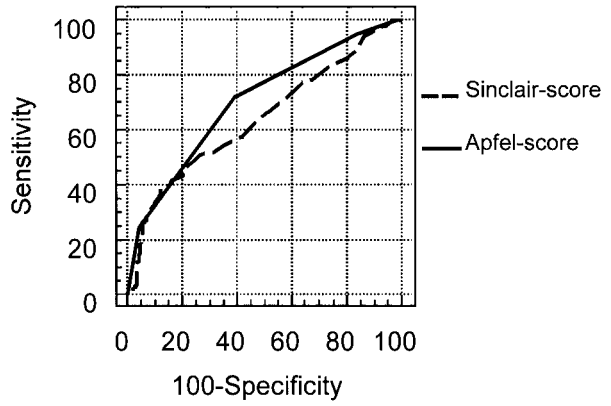

FIGURE 1 Receiver operating characteristic curves for the prediction of postoperative nausea and vomiting (PONV) for the two scores.

When the actual incidence of PONV was plotted against the mean of the predicted incidence in the five risk percentiles (Sinclair-score), the calibration line had a slope of 0.93 and an intercept of 0.27 (R^2=0.95; P<0.001). For the Apfel-score, using weighted linear regression analysis, the slope was 1.08 and the intercept -0.07 (R^2=0.96; P<0.001) (Figure 2). New patient classifications in four groups, according to the suggestion of White and Watcha did not change the calibration curve of the Apfel-score (y=1.026x - 0.0355; R^2=0.98; P<0.001) but decreased the slope of the Sinclair-score (y=0.76x + 0.31; R^2=0.63; P<0.001), because patients classified at extremely high-risk actually did not suffer PONV as frequently as predicted (Figure 3). The 95% CI of the low-risk group of the Apfel-score was very large (0–60%) as a result of the small number of patients included in our clinical sample (four patients).

Discussion

In our in-hospital population, the score of Apfel and colleagues[6] showed a better discriminating power and calibration properties for predicting PONV than the score of Sinclair, Chung and Mezei.[5]

In order to solve the "big little problem" of PONV,[9] we would like to confirm, by an appropriate trial, the effectiveness of the latest strategy suggested by White and Watcha.[3,4] An essential prerequisite is a model for predicting PONV which, as a diagnostic test, must be accurate, applicable to our specific patients and, ideally, easy to use.[10] Several reasons may account for the differences between the two models observed in our study.

First, the scores tested were evaluated in the range of patient, which we have in our clinical practice. Although our patients were mainly females, most other characteristics were comparable to those of the cross-validation study between Finland and Germany,[6] specially the Oulu validation set, and those from the study of Sinclair, Chung and Mezei.[5] The only difference is that the score by Sinclair and co-workers was developed on out-patients, while the score of Apfel *et al.* was developed on in-patients.

Second, the two models were also validated in a second, independent group of patients, as Sinclair, Chung and Mezei adhered to the method previously described by Apfel and co-workers published in 1998.[11,12]

Third, the patients underwent both the predictive tests and measurement of the outcome (reference standard), i.e., in this setting, the recording of PONV. Nevertheless, Sinclair, Chung and Mezei defined PONV as "any volunteered report of nausea or observed active retching or vomiting requiring antiemetics" whereas Apfel and co-workers, like their

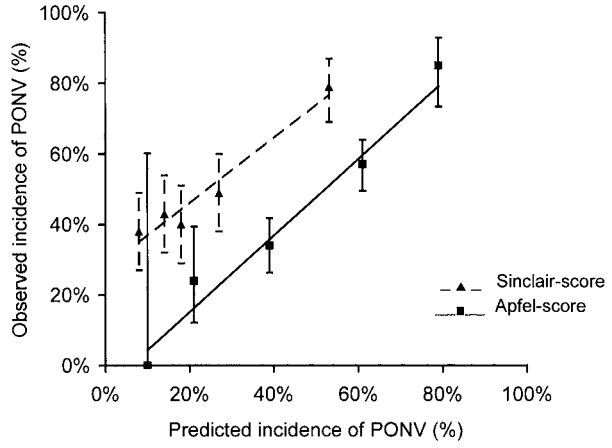

FIGURE 2 Correlation between predicted and observed incidences of postoperative nausea and vomiting (PONV) with 95% confidence intervals.

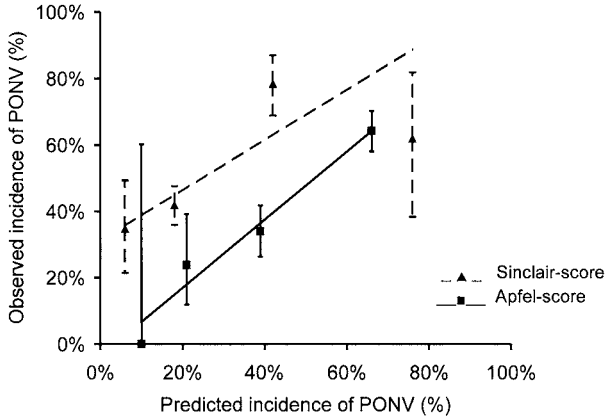

FIGURE 3 Correlation between predicted and observed incidences of postoperative nausea and vomiting (PONV) with 95% confidence intervals after new patient classification according to the suggestion of White and Watcha.[3,4]

predecessors[13,14] systematically asked for and recorded any episode of nausea and vomiting. This "volunteered report" might, to some extent, explain the disappointing calibration curves of the Sinclair-score in our study. The incidence of PONV ranges from 20 to 30% after general anesthesia[1] far more than the 9% reported by Sinclair, Chung and Mezei. Interestingly, this last figure is very close to the incidence of severe nausea (8%), observed in a similar population by Koivuranta and colleagues.[14] Neither patient characteristics, nor type of anesthesia and surgery could explain these differences in the incidence of PONV. Thus, it may be hypothesized that mainly patients with vomiting or severe nausea who voluntarily reported PONV were considered in the survey from Toronto.[5] This could have led to a strong underestimation of this outcome and the upward shift of the calibration line in our validation set.

Furthermore, the strong impact of surgery-related factors may have decreased the accuracy of the model from Sinclair *et al.* since these were not present in our population as shown by the multivariable analysis. Interestingly, the impact of the type of surgery among centres appears to be conflicting[5,12,15] with some studies showing no significant effect at all.[11,14] More interestingly, even if the type of surgery had a statistically significant effect, an operation independent score performed equally well as a more complex score considering the type of surgery.[12] Perhaps Apfel and coworker found the explanation for this phenomenon in their statistical modelling of virtual populations showing that the increase in discriminating power of a score is largest when the first variables are introduced so that more than four or five predictors do not lead to a better prediction unless much stronger predictors are identified.[16]

In summary, the simplified score of Apfel and colleagues offered better discriminating and calibrating properties than that proposed by Sinclair, Chung and Mezei. Furthermore, its simplicity makes it a useful tool for the assessment of the risk of PONV in clinical practice, and for research purposes.

References

1 *Watcha MF, White PF.* Postoperative nausea and vomiting. Its etiology, treatment, and prevention. Anesthesiology 1992; 77: 162–84.

2 *Sadhasivam S, Saxena A, Kathirvel S, Kannan TR, Trikha A, Mohan V.* The safety and efficacy of prophylactic ondansetron in patients undergoing modified radical mastectomy. Anesth Analg 1999; 89: 1340–5.

3 *White PF, Watcha MF.* Postoperative nausea and vomiting: prophylaxis versus treatment (Editorial). Anesth Analg 1999; 89: 1337–9.

4 *Watcha MF.* The cost-effective management of postoperative nausea and vomiting (Editorial). Anesthesiology 2000; 92: 931–3.

5 *Sinclair DR, Chung F, Mezei G.* Can postoperative nausea and vomiting be predicted? Anesthesiology 1999; 91: 109–18.

6 *Apfel CC, Läärä E, Koivuranta M, Greim C-A, Roewer N.* A simplified risk score for predicting postoperative

nausea and vomiting. Conclusions from cross-validations between two centers. Anesthesiology 1999; 91: 693–700.

7 *Eberhart LHJ, Seeling W, Bopp TI, Morin AM, Georgieff M.* Dimenhydrinate for prevention of post- operative nausea and vomiting in female in-patients. Eur J Anaesthesiol 1999; 16: 284–9.

8 *Hanley JA, McNeil BJ.* The meaning and the use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29–36.

9 *Fisher DM.* The "big little problem" of postoperative nausea and vomiting. Do we know the answer yet? (Editorial). Anesthesiology 1997; 87: 1271–3.

10 *Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB.* Diagnosis and screening, *In:* Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB (Eds.). Evidence-Based Medicine. How to Practice and Teach EBM, 2nd ed. Churchill Livingstone Inc., 2000: 67–93.

11 *Apfel CC, Greim CA, Haubitz I, et al.* A risk score to predict the probability of postoperative vomiting in adults. Acta Anaesthesiol Scand 1998; 42: 495–501.

12 *Apfel CC, Greim CA, Haubitz I, et al.* The discriminating power of a risk score for postoperative vomiting in adults undergoing various types of surgery. Acta Anaesthesiol Scand 1998; 42: 502–9.

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, Snäre L, Alahuhta S.* A survey of postoperative nausea and vomiting. Anaesthesia 1997; 52: 443–9.

15 *Cohen MM, Duncan PG, DeBoer DP, Tweed WA.* The postoperative interview: assessing risk factors for nausea and vomiting. Anesth Analg 1994; 78: 7–16.

16 *Apfel CC, Kranke P, Greim C-A, Roewer N.* What can be expected from risk scores for predicting postoperative nausea and vomiting? Br J Anaesth 2001; 86: 822–7.