

# Impact of Risk Assessments on Prophylactic Antiemetic Prescription and the Incidence of Postoperative Nausea and Vomiting

## *A Cluster-randomized Trial*

Teus H. Kappen, M.D., Karel G.M. Moons, Ph.D., Leo van Wolfswinkel, M.D., Ph.D.,  
Cornelis J. Kalkman, M.D., Ph.D., Yvonne Vergouwe, Ph.D., Wilton A. van Klei, M.D., Ph.D.

### ABSTRACT

**Background:** Clinical prediction models have been shown to have moderate sensitivity and specificity, yet their use will depend on implementation in clinical practice. The authors hypothesized that implementation of a prediction model for postoperative nausea and vomiting (PONV) would lower the PONV incidence by stimulating anesthesiologists to administer more “risk-tailored” prophylaxis to patients.

**Methods:** A single-center, cluster-randomized trial was performed in 12,032 elective surgical patients receiving anesthesia from 79 anesthesiologists. Anesthesiologists were randomized to either exposure or nonexposure to automated risk calculations for PONV (without patient-specific recommendations on prophylactic antiemetics). Anesthesiologists who treated less than 50 enrolled patients were excluded during the analysis to avoid too small clusters, yielding 11,613 patients and 57 anesthesiologists (intervention group: 5,471 and 31; care-as-usual group: 6,142 and 26). The 24-h incidence of PONV (primary outcome) and the number of prophylactic antiemetics administered per patient were studied for risk-dependent differences between allocation groups.

**Results:** There were no differences in PONV incidence between allocation groups (crude incidence intervention group 41%, care-as-usual group 43%; odds ratio, 0.97; 95% CI, 0.87–1.1; risk-dependent odds ratio, 0.92; 95% CI, 0.80–1.1). Nevertheless, intervention-group anesthesiologists administered more prophylactic antiemetics (rate ratio, 2.0; 95% CI, 1.6–2.4) and more risk-tailored than care-as-usual-group anesthesiologists (risk-dependent rate ratio, 1.6; 95% CI, 1.3–2.0).

**Conclusions:** Implementation of a PONV prediction model did not reduce the PONV incidence despite increased antiemetic prescription in high-risk patients by anesthesiologists. Before implementing prediction models into clinical practice, implementation studies that include patient outcomes as an endpoint are needed. (*ANESTHESIOLOGY* 2014; 120:343–54)

TO support physicians in their decision making, clinical guidelines increasingly include prediction models or risk scores in their recommendations.<sup>1,2</sup> The growing popularity of prediction models can be explained by their perceived objectivity in modeling complex interactions to predict a patient's risk, in contrast to physicians' clinical judgments, which are typically heuristic and harder to deconstruct. The effects of a prediction model on clinical processes and patient outcomes should be quantified in a so-called “impact analysis,” where the model is implemented in daily practice and compared with care-as-usual.<sup>3–6</sup>

Current guidelines on the management of postoperative nausea and vomiting (PONV) recommend risk-tailored, prophylactic treatment based on risk estimates from a prediction model, to prevent unnecessary costs and possible side effects, in contrast to administering multiple drugs to all patients.<sup>7</sup> Although several PONV prediction models are available,<sup>8–11</sup>

#### What We Already Know about This Topic

- Guidelines on the management of postoperative nausea and vomiting recommend risk-tailored prophylactic treatment based on risk estimates from a prediction model

#### What This Article Tells Us That Is New

- In a single-center, cluster-randomized trial (n = 12,032 patients and 79 anesthesiologists), implementation of a postoperative nausea and vomiting prediction model without specific treatment recommendation did not reduce the incidence of postoperative nausea and vomiting (odds ratio, 0.97; 95% CI, 0.87–1.10)

their actual impact on clinical practice is still being questioned.<sup>12–14</sup> Several studies demonstrated improved guideline adherence when a PONV prediction model was implemented although their effect on clinical practice was limited.<sup>15–17</sup> However, comparative randomized studies assessing the actual impact of risk-dependent prophylaxis on the incidence of

Presented in part at the Annual Meeting of the American Society of Anesthesiologists, Orlando, Florida, October 21, 2008.

Submitted for publication October 22, 2012. Accepted for publication August 15, 2013. From the Department of Anesthesiology, Division of Anesthesiology, Intensive Care and Emergency Medicine (T.H.K., L.v.W., C.J.K., W.A.v.K., K.G.M.M.) and Julius Center for Health Sciences and Primary Care (K.G.M.M., Y.V.), University Medical Center Utrecht, Utrecht, The Netherlands.

Copyright © 2013, the American Society of Anesthesiologists, Inc. Lippincott Williams & Wilkins. *Anesthesiology* 2014; 120:343–54

PONV are rare. Without such studies, one still cannot be confident that PONV prediction models will outperform clinical judgment and improve patient outcomes.

The current study was a cluster-randomized trial, in which we compared a group of physicians who were randomly “exposed” to model-based risk estimates of PONV for their patients with a control group of physicians who provided care-as-usual. We hypothesized that systematic implementation of a validated PONV prediction model would result in improved patient outcome by lowering the incidence of PONV, as a result of an increase in risk-tailored antiemetic prophylactic treatment by physicians. Hence, we also used the current experience to address advantages and disadvantages of such impact studies.

## Materials and Methods

### Design and Participants

Between March 16, 2006 and December 21, 2007, a single-center, cluster-randomized trial was performed to study the effects of implementing a prediction model for PONV on the incidence of PONV and on the administration of antiemetic prophylaxis (Clinicaltrials.gov; Identifier NCT00293618). To prevent contamination and therefore dilution of the implementation effect, the study was cluster-randomized on the physician level rather than on the patient level.<sup>3,18,19</sup> Anesthesiologists of a Dutch university hospital (University Medical Center Utrecht, Utrecht, The Netherlands) were randomly assigned to exposure to predicted risks of PONV as calculated by the prediction model (intervention group) or not (care-as-usual group).

All anesthesiologists and senior residents, henceforth referred to as “anesthesiologists,” were enrolled in the study and randomized by one of the authors. Case mixes among anesthesiologists were expected to differ due to differences in their professional profile, such as differences in experience level or anesthesia subspecialties. Therefore, at the start of the study, permuted-block randomization was used (block size according to the strata; 1:1 allocation ratio, using PASS software; NCSS, Kaysville, UT) to stratify on anesthetic experience (senior resident, junior attending, senior attending) and anesthesia subspecialty (no subspecialty, cardiac anesthesia, pediatric anesthesia). Anesthesiologists could enter or leave the study due to initiation or termination of their employment. Anesthesiologists who entered after the start of the study were randomized only in a stratified way when a sufficiently large block was available, otherwise simple randomization was applied. Allocation sequences were generated by the second author (K.G.M.M.), who was not involved in patient care, and automatically assigned to the name of the anesthesiologist to ensure concealment. The generated list of the assignment of anesthesiologists was then given to the first author (T.H.K.) who enrolled and informed the anesthesiologists.

All adult patients undergoing general anesthesia for elective, noncardiac surgery who had visited the outpatient preoperative evaluation clinic were eligible for this study. For

elective surgery, 98% of patients typically visit the preanesthesia evaluation clinic before their procedure starts. Exclusion criteria were pregnancy, postoperative admission to the intensive care unit, overnight ventilation at the postanesthesia care unit, and inability to communicate in Dutch or English. At the start of the study, patients undergoing intracranial surgery were no longer transferred to the intensive care unit and did no longer require postoperative mechanical ventilation, in contrast to when the study protocol was written. Therefore, the exclusion criterion for intracranial surgery in the study protocol was changed to admission to the intensive care unit.

All eligible patients from the time of study initiation were automatically included using the anesthesia information management system. During the enrollment phase of the study, it became apparent that some anesthesiologists would treat no or only a few patients. Consequently, the number of patients for some anesthesiologists would be too small for an analysis with the anesthesiologists as clusters, for mixed-effects regression models would not converge. Therefore, the study protocol was amended by excluding anesthesiologists from the analysis when they treated less than 50 enrolled patients during the entire study period to enable cluster-based analysis of the trial as it was originally planned.

According to Dutch law, research protocols that do not subject patients to a particular treatment or that require them to behave in a particular way do not apply to the Medical Research Involving Human Subjects Act. As the decision support tool in our study protocol only provided evidence-based information to physicians, the institutional ethical review board waived the need for individual informed consent and approved the study protocol (Medical Ethics Review Board, University Medical Center Utrecht, 05-288).

### The Prediction Model

The implemented prediction model was originally developed in a population of a different university hospital in The Netherlands and had already been externally validated.<sup>20</sup> The model was subsequently updated and optimized for implementation in the University Medical Center Utrecht, where the current study took place.<sup>21</sup> The model consisted of seven predictor variables: age; sex; current smoking; type of surgery; inhalational anesthesia; ambulatory surgery; and history of motion sickness or PONV (for full model description see table 1).

### Interventions

**Intervention Group.** To study the effect of systematically presenting the patient’s PONV risk to the responsible anesthesiologist, we implemented the prediction model as an “assistive” decision support tool. The decision support tool was integrated into our custom-made anesthesia information management system (Vierkleurenpen® software; CarePoint Nederland BV, Ede, The Netherlands), written by one of the authors (L.v.W.). In our hospital, the individual names of each anesthesia team member are registered in the anesthesia information management system at the start of the anesthesia case.

**Table 1.** Regression Coefficients of the Prediction Model for PONV

Predictor	Updated Model
	Odds Ratio (95% CI)
Age (yr)	0.98 (0.98–0.99)
Female sex	1.44 (1.14–1.82)
Current smoking	0.61 (0.48–0.77)
History of PONV/motion sickness	1.82 (1.44–2.31)
Surgery with a high PONV risk*	1.62 (1.14–2.30)
Inhalational anesthesia†	1.42 (1.12–1.79)
Outpatient surgery	0.31 (0.24–0.41)
Intercept‡	1.13 (0.73–1.74)
Model performance characteristics§	
Model discrimination as c-statistic (95% CI)	0.68 (0.66–0.70)
Calibration slope (95% CI)	1.00 (0.89–1.10)
Calibration intercept	0.34

Probability of PONV as estimated by the model =  $1/(1 + \exp(-(-0.12 - 0.017 \times \text{age} + 0.36 \times \text{female sex} - 0.50 \times \text{current smoking} + 0.60 \times \text{history of PONV or motion sickness} + 0.48 \times \text{surgery with a high PONV risk} + 0.35 \times \text{inhalational anesthesia} - 1.16 \times \text{outpatient surgery})))$ .

\* Definition of this predictor was “abdominal or middle ear surgery.” † As compared with intravenous anesthesia using propofol. ‡ For the intercept, the second column represents the baseline odds, not the odds ratio. § Model performance was validated in a subset of patients (between March 2006 and February 2007) treated by anesthesiologists of the care-as-usual group. (Kappen *et al.* Med Decis Making 2012; 32:E1–10.<sup>21</sup>)

exp = exponential; PONV = postoperative nausea and vomiting.

When the anesthesiologist was part of the intervention group, the decision support tool presented the patient’s calculated PONV risk on the computer screen of the anesthesia information management system during the rest of the anesthesia case. Anesthesiologists then decided whether which and how many prophylactic antiemetics would be administered in view of the patient’s individual risk. The presented risk was thus not accompanied with a specific therapeutic recommendation.

Anesthesiologists of the intervention group were provided with several consecutive educational sessions before patient enrollment, at the start, and throughout the study period. These sessions aimed to inform the intervention group about the study background, how the prediction model estimated a patient’s individual risk of PONV, and how the local protocol on antiemetic prophylaxis could be used according to that predicted risk. The local protocol was based on the six-factorial trial of Apfel<sup>22</sup> and consisted of the dosage, timing, and efficacy of prophylactic antiemetic drugs and the use of total intravenous anesthesia (see Materials and Methods, Outcome and Follow-up, fifth paragraph). Although the efficacy of different antiemetics strategies was discussed during the educational sessions, there was no specific recommendation made on how many antiemetic interventions should be applied at a particular predicted PONV risk.

Anesthesiologists of the intervention group were informed of the allocation status of their colleagues to promote discussion among anesthesiologists of the intervention group on how to use the model and its predictions. They were

instructed to avoid discussing PONV with anesthesiologists randomized to the control group. To enable anesthesiologists of the intervention group to reflect on their individual prophylactic management of PONV, they received individualized feedback *via* email after the first 12 months of study, which had been planned before the start of the study. This included the incidence of PONV among the patients they had treated, the overall (hospital wide) incidence of PONV, and the quantity of antiemetics administered.

**Control Group.** Anesthesiologists of the care-as-usual group were not exposed to the patient’s calculated PONV risk. At the start of the study, they only were informed about the goal of the study and their randomization status. Although anesthesiologists of the care-as-usual group were not actively informed of the allocation status of their colleague physicians, additional masking was considered impossible. As antiemetic management was not standardized in any of the two allocation groups, anesthesiologists of the care-as-usual group were simply asked to manage PONV as usual. At that time, the existing, local protocol for administration of PONV prophylaxis only included a preferable order for antiemetic drugs, their dosage, and timing of administration (see Materials and Methods, Outcome and Follow-up, fifth paragraph).

### Outcome and Follow-up

As recommended in various guidelines of prediction-model impact studies,<sup>3,18</sup> implementation of the prediction model was studied in two steps: the effects of the prediction model on patient outcome (the incidence of PONV) as the primary outcome and the change in physician behavior (administration of risk-dependent PONV prophylaxis) caused by the prediction model as the secondary outcome. The other secondary outcomes as stated in the clinical trial registration—that is cost-effectiveness and attitudes of physicians toward prediction models—were considered to be beyond the scope of this article and will be discussed in subsequent articles.

The primary outcome PONV was defined as the occurrence of at least one of the following events within the first 24 h after surgery: an episode of nausea, an episode of vomiting, or the administration of any rescue antiemetic. For nausea, the patient was asked to rate their feeling of nausea on a 3-point verbal rating scale (no/yes, a bit/yes, definitely), and for the analysis, the variable was dichotomized to any nausea (no/yes). Vomiting was defined as the expulsion of gastric contents and was recorded as a binary outcome (no/yes). Research nurses collected data on the occurrence of postoperative nausea using a validated questionnaire.<sup>11,23</sup>

Data were collected at the postanesthesia care unit (30 and 60 min after arrival and when leaving the unit) and after 24 h postsurgery on the ward, or by telephonically when patients had already been discharged. The outcome variable for PONV was coded as missing when any of the follow-up measurements had not been completed. Although research nurses were unlikely to be aware of a patient’s allocation status due to high patient volumes,

active masking of allocation was considered impossible and therefore not performed.

The change in physician behavior caused by the prediction model was defined as the difference in the rate of administration of PONV prophylaxis between study groups. The rate of administration of PONV prophylaxis was defined as the number of interventions that an anesthesiologist applied to a patient with the aim to prevent PONV, that is, the number of prophylactic antiemetics per patient. Administration of ondansetron, droperidol, dexamethasone, or a combination as well as selecting total intravenous anesthesia instead of inhalational anesthesia was considered as prophylactic antiemetics, and their administration was recorded in the anesthesia information system.<sup>23</sup>

Dosage of prophylactic antiemetic drugs was according to the existing local protocol: (1) ondansetron 4 mg IV, 30 min before emergence of anesthesia; (2) droperidol 1.25 mg IV, 30 min before emergence of anesthesia; and (3) dexamethasone 4 mg IV, after induction of anesthesia. At the postanesthesia care unit and the ward, the PONV protocol consisted of rescue treatment with an antiemetic drug: either one of the above antiemetics drugs if not previously administered or metoclopramide 20 mg IV. There was no active surveillance of adverse events during this study.

### Statistical Analysis

Because we aimed to assess the impact of implementing a PONV prediction model on the actual patient outcome (PONV), sample size was based on an estimated PONV incidence of 30% in the control group<sup>8,24</sup> and a relative risk reduction of 25% per antiemetic.<sup>22,25</sup> As intervention-group anesthesiologists were expected to provide more than one antiemetic to high-risk patients, the overall relative risk reduction for the intervention group was estimated at 33%, that is, an absolute risk reduction of 10%. Detection of this 10% reduction in a randomized trial without cluster randomization would require 295 patients per group, using a two-sided  $\alpha$  of 0.05 and power of 0.80. The sample size was adjusted for cluster randomization using an inflation factor based on an average cluster size of 175 patients, and an intraclass correlation coefficient of 0.1, resulting in 5,430 patients for each group.<sup>26,27</sup> The sample size was considered sufficient for physician behavior, as less power was required to detect a difference in prescription of prophylactic antiemetics between allocation groups. Hence, approximately 11,000 patients were expected to be required.

Analysis was performed under the intention-to-treat principle. All statistical analyses were performed with the use of R software (version 2.14.0\*). Statistical significance was defined as a two-sided  $\alpha$  of 0.05. Continuous variables were visually assessed for a normal distribution using histograms, Quantile-Quantile (QQ)-plots. Parametric variables were expressed as means with SDs, nonparametric variables were

expressed as medians with interquartile ranges, and discrete variables were expressed as numbers with percentages.

As this was a cluster-randomized trial, mixed-effects regression analyses were used to take clustering into account: logistic regression for the incidence of PONV and Poisson regression for the number of prophylactic antiemetics (the function for generalized linear mixed-effects models [glmer] from the lme4 library in R software). Results of regression analyses were presented as odds ratios with 95% CIs or rate ratios with 95% CI.

Fixed effects included allocation group, predicted risk of PONV, interaction between allocation group and predicted risk, and study time. The interaction term was included to quantify to what extent the difference in treatment effect (between intervention and care-as-usual) differed across predicted risks; for example, an odds ratio less than 1 would signify that a reduction in PONV due to the directive approach was greater in patients with higher risks. Study time was included in the model to adjust for a possible learning effect for anesthesiologists exposed to the prediction model.

In addition to the fixed effects, random effects were included for the intercept, allocation group, predicted risk, interaction between allocation group and predicted risk, and study time. A random intercept was included for anesthesiologists to account for small differences in PONV risks between their individual patient populations. Random slopes were included to account for different PONV prophylaxis strategies among anesthesiologists. A random slope was included for study time as the possible learning effect was not expected to be similar for each individual anesthesiologist. The intra-class correlation coefficient was calculated for the PONV incidence (using the “glmer” code in R) to quantify the amount of clustering at the anesthesiologists’ level, which represents the resemblance of patients treated by the same anesthesiologist.<sup>28</sup>

Although randomization was stratified for anesthesia subspecialty, some anesthesiologists may still have worked more with particular surgical specialties, treating patients with particular baseline risks of PONV. Due to a relatively small number of anesthesiologists (expected  $n = 60$ ) treating a large number of patients, differences in case mix between anesthesiologists were expected to be magnified in observed baseline characteristics of patients. As the prediction model already included the most important patient characteristics, the mixed-effects analyses were automatically adjusted for differences in patient characteristics by inclusion of the predicted risk as a variable. Therefore, additional analyses for both primary and secondary outcomes were performed to adjust for case-mix differences by including procedure-specific variables as fixed effects: surgical specialty and type of patient (ambulatory surgery yes/no). As the differences in baseline characteristics of patients were expected, these additional adjustments were amended to the study protocol before the start of the analysis.

Before multivariable modeling, all continuous variables were tested for nonlinearity using restricted cubic splines, including predicted PONV risk.<sup>29</sup> Missing data were multiple

\* Available at: <http://www.r-project.org>. Accessed May 10, 2013.



imputed ( $n = 10$ ) using a regression approach in R (the function for multiple imputation [aregImpute] from the library Hmisc in R software). Imputation of missing study variables was based on predictors, outcome variables, and other perioperative data.<sup>30–33</sup> As PONV was coded missing when any of the follow-up measurements was incomplete, nonmissing follow-up measurements of PONV were added to the imputation process to serve as auxiliary variables to impute missing values for PONV. Subsequently, the imputed values for PONV were included into the mixed-effects regression model, instead of deleted. The anesthesiologists were added as an extra variable in the imputation model to take into account the multilevel structure of the data.

Three *post hoc* sensitivity analyses were performed to test the robustness of the results on three different areas of possible uncertainty. The three areas of possible uncertainty were the effects of multiple imputation of missing values; the definition of the outcome variable PONV and its consequence for the results; and the exclusion of anesthesiologists who treated less than 50 enrolled patients during the study period.

The sensitivity analysis on missing values was performed using a similar mixed-effects model as the main analysis, but on complete cases only. For PONV, patients with any missing follow-up time points were discarded during the complete case analysis. For the analysis on prophylactic antiemetic interventions, all data were available and no patients were discarded.

For the sensitivity analysis on the primary outcome definition, we restricted the definition of PONV to serious nausea or vomiting at any of the follow-up time points. In contrast to the main analysis, minor nausea (the middle category of the 3-point nausea scale) and the use of rescue antiemetics were not considered PONV for this sensitivity analysis.

As the exclusion of anesthesiologists with a low patient count during the study resulted in reduction in a sample size, and as the discarded group of patients may be selective, a sensitivity analysis examined the impact of this exclusion. The reason to exclude anesthesiologists with a low patient count was to enable performing a cluster-based analysis, that is, the mixed-effects models. Therefore, this sensitivity analysis consisted of conventional logistic and Poisson regression analyses in all randomized patients, that is, without the random effects and the anesthesiologists with a low patient count not excluded. Furthermore, anesthesiologists who treated no patients during the study period were still excluded from the analysis.

## Results

A total of 79 anesthesiologists were randomized, who together treated 12,032 enrolled patients (fig. 1). Of the 79 anesthesiologists, 6 were excluded as they treated no patients during the study period and 16 were excluded as they treated less than 50 patients per individual during the study period (in total 397 patients). An additional 22 patients were excluded because of a technical error in the anesthesia information

management system. This resulted in 11,613 patients treated by 57 anesthesiologists (31 in intervention group and 26 in care-as-usual group) to be analyzed. Anesthesiologists of the intervention group treated fewer patients per anesthesiologist (median 162 patients per anesthesiologist, interquartile range 120–207) than anesthesiologists of the care-as-usual group (median 236 patients per anesthesiologist, interquartile range 128–300; table 2).

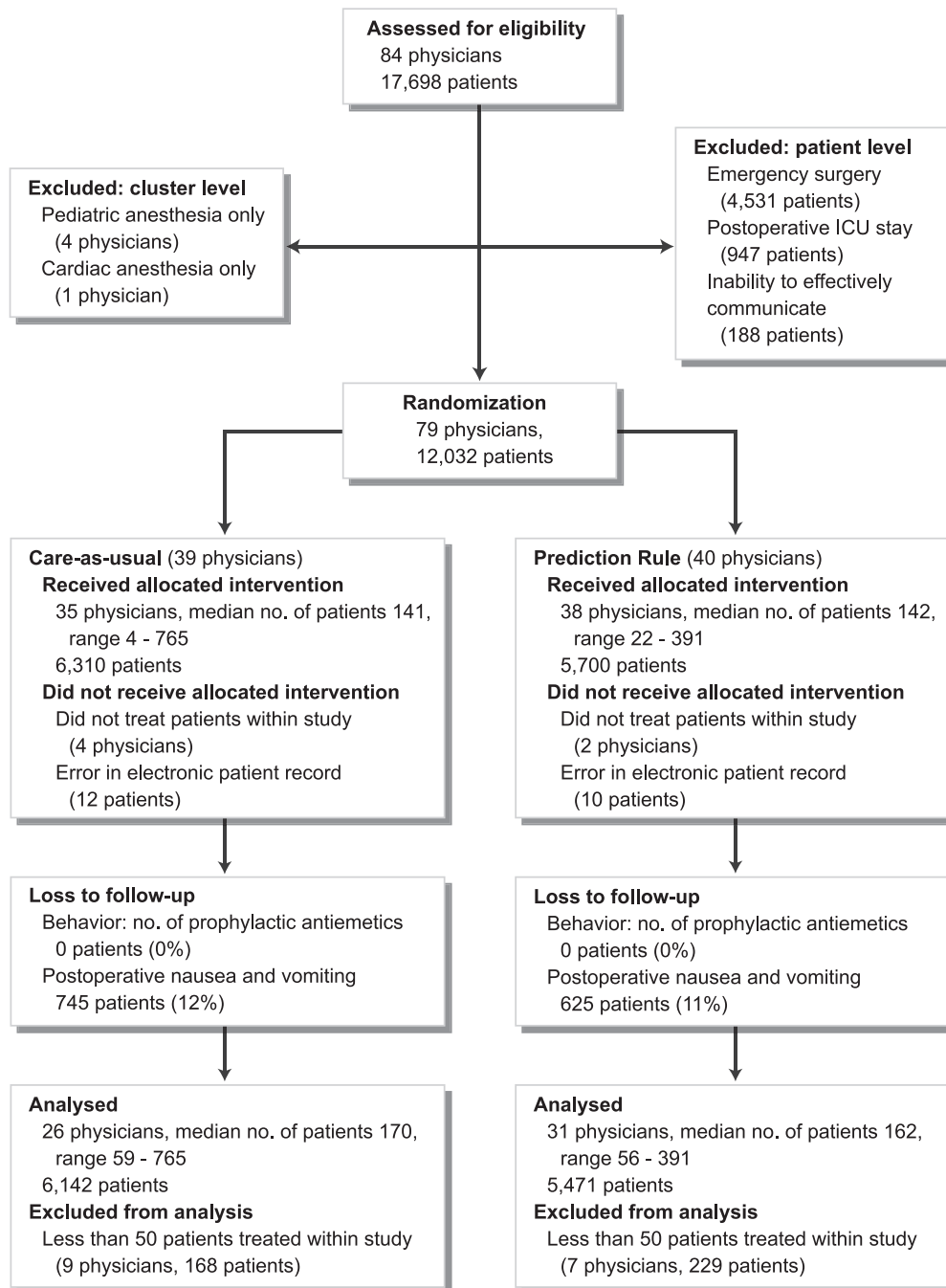
Of the 11,613 patients, 5,471 (47%) were treated by intervention-group anesthesiologists and 6,142 (53%) by “care-as-usual” anesthesiologists (table 3). The patients’ mean ages and sex distribution were comparable. The intervention group included more outpatients (38 *vs.* 27%) and fewer procedures with a high PONV risk (11 *vs.* 15%); all other predictors were comparable between the two groups. The predicted PONV risk was slightly lower in the intervention group: 37% (SD, 15%) *versus* 39% (SD, 15%). No adverse events were reported by the study participants.

In total, 80% of all follow-up measurements were completed (intervention group 80% and care-as-usual group 79%), with 88% of all patients having at least one follow-up measurement completed (intervention group 89% and care-as-usual group 88%). As the primary outcome for a single patient was considered missing when any of the follow-up measurements were missing, 70% of patients had all follow-up measurements completed ( $n = 8,104$ ), whereas the remaining 30% had their outcome variables imputed. The crude effect of prediction model implementation on patient outcome is shown in table 4. The crude incidence of PONV within the first 24 h after surgery was 41% for patients in the intervention group and 43% in the care-as-usual group. Intraclass correlation was low (0.020).

Differences in the occurrence of PONV were small and not statistically significant between intervention and care-as-usual groups (odds ratio for allocation group, 0.97; 95% CI, 0.87–1.1; odds ratio for the interaction term of allocation group and predicted risk, 0.92; 95% CI, 0.80–1.1). The absence of statistical significance is reflected in figure 2A as 95% CIs of both groups almost fully overlap. Prespecified adjustment for baseline characteristics did not change these results and inferences. Moreover, the sensitivity analyses on missing data, the outcome definition, and exclusion of anesthesiologists and patients showed small and nonsignificant effects on the PONV incidence, similar to the main analyses. Numerical descriptions of both adjusted and unadjusted models, as well as the three sensitivity analyses, can be found in table 5.

The crude effect of the group assignment on physician behavior is shown as the percentage of patients who received a particular number of administered prophylactic antiemetics (table 4). The number and type of prophylactic antiemetics were documented for all patients.

In the main analysis, anesthesiologists of the intervention group administered more antiemetic prophylaxis than anesthesiologists of the care-as-usual group (rate ratio for allocation



**Fig. 1.** Flow diagram of the progress of anesthesiologists and patients through the phases of the study. Anesthesiologists represent either attending anesthesiologists or senior anesthesiology residents. ICU = intensive care unit.

group, 2.0; 95% CI, 1.6–2.4). Moreover, when administering antiemetic prophylaxis, intervention-group anesthesiologists discriminated more between patients with a high or low predicted risk than anesthesiologists of the care-as-usual group (rate ratio for the interaction term of allocation group and predicted risk, 1.6; 95% CI, 1.3–2.0; fig. 2B). Compared with anesthesiologists of the care-as-usual group, intervention-group anesthesiologists administered a higher number of prophylactic antiemetics to patients with a high predicted risk, whereas low-risk patients received a lower number of antiemetics. Both

overall and risk-dependent differences in physician behavior were statistically significant (fig. 2B; 95% CI areas of both groups cross each other and only slightly overlap). Neither adjustment for baseline differences nor the three sensitivity analyses changed the results and their inferences (table 6).

## Discussion

Following recent guidelines on studying the impact of clinical prediction models, we performed a cluster-randomized

**Table 2.** Baseline Characteristics of Anesthesiologists

	Care-as-Usual Group (n = 26)	Intervention Group (n = 31)
Number of patients treated, median (IQR)	170 (128–300)	162 (120–207)
Age*, yr, mean (SD)	43 (9)	43 (10)
Female sex, n (%)	11 (42)	15 (48)
Entered during study, n (%)	5 (19)	8 (26)
Left department during study, n (%)	4 (15)	7 (23)
Specialty: cardiac anesthesia*, n (%)	2 (8)	2 (6)
Pediatric anesthesia*, n (%)	3 (12)	4 (13)
Pain medicine*, n (%)	1 (4)	4 (13)
Senior resident*, n (%)	12 (46)	11 (35)
Entered as senior resident during study, n (%)	4 (15)	4 (13)
Senior resident became attending anesthesiologist during study, n (%)	5 (19)	4 (13)

\* Characteristic was documented at the moment of inclusion in the study.

IQR = interquartile range.

**Table 3.** Baseline Characteristics of Patients

	N*	Care-as-Usual Group (n = 6,142)	Intervention Group (n = 5,471)
Age, yr, mean (SD)	11,613	51 (17)	50 (17)
Female sex, n (%)	11,613	3,307 (54)	2,988 (55)
Outpatients, n (%)	11,613	1,646 (27)	2,086 (38)
ASA class, n (%)	11,325		
1		2,211 (37)	1,985 (37)
2		3,111 (52)	2,765 (52)
3		648 (11)	581 (11)
4		15 (0)	9 (0)
Current smoking, n (%)	11,134	1,765 (30)	1,598 (30)
Surgery with a high PONV risk, n (%)	10,545	807 (15)	541 (11)
History of PONV/motion sickness, n (%)	8,837	1,511 (33)	1,397 (33)
Use of inhalational anesthesia†, n (%)	11,613	3,212 (52)	3,321 (61)
Predicted risk of PONV, %, mean (SD)	11,613	39 (15)	37 (15)
Apfel risk score‡, %, mean (SD)	11,613	42 (19)	42 (19)
Operation duration, min, median (IQR)	11,613	101 (63–167)	95 (59–151)

\* N represents the total number of nonmissing observations for that characteristic. † As compared with intravenous anesthesia using propofol. ‡ For calculation of the Apfel risk score, the observed use of postoperative opioids was used instead of the expected.

ASA = American Society of Anesthesiologists; IQR = interquartile range; PONV = postoperative nausea and vomiting.

trial on the implementation of a validated prediction model for PONV. We quantified the effects of implementing such a prediction model on both physician behavior and patient outcome. Patients did not have a substantially lower incidence of PONV when their anesthesiologists were provided with an intraoperative predicted PONV risk despite an increased administration of risk-tailored prophylaxis by these anesthesiologists. In other words, anesthesiologists of the intervention group administered more prophylactic antiemetics to patients at higher risk and fewer antiemetics to patients at low risk in comparison with their colleagues of the care-as-usual group. However, the tailored prescription of antiemetics in the intervention group did not result in a substantially lower incidence of PONV.

The discrepancy in results between patient outcome and physician behavior was unexpected. At the start of the study, we assumed that all conditions to proceed with an impact

study of our PONV prediction model had been met: the implemented prediction model had been externally validated, several clinical guidelines advocated the use of prediction models for PONV, and the efficacy of prophylactic antiemetics had been well established by several randomized clinical trials and meta-analyses.<sup>22,25,34–36</sup> Because this study showed that application of the PONV prediction model did not translate into a clear benefit for patients, it is conceivable that one or more of these presumptions were wrong.

First, the predictive performance of the prediction model may actually have been insufficient to improve clinical decision making. In previous validation studies, predictive performance of all existing PONV models was typically moderate (c-statistic approximately 0.70).<sup>9,11</sup> Our prediction model had comparable discrimination (c-statistic of 0.68), and it slightly underestimated the actual PONV risk.<sup>21</sup> With a moderate predictive performance, decisions based on the

**Table 4.** Primary and Secondary Outcomes by Deciles of Predicted Risk of PONV

Predicted Risk Deciles, %	Incidence of PONV			No. of Prophylactic Antiemetics*										
	N†	Care-as-Usual Group‡ (n = 6,142)	Intervention Group‡ (n = 5,471)	N†§	Care-as-Usual Group   (n = 6,142)					Intervention Group   (n = 5,471)				
					0	1	2	3	4	0	1	2	3	4
6–18	790	67 (18)	82 (20)	1,162	75	24	1	0	0	82	16	2	0	0
18–25	750	102 (28)	102 (27)	1,163	62	33	5	0	0	69	27	4	0	0
25–30	795	101 (26)	113 (28)	1,163	53	41	6	0	0	56	38	5	1	0
30–33	795	123 (29)	117 (31)	1,159	48	45	6	0	0	52	40	8	1	0
33–37	823	165 (38)	133 (34)	1,161	45	47	7	1	0	51	37	10	2	0
37–42	831	199 (45)	166 (43)	1,160	42	52	6	0	0	42	42	13	2	0
42–47	815	195 (44)	163 (44)	1,159	42	48	9	0	0	39	40	18	3	0
47–52	809	251 (55)	202 (58)	1,163	35	51	13	1	0	32	42	22	4	0
52–59	835	294 (64)	216 (58)	1,162	31	50	16	3	0	21	44	29	5	1
59–81	861	337 (69)	253 (68)	1,161	23	54	21	2	0	12	36	42	9	1
Overall	8,104	1,834 (43)	1,547 (41)	11,613	45	45	9	1	0	47	36	15	3	0

\* The sum of the prophylactic use of ondansetron, droperidol, dexamethasone, and/or total intravenous anesthesia. † N represents the total number of non-missing observations per decile of predicted risk. ‡ Data represent absolute numbers of PONV (%). § For the no. of prophylactic antiemetics, all observations were available. || Data represent percentages within risk decile.

PONV = postoperative nausea and vomiting.

model may not have been superior to care-as-usual, that is, clinical judgment.

Second, despite a statistically significant impact on physician behavior, the absolute impact was relatively small. In this study, four different prophylactic antiemetics were available. However, anesthesiologists of the intervention group mostly administered up to two antiemetics to high-risk patients, whereas more than two antiemetics may be indicated.<sup>22</sup> From the results of our study, we cannot infer why anesthesiologists of the intervention group were reluctant to administer more than two antiemetics. Nonetheless, one might expect a decrease in PONV occurrence when anesthesiologists of the intervention group administered more prophylactic antiemetics to their high-risk patients.

Third, proven efficacy in randomized clinical trials is no guarantee for effectiveness in daily practice.<sup>37–40</sup> In previous efficacy trials, prophylactic antiemetics reduced the risk of PONV by approximately 30%, with a wide variation between different studies.<sup>25</sup> However, study populations of these efficacy trials were often restricted to specific high-risk groups, such as women undergoing laparoscopic procedures. As the current study included a large sample of surgical patients with minimal inclusion restrictions, the actual effectiveness of “proven” prophylactic antiemetics may indeed be lower than reported earlier.

The primary goal of the current study was to quantify the effects of implementing a clinical prediction model on patient outcome (PONV) and to interpret these effects in view of the change in physician behavior caused by the prediction model, that is, administration of antiemetics.

The absence of a direct link between increased administration of antiemetics and reduced PONV should not be interpreted as evidence that PONV prediction models—or prediction models in general—are ineffective in clinical

practice. This study is only one example, and its limitations should be considered when applying the results to different settings.

First, although we assessed a large number of outcomes, this implementation study included a single center.<sup>3</sup> Second, despite a large study population, differences in baseline characteristics occurred, probably due to cluster randomization. By randomizing a relatively small sample of anesthesiologists, differences in anesthesiologists’ baseline characteristics may have been magnified at the patient level although adjustment for baseline differences in patient case mix among anesthesiologists did not change the results.

Third, anesthesiologists were not naive to prophylactic PONV management. As table 4 and figure 2A clearly indicate, anesthesiologists of the care-as-usual group also provided prophylaxis to their patients in a risk-dependent manner, reducing the possible impact of a prediction model on the incidence of PONV. It is conceivable that anesthesiologists were already able to (heuristically) identify high-risk patients before the start of the study, as a consequence of training, clinical experience, medical literature, and the pre-existent local PONV protocol. In addition, there might have been increased knowledge and experience with the use of risk scores for PONV prophylaxis, as the implemented prediction model was updated and optimized in the same hospital and the anesthesiologists may already have been using other risk scores before the start of the study. Therefore, a Hawthorne effect cannot be ruled out as a possible explanation for the results of this study. An alternative explanation may be that some contamination may have occurred, as anesthesiologists of the intervention group may occasionally have discussed PONV and its prophylactic treatment with anesthesiologists of the care-as-usual group. However, true



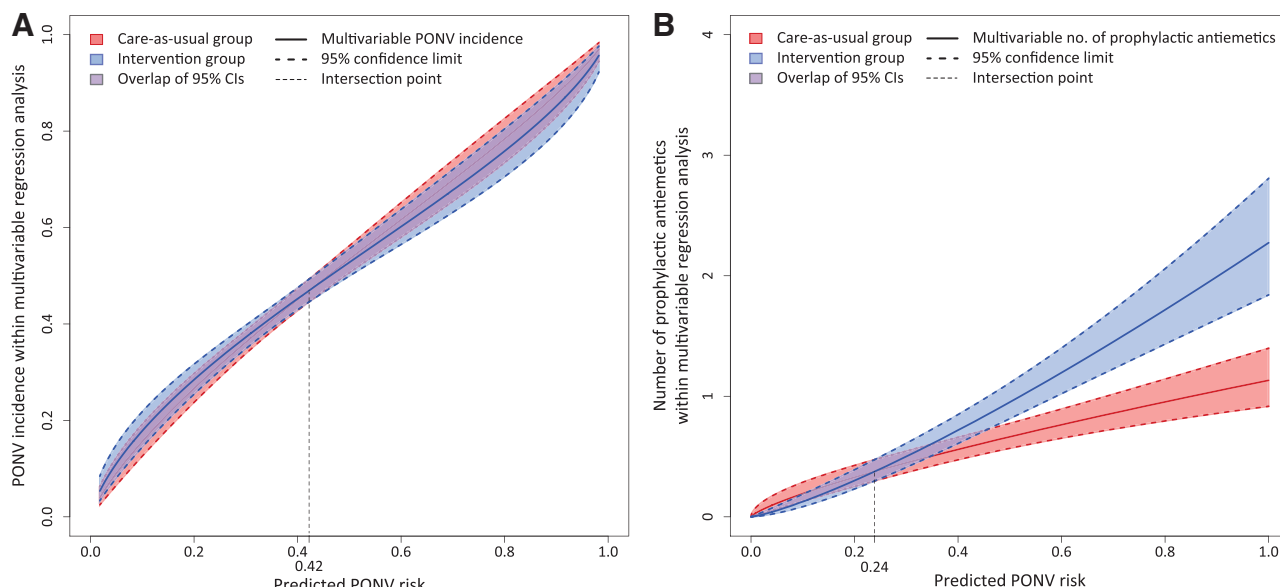
**Table 5.** Mixed Effects Regression Analysis on the Incidence of PONV

	Main Analysis*†	Sensitivity Analysis		
		Complete Case†‡	Serious PONV*§	All Physicians and Patients*†
<b>Unadjusted analysis</b>				
Intervention group	<b>0.97 (0.87–1.10)</b>	0.93 (0.81–1.06)	0.99 (0.87–1.13)	0.98 (0.89–1.07)
Predicted risk#	<b>2.29 (2.07–2.54)</b>	2.83 (2.53–3.17)	2.27 (2.02–2.55)	2.24 (2.06–2.44)
Interaction: study group and predicted risk#	<b>0.92 (0.80–1.05)</b>	0.88 (0.74–1.04)	0.89 (0.75–1.04)	0.92 (0.82–1.03)
<b>Analysis adjusted for baseline differences**</b>				
Intervention group	<b>0.97 (0.86–1.09)</b>	0.93 (0.81–1.06)	0.98 (0.86–1.13)	0.94 (0.86–1.04)
Predicted risk#	<b>2.37 (2.10–2.66)</b>	2.92 (2.57–3.31)	2.41 (2.12–2.74)	2.36 (2.15–2.59)
Interaction: study group and predicted risk#	<b>0.91 (0.80–1.05)</b>	0.87 (0.74–1.03)	0.90 (0.76–1.06)	0.91 (0.81–1.02)

Numbers represent odds ratios with 95% CIs for the fixed effects. All analyses included time as a covariate. Bold numbers signify main results.

\* After multiple imputations using 10 imputation datasets. † Patients with any nausea, vomiting, or having received rescue antiemetics were considered cases. ‡ Cases with missing variables were discarded, resulting in an analysis with  $n = 8,104$  patients. § Only patients who reported definite nausea or vomiting were considered cases. || Unclustered logistic regression analysis in all physicians ( $n = 73$ ) and patients ( $n = 12,010$ ), without exclusion because of too small clusters. # Odds ratios represent predicted risks of 100% (a probability of 1). \*\* Baseline differences were adjusted by including surgical specialty and ambulatory surgery as covariates.

PONV = postoperative nausea and vomiting.



**Fig. 2.** Graphical representation of the mixed-effects regression analysis on the incidence of postoperative nausea and vomiting (PONV) (A) and on administration of prophylactic antiemetics by anesthesiologists (B). The *solid lines* and their 95% CIs represent the fixed effects of the mixed-effects regression analyses. The *dotted vertical line* shows the intersection point of both groups. The mixed-effects models included fixed effects for the variables such as allocation group, predicted risk of PONV, interaction between allocation group and predicted risk, and time. Random effects were included for the intercept, predicted risk, interaction between allocation group and predicted risk, and time. The *lines* were calculated using the average for the variable time. The 95% CIs were calculated from the covariance matrix for the variables allocation group and its interaction term with predicted risk. (A) May be interpreted as the occurrence of PONV after receiving prophylaxis, in patients with a particular predicted risk within each group. The differences between the *blue and red areas* represent the effect of implementation of the prediction model on the occurrence of PONV in patients with a particular predicted risk. (B) May be interpreted as the number of prophylactic antiemetics a patient with a particular predicted risk of PONV would receive from any anesthesiologist within each group. The differences between the *blue and red areas* therefore represent the changes in physician behavior concerning prescription of antiemetic prophylaxis, caused by implementation of the prediction model.

contamination is unlikely, as the predicted risk is not easily calculated without a proper decision support tool.

Fourth, we used an “assistive” implementation strategy as opposed to a “directive” approach: anesthesiologists were free to decide how to interpret the predicted risk and how

to assess the need for antiemetics. Additional interventions to increase the impact on physician behavior may include a more intensive education and feedback programme.<sup>15,41</sup> A directive approach, which includes actionable recommendations, may have a larger impact on physician behavior.<sup>42</sup>

**Table 6.** Mixed Effects Regression Analysis on the Number of Prophylactic Antiemetics

	Main Analysis*	Sensitivity Analysis		
		Complete Case†	Serious PONV‡	All Physicians and Patients*§
<b>Unadjusted analysis</b>				
Intervention group	<b>2.03 (1.64–2.50)</b>	2.03 (1.64–2.50)	2.03 (1.64–2.5)	1.76 (1.58–1.96)
Predicted risk	<b>2.16 (1.84–2.53)</b>	2.16 (1.84–2.53)	2.16 (1.84–2.53)	2.25 (2.07–2.43)
Interaction: study group and predicted risk	<b>1.63 (1.30–2.03)</b>	1.63 (1.30–2.03)	1.63 (1.30–2.03)	1.60 (1.43–1.80)
<b>Analysis adjusted for baseline differences#</b>				
Intervention group	<b>1.87 (1.55–2.25)</b>	1.87 (1.55–2.25)	1.87 (1.55–2.25)	1.88 (1.69–2.11)
Predicted risk	<b>1.81 (1.57–2.09)</b>	1.81 (1.57–2.09)	1.81 (1.57–2.09)	1.71 (1.56–1.87)
Interaction: study group and predicted risk	<b>1.57 (1.30–1.89)</b>	1.57 (1.30–1.89)	1.57 (1.3–1.89)	1.63 (1.45–1.83)

Numbers represent rate ratios with 95% CIs for the fixed effects. All analyses included time as a covariate. Bold numbers signify main results.

\* After multiple imputations using 10 imputation datasets. † Identical to the main analysis as there were no missings for the secondary outcome. ‡ Identical to the main analysis as the secondary outcome definition was not changed. § Unclustered Poisson regression analysis in all physicians (n = 73) and patients (n = 12,010), without exclusion because of too small clusters. || Rate ratios represent predicted risks of 100% (a probability of 1). # Baseline differences were adjusted by including surgical specialty and ambulatory surgery as covariates.

PONV = postoperative nausea and vomiting.

Fifth, the observed PONV incidence within 24 h after surgery in the care-as-usual group of our study is 43%, which is high for a study population which was not selected based on a preoperative risk of PONV, such as Apfel's trial.<sup>22</sup> The sensitivity analysis in which we defined PONV as only serious nausea or vomiting did not change the results (table 5). The average predicted risks of PONV from both the implemented prediction model and Apfel's<sup>8</sup> simplified risk score are very close to the observed PONV incidence in our study (table 3). Consequently, our study should be considered a high-risk population for PONV, and results may not necessarily translate to population with a lower average PONV risk.

Sixth, the results for the PONV incidence bordered on statistical significance. The exclusion of almost a quarter of initially randomized anesthesiologists and approximately 400 patients after randomization may have prevented the results to become significant. Therefore, an unclustered sensitivity analysis in all randomized patients was performed. The sensitivity analysis produced similar results as the main analysis (table 5). Even after adding the 400 patients who were excluded, the effect of the implemented prediction model on PONV remained very small in the sensitivity analysis, and it would require a much larger set of patients to become statistically significant. The small effect size is reflected in its clinical relevance, as a crude 2% reduction in PONV is not a substantial effect from a patient perspective.

In conclusion, implementation of a previously "validated" prediction model for PONV did not result in a clinically relevant decrease in PONV incidence despite an increase in "risk-tailored" application of prophylactic antiemetic strategies by anesthesiologists. Even when the use of a prediction model is consolidated in several guidelines, the discrepancy in the results of this study underscores the need to perform a formal impact analysis which includes patient outcome, before attempting to implement a prediction model into clinical practice.

## Acknowledgments

This study was funded by The Netherlands Organisation for Health Research and Development (ZonMW), The Hague, South Holland, The Netherlands. ZonMW Project Numbers: 945.16.202; 912.08.004; 918.10.615. Trial registration: Clinicaltrials.gov; Identifier NCT00293618: <http://clinicaltrials.gov/ct2/show/NCT00293618>.

## Competing Interests

The authors declare no competing interests.

## Correspondence

Address correspondence to Dr. Kappen: University Medical Center Utrecht, Department of Anesthesiology, P.O. Box 85500, Mail Stop F.06.149, 3508 GA Utrecht, The Netherlands. [t.kappen@umcutrecht.nl](mailto:t.kappen@umcutrecht.nl). Information on purchasing reprints may be found at [www.anesthesiology.org](http://www.anesthesiology.org) or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

## References

- Graham I, Atar D, Borch-Johnsen K, Boysen G, Burell G, Cifkova R, Dallongeville J, De Backer G, Ebrahim S, Gjelsvik B, Herrmann-Lingen C, Hoes A, Humphries S, Knäpion M, Perk J, Priori SG, Pyörälä K, Reiner Z, Ruilope L, Sans-Menendez S, Scholte op Reimer W, Weissberg P, Wood D, Yarnell J, Zamorano JL, Walma E, Fitzgerald T, Cooney MT, Dudina A; European Society of Cardiology (ESC) Committee for Practice Guidelines (CPG): European guidelines on cardiovascular disease prevention in clinical practice: Executive summary: Fourth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (Constituted by representatives of nine societies and by invited experts). *Eur Heart J* 2007; 28:2375–414
- Qaseem A, Snow V, Barry P, Hornbake ER, Rodnick JE, Tobolic T, Ireland B, Segal JB, Bass EB, Weiss KB, Green L, Owens DK; Joint American Academy of Family Physicians/ American College of Physicians Panel on Deep Venous Thrombosis/Pulmonary Embolism: Current diagnosis of venous thromboembolism in primary care: A clinical practice

- guideline from the American Academy of Family Physicians and the American College of Physicians. *Ann Intern Med* 2007; 146:454–8
3. Reilly BM, Evans AT: Translating clinical research into clinical practice: Impact of using prediction rules to make decisions. *Ann Intern Med* 2006; 144:201–9
  4. Hlatky MA, Greenland P, Arnett DK, Ballantyne CM, Criqui MH, Elkind MS, Go AS, Harrell FE Jr, Hong Y, Howard BV, Howard VJ, Hsue PY, Kramer CM, McConnell JP, Normand SL, O'Donnell CJ, Smith SC Jr, Wilson PW; American Heart Association Expert Panel on Subclinical Atherosclerotic Diseases and Emerging Risk Factors and the Stroke Council: Criteria for evaluation of novel markers of cardiovascular risk: A scientific statement from the American Heart Association. *Circulation* 2009; 119: 2408–16
  5. Moons KG: Criteria for scientific evaluation of novel markers: A perspective. *Clin Chem* 2010; 56:537–41
  6. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M: Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012; 98:691–8
  7. Gan TJ, Meyer TA, Apfel CC, Chung F, Davis PJ, Habib AS, Hooper VD, Kovac AL, Kranke P, Myles P, Philip BK, Samsa G, Sessler DI, Temo J, Tramer MR, Vander Kolk C, Watcha M; Society for Ambulatory Anesthesia: Society for Ambulatory Anesthesia guidelines for the management of postoperative nausea and vomiting. *Anesth Analg* 2007; 105:1615–28
  8. Apfel CC, Läärä E, Koivuranta M, Greim CA, Roewer N: A simplified risk score for predicting postoperative nausea and vomiting: Conclusions from cross-validations between two centers. *ANESTHESIOLOGY* 1999; 91:693–700
  9. Apfel CC, Kranke P, Eberhart LH, Roos A, Roewer N: Comparison of predictive models for postoperative nausea and vomiting. *Br J Anaesth* 2002; 88:234–40
  10. Pierre S, Benais H, Pouymayou J: Apfel's simplified score may favourably predict the risk of postoperative nausea and vomiting. *Can J Anaesth* 2002; 49:237–42
  11. van den Bosch JE, Kalkman CJ, Vergouwe Y, Van Klei WA, Bonsel GJ, Grobbee DE, Moons KG: Assessing the applicability of scoring systems for predicting postoperative nausea and vomiting. *Anaesthesia* 2005; 60:323–31
  12. Eberhart LH, Morin AM: Risk scores for predicting postoperative nausea and vomiting are clinically useful tools and should be used in every patient: Con-'Life is really simple, but we insist on making it complicated'. *Eur J Anaesthesiol* 2011; 28:155–9
  13. Kranke P, Eberhart LH: Possibilities and limitations in the pharmacological management of postoperative nausea and vomiting. *Eur J Anaesthesiol* 2011; 28:758–65
  14. Pierre S: Risk scores for predicting postoperative nausea and vomiting are clinically useful tools and should be used in every patient: Pro-'Don't throw the baby out with the bath-water'. *Eur J Anaesthesiol* 2011; 28:160–3
  15. Frenzel JC, Kee SS, Ensor JE, Riedel BJ, Ruiz JR: Ongoing provision of individual clinician performance data improves practice behavior. *Anesth Analg* 2010; 111:515–9
  16. Kooij FO, Klok T, Hollmann MW, Kal JE: Automated reminders increase adherence to guidelines for administration of prophylaxis for postoperative nausea and vomiting. *Eur J Anaesthesiol* 2010; 27:187–91
  17. Kooij FO, Vos N, Siebenga P, Klok T, Hollmann MW, Kal JE: Automated reminders decrease postoperative nausea and vomiting incidence in a general surgical population. *Br J Anaesth* 2012; 108:961–5
  18. Moons KG, Altman DG, Vergouwe Y, Royston P: Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. *BMJ* 2009; 338:b606
  19. Campbell MK, Grimshaw JM: Cluster randomised trials: Time for improvement. The implications of adopting a cluster design are still largely being ignored. *BMJ* 1998; 317: 1171–2
  20. Van den Bosch JE, Moons KG, Bonsel GJ, Kalkman CJ: Does measurement of preoperative anxiety have added value for predicting postoperative nausea and vomiting? *Anesth Analg* 2005; 100:1525–32
  21. Kappen TH, Vergouwe Y, van Klei WA, van Wolfswinkel L, Kalkman CJ, Moons KG: Adaptation of clinical prediction models for application in local settings. *Med Decis Making* 2012; 32:E1–10
  22. Apfel CC, Korttila K, Abdalla M, Kerger H, Turan A, Vedder I, Zernak C, Danner K, Jokela R, Pocock SJ, Trenkler S, Kredel M, Biedler A, Sessler DI, Roewer N; IMPACT Investigators: A factorial trial of six interventions for the prevention of postoperative nausea and vomiting. *N Engl J Med* 2004; 350:2441–51
  23. Visser K, Hassink EA, Bonsel GJ, Moen J, Kalkman CJ: Randomized controlled trial of total intravenous anesthesia with propofol *versus* inhalation anesthesia with isoflurane-nitrous oxide: Postoperative nausea with vomiting and economic analysis. *ANESTHESIOLOGY* 2001; 95:616–26
  24. Gan TJ: Postoperative nausea and vomiting—Can it be eliminated? *JAMA* 2002; 287:1233–6
  25. Carlisle JB, Stevenson CA: Drugs for preventing postoperative nausea and vomiting. *Cochrane Database Syst Rev* 2006; 3:CD004125
  26. Donner A: Sample size requirements for stratified cluster randomization designs. *Stat Med* 1992; 11:743–50
  27. Campbell MK, Thomson S, Ramsay CR, MacLennan GS, Grimshaw JM: Sample size calculator for cluster randomized trials. *Comput Biol Med* 2004; 34:113–25
  28. Campbell MK, Grimshaw JM, Elbourne DR: Intraclass correlation coefficients in cluster randomized trials: Empirical insights into how should they be reported. *BMC Med Res Methodol* 2004; 4:9
  29. Harrell FE Jr, Lee KL, Mark DB: Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15:361–87
  30. Little RJA: Regression with missing X's: A review. *J Am Stat Assoc* 1992; 87:1227–37
  31. Greenland S, Finkle WD: A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am J Epidemiol* 1995; 142:1255–64
  32. Donders AR, van der Heijden GJ, Stijnen T, Moons KG: Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006; 59:1087–91
  33. Moons KG, Donders RA, Stijnen T, Harrell FE Jr: Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006; 59:1092–101
  34. Leslie JB, Gan TJ: Meta-analysis of the safety of 5-HT<sub>3</sub> antagonists with dexamethasone or droperidol for prevention of PONV. *Ann Pharmacother* 2006; 40:856–72
  35. Neufeld SM, Newburn-Cook CV: The efficacy of 5-HT<sub>3</sub> receptor antagonists for the prevention of postoperative nausea and vomiting after craniotomy: A meta-analysis. *J Neurosurg Anesthesiol* 2007; 19:10–7
  36. Karanickolas PJ, Smith SE, Kanbur B, Davies E, Guyatt GH: The impact of prophylactic dexamethasone on nausea and vomiting after laparoscopic cholecystectomy: A systematic review and meta-analysis. *Ann Surg* 2008; 248:751–62
  37. Bauer MS, Williford WO, Dawson EE, Akiskal HS, Altschuler L, Fye C, Gelenberg A, Glick H, Kinosian B, Sajatovic M: Principles of effectiveness trials and their implementation in VA Cooperative Study #430: 'Reducing the efficacy-effectiveness gap in bipolar disorder'. *J Affect Disord* 2001; 67:61–78

38. Godwin M, Ruhland L, Casson I, MacDonald S, Delva D, Birtwhistle R, Lam M, Seguin R: Pragmatic controlled clinical trials in primary care: The struggle between external and internal validity. *BMC Med Res Methodol* 2003; 3:28
39. Hallfors D, Cho H, Sanchez V, Khatapoush S, Kim HM, Bauer D: Efficacy *vs* effectiveness trial results of an indicated "model" substance abuse program: Implications for public health. *Am J Public Health* 2006; 96:2254–9
40. Karanickolas PJ, Montori VM, Devereaux PJ, Schünemann H, Guyatt GH: A new "mechanistic-practical" framework for designing and interpreting randomized trials. *J Clin Epidemiol* 2009; 62:479–84
41. Weingarten SR, Henning JM, Badamgarav E, Knight K, Hasselblad V, Gano A Jr, Ofman JJ: Interventions used in disease management programmes for patients with chronic illness-which ones work? Meta-analysis of published reports. *BMJ* 2002; 325:925
42. Kawamoto K, Houlihan CA, Balas EA, Lobach DF: Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ* 2005; 330:765