



ELSEVIER

Controlled Clinical Trials 24 (2003) 736–751

Controlled
Clinical
Trials

Design paper

An International Multicenter Protocol to Assess the single and combined benefits of antiemetic interventions in a Controlled clinical Trial of a $2 \times 2 \times 2 \times 2 \times 2 \times 2$ factorial design (IMPACT)

Christian C. Apfel, M.D.^{a,b,*}, Kari Korttila, F.R.C.A., Ph.D.^c,
Mona Abdalla, Ph.D.^d, Andreas Biedler, M.D.^e, Peter Kranke, M.D.^a,
Stuart J. Pocock, Ph.D.^d, N. Roewer, M.D.^a

^a*Department of Anesthesiology, University of Wuerzburg, Wuerzburg, Germany*

^b*Outcomes Research Institute and Department of Anesthesiology, University of Louisville, Kentucky, USA*

^c*Department of Anesthesiology and Intensive Care, University of Helsinki, Helsinki, Finland*

^d*Statistics Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom*

^e*Department of Anesthesiology and Intensive Care, University Hospital of Saarland, Homburg, Germany*

Manuscript received October 10, 2002; manuscript accepted June 23, 2003

Abstract

For various diseases clinicians have to combine different drugs or interventions when a single drug or intervention does not lead to satisfactory results. However, quantifying the relative benefit of certain drugs or interventions when given alone and in combination under controlled conditions requires a complex factorial design. This paper describes such a method applied to a large multicenter trial for the prevention of postoperative nausea and vomiting (PONV), which may be of great interest for other specialties. Approximately 28 million operations are performed annually in the United States, mainly under general anesthesia with volatile anesthetics. Unfortunately, one-third of these patients suffer from PONV. This prompted extensive research of antiemetic and anesthetic drugs, but none of the interventions appeared to satisfactorily prevent PONV. Scuderi et al. were the first to almost eliminate PONV by combining various antiemetic interventions; however, the relative benefit of each intervention remained unclear. Accordingly, we have designed a large randomized controlled trial studying six different antiemetic interventions—three involving use of various antiemetic drugs and three involving choice of anesthetic drugs—to answer the following main questions: (1) What is the relative benefit of each of the antiemetic intervention? (2) Are certain combinations of antiemetic

* Corresponding author: Dr. Christian C. Apfel, Department of Anesthesiology and Outcomes Research Institute, University of Louisville, 501 E Broadway, Suite 210, Louisville, KY 40202. Tel.: 502-852-4055; mobile: 502-298-8932; fax: 502-852-2610.

E-mail address: apfel@ponv.org and christian.apfel@louisville.edu

interventions more effective than others? Using a complete factorial design this leads to $2^3 = 8$ antiemetic combinations, which multiply with the $2^3 = 8$ combinations of anesthetic drugs, leading to a total of $2^6 = 64$ possible combinations. The six factors are the antiemetics ondansetron (versus control), dexamethasone (versus control), droperidol (versus control), and the intravenous anesthetic propofol (versus volatile anesthetics), air (versus nitrous oxide), and remifentanyl (versus fentanyl). The primary outcome is freedom from PONV within the first 24 hours after anesthesia. Eligible patients are adults scheduled for elective surgery under general anesthesia with an increased risk for PONV so that the expected incidence in the control group (with none of the six antiemetic interventions) is approximately 60%. In order to allow analyses for up to three factor interactions, a sample size was estimated to be in the range of approximately 5000 patients. To the best of our knowledge this is the first randomized controlled trial of a six-way factorial design that may serve as an example for numerous other medical specialties. © 2003 Elsevier Inc. All rights reserved.

Keywords: Randomized controlled trial; Factorial design; Interaction analysis; Logistic regression analysis; Postoperative nausea and vomiting; Antiemetics; Ondansetron; Dexamethasone; Droperidol; Propofol

Introduction

There are numerous drugs or interventions that—when given alone—do not satisfactorily prevent or treat diseases. When various interventions with different sites of action are available, their combined use may result in increased efficacy. However, even if only a few approaches are available, the exponentially increasing number of possible combinations renders the identification of the “ideal” approach an almost impossible task. Accordingly, most “combination” studies are simple comparisons of a single versus a double intervention with usually no more than two or three arms. Unfortunately, the results of these studies can hardly be compared since they are derived from different populations and background factors. The only solution to this problem is the usage of a complete factorial design that allows quantifying the relative benefit of certain interventions when they are given alone and in combination under controlled conditions (see [Appendix A](#)). This paper describes this unique method applied to a large multicenter trial for the prevention of nausea and vomiting after anesthesia, which may be of interest for other specialties.

Postoperative nausea and vomiting (PONV) are well-known problems after inhalational anesthesia [1]. Institutional incidences vary considerably but on average every third patient suffers from PONV [2,3]. The individual risk of PONV can be predicted by operation-independent risk scores developed in the last decade [3–5], and recent validations suggest that simplified risk scores allow a correct classification in approximately 70% of cases [6,7]. Different antiemetic interventions have been shown to reduce PONV [8–10] but, because of their limited efficacy, unselective routine application has been questioned for medical and economical reasons [11,12]. However, as the number of patients needed to be treated to prevent one patient from experiencing PONV decreases when the control event rate (i.e., standard risk without the use of any antiemetic intervention) is high, patients at high risk for PONV may well profit from an effective prophylactic antiemetic intervention [13,14].

Strategies to reduce PONV can be categorized into (1) the prophylactic application of antiemetics drugs and (2) lowering the standard risk by the usage of less emetogenic anesthetics. The most promising antiemetics as quantified by meta-analyses are ondansetron, dexamethasone, and droperidol [8,10,15]. Correspondingly, the three most promising anesthetics

to lower the baseline risk involve the use of propofol instead of volatile anesthetics, the omission of nitrous oxide, and the use of remifentanyl instead of fentanyl [9,16–18]. The limited efficacy of a single intervention has prompted studies that combine two antiemetics with better but still not completely satisfactory results [19,20]. Accordingly, Scuderi et al. combined numerous approaches at the same time to reduce PONV and were able to demonstrate an almost complete elimination of PONV [21]. However, such multimodal approaches do not allow the quantification of the relative benefit of each antiemetic intervention so that some interventions may well be ineffective while a limited number of interventions may have led to a very similar result.

Study design

The above outlined problem can only be solved by a study stratified for different types of general anesthesia as well as for different antiemetics. Accordingly, this paper describes the International Multicenter Protocol to Assess the single and combined benefits of antiemetic interventions in a Controlled clinical Trial of $2 \times 2 \times 2 \times 2 \times 2 \times 2$ factorial design (IMPACT), which enables subanalyses of up to three factor interactions. Every patient will be randomized to receive one combination of the six influencing factors, i.e., allocated to 1 of the 64 groups.

- I. 4 mg ondansetron versus control
- II. 4 mg dexamethasone versus control
- III. 1.25 mg droperidol versus control
- IV. Propofol (Diprivan, Disoprivan) versus a volatile anesthetic
- V. Air versus nitrous oxide
- VI. Remifentanyl versus fentanyl

Outcome measures

The primary outcome is the incidence of nausea and/or vomiting in the first 24 hours after anesthesia. In the medical literature this is usually abbreviated as “PONV.”

The secondary outcomes are the incidences of postoperative vomiting and PONV in the early (0–2 hours), middle (2–6 hours), and delayed (6–24 hours) postoperative period as well as patient satisfaction.

Objectives

The primary objectives are to demonstrate that:

1. Each single intervention leads to an overall reduction of PONV. These “main effects” are to be quantified.
2. There are no clinically relevant two or three factor interactions, i.e., the factors act synergistically (additive synergism) without antagonism or potentiation (supra-additive synergism).
3. A decreased incidence of PONV will be associated with increased patient satisfaction.

The secondary objectives are to demonstrate that:

1. Volatile anesthetics cause dose-dependent early PONV with no clinically relevant differences between various specific volatile anesthetics.
2. The relative reduction rate of all antiemetic interventions (except dexamethasone) are highest in the early (0–2 hours) postoperative period.
3. The relative reduction rate of antiemetic interventions for PONV does not depend on the type of surgery, nor on patient-related risk factors such as female gender, smoking status, history of motion sickness or PONV, or young age.
4. The cost-benefit relation differs between various antiemetic interventions, even when wake-up times, nursing, bed cleaning, etc. are considered.
5. Additional consideration of antiemetic interventions such as antiemetics or the total intravenous anesthesia can significantly improve the prediction of a currently established simplified risk score for PONV [5].
6. The type of surgery is not an independent predictor for PONV and does not increase predictability significantly if included in a risk model.

Patient selection

Inclusion criteria included: male and female adults of at least 18 years of age, scheduled as in- or outpatients for elective surgery under general anesthesia of at least 1 hour with a risk of PONV of about 40% or higher according to Apfel et al. (two or more risk factors) [5].

Exclusion criteria included: all patients who have contraindications for any of the drugs used in the study, e.g., Parkinson's disease (droperidol), severe liver or kidney insufficiency (propofol, droperidol, ondansetron), severe cardiac arrhythmia (droperidol, ondansetron); pregnant or lactating females; patients under chemotherapy with emetogenic drugs; patients given antiemetic drugs within 24 hours before the operation; and patients who are anticipated to require postoperative artificial ventilation.

Stratification and randomization

Due to the factorial design the number of treatment combinations to be performed will be $2^6 = 64$. The centers are also offered an additional potentially influencing factor to consider in the randomization list, i.e., to titrate the depth of anesthesia as measured by an automated electroencephalographic algorithm called Bispectral index (BIS index, Aspect Medical Inc., Newton, MA, USA) to the range of 30–40 or 55–65, which leads to $2^7 = 128$ treatment combinations. For reasons described in the statistics section, the ratio of patients receiving propofol or volatile anesthetics as treatment will be 2:1. To achieve this ratio, twice as many patients who receive the 2^6 other treatment combinations should receive propofol in comparison to patients receiving volatile anesthesia. This results in a minimum number of $2^6 \cdot (2+1)^1 = 192$ randomized patients to receive the 128 treatment combinations in the required ratio. The predefined 192 combinations are randomized using the random function in Excel and doubled to obtain a list for 384 patients. Each center will be provided with a different randomization list.

Flanking study

Besides using nitrous oxide versus air with 30% oxygen, a third alternative is the use of air with 80% oxygen since recent results suggest the latter to be equally effective to the antiemetic ondansetron [22,23]. Three centers are interested in having this third alternative implemented in the randomization, which results in 96 patients with nitrous oxide, 96 patients with air (30% oxygen), and an identical branch of 96 patients with air (80% oxygen). This leads to 284 instead of 192 combinations for which a separate randomization will be created.

Anesthesia

On the morning of surgery, all patients will receive premedication with a benzodiazepine that is normally used in the department. Three minutes before induction either a bolus 100–200 µg fentanyl or a perfusor of 0.25 µg/kg/min remifentanyl will be applied (factor VI). The induction of anesthesia will be performed with 2–3 mg/kg propofol (Diprivan, Disoprivan; factor IV) IV. The intubation will be facilitated with rocuronium according to clinical needs. Normocapnic mechanical ventilation will be performed with nitrous oxide:oxygen or air:oxygen after intubation (factor V). Maintenance of hypnosis will be obtained with either propofol (Diprivan, Disoprivan) starting at 5 mg/kg/h or volatile anesthetics (factor IV) starting at 1 minimal alveolar concentration (MAC). If a center is also randomized for BIS-levels, hypnosis with propofol or the volatile anesthetic will be titrated to achieve BIS = 55–65 or 30–40 (factor VII); if not, the dosage will be left to the discretion of the anesthesiologist. Analgesia will be maintained according to stratification with boluses of 50–100 µg fentanyl IV if heart rate or blood pressure increases >20% of preoperative value. If randomized for remifentanyl, the infusion rate will continue with 0.25 µg/kg/min. This can be adjusted according to clinical needs (heart rate or blood pressure) within 0.1 to 0.5 µg/kg/min. If a center was randomized for BIS and a patient has values below 65 (indicating sufficient depths of hypnosis) while developing signs of insufficient anesthesia (e.g., movement on surgical stimulation), this should be treated with opioids as it most likely indicates insufficient analgesia. All patients should receive a loading dose of a nonopioid analgesic, preferably 2.5 g metamizol IV, 1 g paracetamol (=acetaminophen) rectally, or alternatively 2 g propacetamol IV depending on availability in that country. If the patient receives remifentanyl, up to 0.1 mg/kg piritramide, oxycodone, or morphine should be administered IV about half an hour before the end of surgery according to the expected requirement. According to the randomization 4 mg dexamethasone (factor II) or 1.25 mg droperidol (factor III) will be given intravenously within 20 minutes at the start of the case [24,25], while 4 mg ondansetron (factor I) will be given intravenously within 20 minutes before the end of surgery [26].

Blinded forms holding information about the investigated variables will be kept with each patient during the whole period. Additional postoperative analgesia will be provided with boluses of 0.05 mg/kg piritramide, oxycodone, or morphine intravenously and a lockout time set to 5 to 10 minutes as clinically required.

Postoperative assessments

Patients will be monitored for any emetic episodes for the first 24 postoperative hours and any emetic episode will be recorded with the time, severity, and its characteristics on

standardized forms. Emetic episodes will be treated first with ondansetron, second with dexamethasone, and third with droperidol as rescue medication. Treatment of further emetic episodes should be decided by the investigator free of choice. Maximum pain and nausea levels perception will be assessed at 120 minutes after anesthesia, along with the Myles recovery score [27]. Information on the following 22 hours will be obtained either by personal interview or written questionnaire including questions on pain, nausea, and emetic episodes as well as a repetition of the Myles recovery score and an anesthesia satisfaction questionnaire.

Statistical methods

The number and proportion of patients suffering from PONV during the 24 hours after anesthesia will be summarized and compared between any two treatment groups (e.g., ondansetron versus control) after controlling for the other treatment effects using the Mantel Haenszel test [28]. The Mantel Haenszel risk ratio and 95% confidence interval between these two treatment groups will be computed controlling for the other treatments, and a test of homogeneity of the risk ratios across the other treatments performed (this corresponds to a two-factor interaction test between this treatment and the other randomized treatments). This analysis is the simple method used to answer the first objective, allowing us to test the first-order interaction between these treatments. If no first-order interaction is present, due to the expected balanced randomization between the various treatments, we would expect this analysis to yield similar results to a chi-square test and corresponding 95% confidence interval for the risk ratio between the specific investigated treatments. In that case, the simple chi-square test and corresponding risk ratio would be summarized and presented to the reader due to its simplicity.

To fully investigate the second primary objective, a multivariate analysis in the form of a logistic regression analysis has to be defined [29]. A number of logistic regression models have been defined in Table 1 to estimate the following parameters:

Table 1. Models to investigate main objective 2

Model (number)	Notation of main effects and assumed interaction
Model 1: Log (odds PONV _{<i>i</i>})	$= \alpha + \beta_1 \cdot T1_i + \beta_2 \cdot T2_i + \beta_3 \cdot T3_i + \beta_4 \cdot T4_i + \beta_5 \cdot T5_i + \beta_6 \cdot T6_i$
Model 2: Log (odds PONV _{<i>i</i>})	$= \alpha + \beta_1 \cdot T1_i + \beta_2 \cdot T2_i + \beta_3 \cdot T3_i + \beta_4 \cdot T4_i + \beta_5 \cdot T5_i + \beta_6 \cdot T6_i + \lambda_{12} \cdot T1_i \cdot T2_i + \lambda_{13} \cdot T1_i \cdot T3_i + \lambda_{14} \cdot T1_i \cdot T4_i + \lambda_{15} \cdot T1_i \cdot T5_i + \lambda_{16} \cdot T1_i \cdot T6_i + \lambda_{23} \cdot T2_i \cdot T3_i + \lambda_{24} \cdot T2_i \cdot T4_i + \lambda_{25} \cdot T2_i \cdot T5_i + \lambda_{26} \cdot T2_i \cdot T6_i + \lambda_{34} \cdot T3_i \cdot T4_i + \lambda_{35} \cdot T3_i \cdot T5_i + \lambda_{36} \cdot T3_i \cdot T6_i + \lambda_{45} \cdot T4_i \cdot T5_i + \lambda_{46} \cdot T4_i \cdot T6_i + \lambda_{56} \cdot T5_i \cdot T6_i$
Model 3: Log (odds PONV _{<i>i</i>})	$= \alpha + \beta_1 \cdot T1_i + \beta_2 \cdot T2_i + \dots + \beta_6 \cdot T6_i + \lambda_{12} \cdot T1_i \cdot T2_i + \lambda_{13} \cdot T1_i \cdot T3_i + \dots + \lambda_{56} \cdot T5_i \cdot T6_i + \mu_{123} \cdot T1_i \cdot T2_i \cdot T3_i + \mu_{124} \cdot T1_i \cdot T2_i \cdot T4_i + \dots + \mu_{456} \cdot T4_i \cdot T5_i \cdot T6_i$
Model 4: Log (odds PONV _{<i>i</i>})	$= \alpha + \beta_1 \cdot T1_i + \beta_2 \cdot T2_i + \dots + \beta_6 \cdot T6_i + \lambda \cdot (T1_i \cdot T2_i + T1_i \cdot T3_i + T2_i \cdot T3_i) + \lambda_2 \cdot (T1_i \cdot T2_i + T1_i \cdot T3_i + T2_i \cdot T3_i) \cdot T4_i + \mu \cdot (T1_i \cdot T2_i \cdot T3_i) + \mu_2 \cdot (T1_i \cdot T2_i \cdot T3_i) \cdot T4_i$

Where: T1_{*i*} is the treatment defined as 1 if patient *i* received 4 mg ondansetron and 0 if patient received the control; T2_{*i*} is the treatment defined as 1 if patient *i* received 4 mg dexamethasone and 0 if patient received the control; T3_{*i*} is the treatment defined as 1 if patient *i* received 1.25 mg droperidol and 0 if patient received the control; T4_{*i*} is the treatment defined as 1 if patient *i* received propofol and 0 if patient received volatile anesthetic; T5_{*i*} is the treatment defined as 1 if patient *i* received air and 0 if patient received nitrous oxide; T6_{*i*} is the treatment defined as 1 if patient *i* received remifentanyl and 0 if patient received fentanyl.

In Model 1, the six main treatment effects adjusting for the other treatment effects are to be estimated using logistic regression analysis. They are denoted by β_1 to β_6 in the table. In Model 2, in addition to quantifying the six main effects, all possible two-factor interactions will be considered as denoted by $\lambda_{12} \dots \lambda_{56}$. Nonsignificant interactions will be dropped in the resulting equation (backward selection).

In Model 3, three factor interactions will be introduced and denoted by $\mu_{123} \dots \mu_{456}$. However, variables that failed statistical significance in main effects and in the two-factor interactions will not be considered. In addition, nonsignificant three-factor interactions will be dropped in the resulting equation (backward selection).

Using Model 4, we investigate whether the effectiveness of two or three antiemetics as a whole depend on the maintenance of anesthesia with propofol.

The second primary objective concerning patient satisfaction will be tested by comparing the distribution of the rank of patient satisfaction with the anesthesia between patients who suffered from a PONV episode with those who did not using the Wilcoxon sum of rank test. In order to perform multiple logistic regression on patient satisfaction as a binary outcome, the ranked satisfaction data will be dichotomized by performing a “median split.”

Sample size estimations

According to the patient selection, an average incidence of PONV of approximately 60% is assumed in the reference group (with no antiemetic intervention, i.e., 1 of 64 combinations). A decrease to 2/3 of the reference incidence is regarded as realistic. A type I error of 0.05 is defined as statistically significant. The type II error is defined to be 0.2, i.e., there will be a 80% power for detecting such a difference if it exists. The following assumptions were made.

Incidence of PONV for patients undergoing inhalational anesthesia is expected to be 60%. The expected reduction to 2/3 due to the application of propofol instead would lead to 40%. The addition of an antiemetic should lead to an expected PONV incidence of 26.7%. The addition of a second antiemetic should further reduce PONV to 17.8%, whereas the addition of the third antiemetic should result in 11.9% incidence. These calculations were made on the assumption of no two- or three-factor interactions between the three antiemetics and/or the type of anesthesia. Based on these assumptions, investigating whether a three-drug antiemetic combination is superior to a specific two-drug combination after propofol anesthesia, 598 patients per group are necessary to detect the difference of 5.9% between the incidence of 17.8% of the double combination and 11.9% of the triple combination. We hypothesize that the three double combinations of any two antiemetic treatments (e.g., ondansetron + dexamethasone, ondansetron + droperidol, and dexamethasone + droperidol) are equally effective. If the data are consistent with this assumption then the three double combinations would be grouped and compared with the one triple combination. The necessary number would be 409 patients per group, i.e., about 1227 for the double combination versus 409 patients for the triple combination. As 2^3 groups are existing, a total of $409 \cdot 2^3 = 3272$ patients are necessary to investigate the effect of the three antiemetics in patients receiving propofol for maintenance.

patients receiving the first or second antiemetics, respectively. Also let δ_{AB} be as $\delta_{..}$ but for patients receiving the (first+second) antiemetics, respectively. Under these assumptions, the two-way interaction term Δ can be written as follows: $\text{Log}(\Delta) = [(\delta_{AB} - \delta_{A.}) - (\delta_{.B} - \delta_{..})]$.

Given the assumption of a positive two-way interaction as mentioned above, then $\Delta = 34.3\%$. Now, the required sample size n per group needs to be estimated, so that the null hypothesis $\Delta = 0$ against a two-sided alternative at the 5% significance level and with an 80% power can be tested. The standard error of $\log(\hat{\Delta})$ is easily written from basic knowledge of the variance of the log odds ratio statistics from any 2×2 table [30]. The required sample size per group for an 80% power is thus equal to 316 patients (i.e., a total sample size of $316 \cdot 2^4 = 5056$). As a consequence of the factorial design, the total number of patients must be obtained by a multiple of 2·192 patients (the number of patients every center is going to enroll), so that 14 randomization lists (and centers) are needed, which results in 5376 patients. It should be noted that if we assume a negative two-way interaction term, then the sample size needed to detect this interaction would be larger than the one postulated above.

Ethics committee, informed consent, and patient insurance

Ethics committee approval was initially obtained by the University of Wuerzburg and, if necessary, by the local ethics committees of the participating centers. Informed oral and written consent will be obtained from all patients prior to their randomization. Patients may be contacted later in the course of the study for further investigations (second satisfaction assessment after some months, or cross-checking of evaluated data). Patient insurance, necessary in some participating countries due to national drug administration rules, is arranged by the principal investigator.

Conduct of the study

In general

The data will be primarily recorded at each center on specially designed forms and entered into local databases by the investigators. The data recorded on paper will remain at the centers while the databases will be mailed to the principal investigator in monthly intervals.

Fifteen European centers are participating and received a randomization list of 384 patients each. To accelerate patient acquisition up to 15 additional centers may be accepted by the principal investigator with a randomization list of 192 patients each. Randomization will be stopped when 5264 patients are enrolled.

Within the center

Eligible patients who give oral and written informed consent will be assigned to the next consecutive number (“code”) on the center’s randomization list, which will be noted on the anesthesia form and on the “preoperative assessment form.”

In the operating room or in the room for induction, the envelope for that code number will be opened to inform the anesthesiologist of the anesthesia and prophylactic antiemetic interventions to use. Accordingly, the anesthesiologist performing the anesthesia will not be blinded toward the drugs and is asked to fill out the “intraoperative assessment form.” However, information regarding the investigated variables will not be reported during the transfer from the operating room to the recovery room so that nurses and investigators of the postoperative period will remain blinded. Until the end of the study period, anesthesia and study forms containing information regarding the investigated variables will be kept in a sealed envelope near the patient to be directly accessible if unblinding becomes necessary.

The postoperative assessment will be performed at 2 and 24 hours postoperatively. Any emetic event, pain event, or adverse event should be documented with their characteristics and severity as well as the recovery questionnaire from Myles et al. [27] and the patient satisfaction questionnaire. Rescue treatment is recommended in case of emetic episodes, nausea levels ≥ 4 , or on the patient’s request. Within the whole 24-hour study period, the first episode of PONV should be treated with 4 mg ondansetron IV. If no satisfactory improvement occurs within 10 minutes the patient should receive 4 mg dexamethasone. If that fails also, 0.625 mg droperidol may be given. Further therapy is the choice of the investigator.

Adverse and serious adverse events

Adverse events will be recorded with the other outcome variables on the postoperative form, irrespective of whether they occurred intra- or postoperatively. All applicable details (time and kind of adverse event, severity, treatment, outcome, etc.) should be recorded. In case of a serious adverse event the information must be forwarded to the local ethics committee and to the principal investigator in Wuerzburg within 24 hours on working days or by 16:00 local time the next working day.

Monitoring

The study complies with the Good Clinical Practice/International Conference on Harmonisation guidelines. All data will be primarily recorded on specially designed forms and entered into local databases by the investigators at each center. The local databases have integrated plausibility checks ($n = 123$). Some fields are compulsory fields that must be entered and fulfill certain limits in order to enable saving the record (e.g., date of birth). Others will just give a warning if entries are unlikely but still possible (e.g., weight limits). The data recorded on paper will remain at the centers while the databases will be mailed to the data management center in monthly intervals. The data will be screened for completeness and plausibility, and discrepancy reports will be mailed to the investigators within 4 weeks to be corrected or explained. These include 151 automatically generated plausibility checks of cross-linked data (e.g., body mass index, dosage per kg body weight). Quality assurance will be achieved by visiting centers and cross-checking patient records with the study records (audit). In cases where discrepancies cannot be solved by comparison of records, patients may be contacted.

Discussion

To our knowledge, the current protocol is the largest randomized controlled trial on PONV. However, it is not size but factorial design that warrants attention in attempting to quantify the relative benefit of combining different interventions.

Assessing interactions

The concern of potential interactions has led researchers to control for all potentially influencing factors. The study of such homogeneous populations increases internal validity but at the same time decreases the external validity for a heterogeneous population in “real life.” In many cases it may therefore be more appropriate to investigate a mixed population from a wide spectrum to increase external validity. In such studies, the random treatment allocation should balance for all other potentially influencing factors. Usually the sample size of such studies is tailored to detect the main effect and not to investigate potential interaction. The difference in a factorial trial is that other potentially influencing factors are controlled to be equally distributed and not left to random chance (see [Appendix A](#)). If no interaction analyses are performed, the external validity is similar to a simple randomized controlled trial that balances other factors by random chance. This results in a highly efficient and economic approach to investigate multiple hypotheses “for the price of one,” since every patient contributes information to each of the randomized factors simultaneously. In addition, only factorial designs allow systematic interaction analyses and have high internal and external validities at the same time.

This trial has a six-way factorial design to quantify the benefits of each factor. The main effects could be quantified with a few hundred patients because groups can be lumped together while comparability will be preserved (see [Appendix A](#)). However, this investigation goes far beyond that by seeking to investigate interactions. As this leads to comparisons between smaller groups of patients, a much larger total number of patients is required. Thus, a total number of 5376 patients was estimated to have an adequate power to detect differences between three and two antiemetic treatment combinations given a specific anesthetic policy. In addition, this sample size is also sufficient to investigate three-factor (second-order) interactions. Ideally, an investigation of six-factor (fifth-order) interactions would provide an answer to every possible question; however, its required sample size would be far beyond any practical possibilities. In addition, according to previous investigations there are only a few situations where interactions may have a significant impact and in real life higher-degree interactions are extremely unlikely in the absence of lower-degree interactions [17].

Interactions in PONV

Nitrous oxide has been used and is still used to supplement anesthesia because of its analgesic properties (intraoperative opioid-sparing effect). During the last few years its continued use has been repeatedly questioned for causing PONV. One systematic review claimed that nitrous oxide increases the incidence of PONV only in patients at a very high risk for PONV while it has no significant effect when the risk of PONV is low [31]. Another systematic review published in the same year concluded that nitrous oxide increases the incidence of

PONV by about 1.4 [18]. Subgroup analyses found similar effects irrespective of type of surgery or whether propofol or volatile anesthetics were used, although the discussion is still ongoing [32]. However, effect “appeared” to be largest in female patients [18].

A very recent analysis of a single-center study with 1180 patients was unable to detect first-order (two-factor) interactions between antiemetics given prophylactically and the type of surgery or the type of anesthesia [17]. This was unexpected because pathophysiological hypotheses supported the idea that certain antiemetics were more effective after certain types of surgery or in patients with certain risk characteristics (e.g., females versus males). In addition, despite the plethora of published trials, there are almost no data suggestive of first-order interactions. Despite there being virtually no data available to support the existence of first-order interactions [12], they would be of great clinical interest.

The above-mentioned difficulties in interpreting results of small single-center studies as well as from meta-analyses demonstrate that a factorial approach is desperately needed to detect potential interactions of antiemetic interventions. However, its potential clinical implication should be discussed when the results from this promising project are available.

Timing and dose response of antiemetics

Splinter et al. demonstrated that low-dose ondansetron with dexamethasone more effectively reduced PONV than high-dose ondansetron alone [19]. Such phenomena are well known in other medical specialties and used to reduce unwarranted side effects. Thus, the ideal dosage of each antiemetic given in a triple combination may be considerably lower than if given alone. However, before an ideal dosage of a combination can be investigated, the most promising combination has to be identified in advance. For this, each antiemetic drug needs to be given in a sufficiently high dosage so that a further increase would not increase the antiemetic effect (ceiling effect) and that an increased antiemetic effect by the addition of another antiemetic is not just due to underdosage of the first antiemetic.

Current status of the trial

This manuscript was submitted to *Controlled Clinical Trials* after some centers had already started to recruit patients and we are pleased to report that enrollment was just completed while revising the print proofs of this manuscript. Thus, we expect our first results very shortly.

Conclusions

Combinations of interventions are frequently used in various fields of medicine. The only way to systematically quantify the relative benefit of such combinations is a randomized controlled trial of a complete factorial design. To our knowledge this is the first complete six-way factorial design with adequate power to investigate up to three-factor interactions. We are convinced that this highly efficient approach may serve as an example for other specialties.

Acknowledgments

This study was supported by resources from Department of Anesthesiology, University of Wuerzburg, Grant 1518 TG 72, Josef-Schneider Str. 2, D-97080 Germany; Helsinki University Central Hospital, HUS-EVO Grant TYH 0324, University of Helsinki, Finland; AstraZeneca, Tinsdaler Weg 183, D-22876 Wedel, Germany; and GlaxoSmithkline, Alsterufer 1, D-20354 Hamburg, Germany.

Appendix A. Principle of the factorial design

Factorial designs are well known by statisticians. However, during the phase of center recruitment we realized that very few anesthesiologists (and perhaps other clinicians as well) are familiar with this approach. Thus, a short explanation will be given here.

Basically, in a factorial design several treatments (factors) are randomized to the same patient [33,34]. This is an important approach when the benefit of a single treatment is limited. For the sake of simplicity we will start with two interventions (antiemetics): ondansetron versus placebo and dexamethasone versus placebo. When every patient is randomized twice, i.e., to both possible interventions, this is a 2×2 factorial design and leads to four possible groups: ondansetron+dexamethasone, ondansetron+placebo, placebo+dexamethasone, and placebo+placebo (Table 3). These four groups may be analyzed separately but this is not the intention. When all groups have the same size, the ondansetron+dexamethasone and ondansetron+placebo can be lumped together and compared with placebo+dexamethasone and placebo+placebo lumped together. This enables us to quantify the average effect of ondansetron (main effect). At the same time ondansetron+dexamethasone and placebo+dexamethasone can be lumped together and compared with ondansetron+placebo and placebo+placebo lumped together. This enables us to quantify the average effect of dexamethasone (main effect), which can now be compared with the main effect of ondansetron at the same time. Thus, if the main goal is to quantify the main effects of interventions, a factorial design offers to test multiple hypotheses “for the price of one” since it requires considerably fewer patients than if investigated in separate studies. Furthermore, as the interventions were given to the same patients in one trial, the comparison of the main treatment effects is more valid than from two different trials (or meta-analyses).

The main effects may be confounded if one treatment effect influences the other treatment effect (interaction). As the group in which an interaction may exist is only a fraction of the combined groups that are used to quantify the main effects, a considerably higher number of patients are required in order to have sufficient power ($1 - \beta$) to detect potentially existing interactions. At first sight this seems to contradict the above-mentioned decreased

Table 3. Factorial design with two factors: ondansetron and dexamethasone

	Ondansetron	Control	Sum
Dexamethasone	n	n	$2n$
Control	n	n	$2n$
Sum	$2n$	$2n$	$4n$

need for patients. However, it should be stressed that the factorial design is the only reliable approach to investigate interactions at all. Accordingly, it can be summarized that factorial designs have a very favorable cost-benefit ratio as they provide considerably more information (main effects plus the clinically important effect of the combination) by controlling the randomization of more than one factor compared to the conventional randomized controlled trial designs.

Appendix B. Contributors

Steering committee

Christian C. Apfel, M.D., Department of Anesthesiology, University of Wuerzburg, Germany
Kari Korttila, M.D. Ph.D. F.R.C.A., Department of Anesthesiology, University of Helsinki, Finland
Andreas Biedler, M.D., Department of Anesthesiology, University Hospital of the Saarland, Germany

Data management and monitoring

Christian C. Apfel, M.D., E. Kaufmann, M.D., M. Kredel, M.D., A. Schmelzer, M.D., J. Wermelt, M.D., Department of Anesthesiology, University of Wuerzburg, Germany
G. Link, Database-Engineering, Rimpär, Germany
Data-transfer to: PONV-research@mail.uni-wuerzburg.de

Statistical committee

Stuart J Pocock, Ph.D., Mona Abdalla, Ph.D., Statistics Unit, London School of Hygiene and Tropical Medicine, London, U.K.
Christian C. Apfel, M.D., Department of Anesthesiology, University of Wuerzburg, Germany

On site investigators (ordered by name of city)

Augsburg (Germany): H. Bartsch, M.D., H. Forst, M.D.
Barcelona (Spain): P. de la Mora, M.D., R. Rincon, M.D., P. Vila, M.D.
Berlin (Germany): G. Fritz, M.D., C. Hoehne, M.D.
Bielefeld (Germany): F. Bach, M.D., D. Buschmann, M.D., F. Mertzluft, M.D., I. Vedder, M.D.
Coburg (Germany): U. Jakob, M.D., S. Probst, M.D.
Duisburg (Germany): G. Frings, M.D.
Edirne (Turkey): A. Turan, M.D.
Eisenberg (Germany): M. Lange, M.D.
Freiburg (Germany): O. Danzeisen, M.D., C. Heringhaus, M.D., I. Schramm

Garmisch-Patenkirchen (Germany): W. Leidinger, M.D., J.N. Meierhofer, M.D., U. Rupert, M.D., C. Zernak, M.D.
Hamburg-Altona (Germany): H. Bause, M.D., H. Bordon, M.D., K. Stoecklein, M.D.
Heidelberg (Germany): H. Feierfeil, M.D., J. Motsch, M.D.
Helsinki (Finland): R.M. Jokela, M.D., A. Soikkeli, M.D., K. Korttila, M.D., F.R.C.A., Ph.D.
Homburg (Germany): A. Biedler, M.D., D. Detzel, M.D., W. Wilhelm, M.D.
Kaiserslautern (Germany): K. Danner, M.D., C. Madler, M.D., B. Steinbrecher, M.D.
Langen (Germany): H.-B. Hopf, M.D., S. Pohl, M.D.
Lauterbach (Germany): A. Goebel, M.D.
Lingen (Germany): B. Book, M.D., W. Hoeltermann, M.D., C. Prause, M.D.
Lueneburg (Germany): C. Frenkel, M.D., A. Paura, M.D.
Mannheim (Germany): C. Isselhorst, B. Fritz, M.D., H. Kerger, M.D.
Oulu (Finland): S. Alahuhta, M.D., T. Kangas-Saarela, M.D., P. Karjaleinen, M.D.
Plymouth (Great Britain): R. Sneyd, M.D.
Presov (Slovakia): E. Palencikova, M.D., S. Trenkler, M.D.
Rovaniemi (Finland): M. Koivuranta, M.D.
Schwerin (Germany): M. Hergert, M.D., F. Liebenow, M.D.
Tübingen (Germany): A. Kimmich, E. Schneider, M.D., M. Trick, M.D.
Ulm (Germany): L. Eberhart, M.D., K. Werthwein
Vienna (Austria): A. Bacher, M.D.
Wuerzburg (Germany): O. Eichelbrönnner, M.D., C.-A. Greim, M.D., E. Kaufmann, M.D., P. Kranke, M.D., M. Kredel, M.D., N. Roewer, M.D., A. Schmelzer, M.D., J. Wermelt, M.D.

References

- [1] Kovac AL. Prevention and treatment of postoperative nausea and vomiting. *Drugs* 2000;59:213–243.
- [2] Sinclair DR, Chung F, Mezei G. Can postoperative nausea and vomiting be predicted? *Anesthesiology* 1999;91:109–118.
- [3] Koivuranta M, Laara E, Snare L, Alahuhta S. A survey of postoperative nausea and vomiting. *Anaesthesia* 1997;52:443–449.
- [4] 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.
- [5] Apfel CC, Laara 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.
- [6] Pierre S, Benais H, Pouymayou J. Apfel's simplified score may favourably predict the risk of postoperative nausea and vomiting. *Can J Anesth* 2002;49:237–242.
- [7] Apfel CC, Kranke P, Eberhart LHJ, Roos IA, Roewer N. A comparison of predicting models for postoperative nausea and vomiting. *Br J Anaesth* 2002;88:234–240.
- [8] Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Efficacy, dose-response, and safety of ondansetron in prevention of postoperative nausea and vomiting: a quantitative systematic review of randomized placebo-controlled trials. *Anesthesiology* 1997;87:1277–1289.
- [9] Sneyd JR, Carr A, Byrom WD, Bilski AJ. A meta-analysis of nausea and vomiting following maintenance of anaesthesia with propofol or inhalational agents. *Eur J Anaesthesiol* 1998;15:433–445.
- [10] Henzi I, Walder B, Tramer MR. Dexamethasone for the prevention of postoperative nausea and vomiting: a quantitative systematic review. *Anesth Analg* 2000;90:186–194.

- [11] Scuderi PE, James RL, Harris L, Mims GR. Antiemetic prophylaxis does not improve outcomes after outpatient surgery when compared to symptomatic treatment. *Anesthesiology* 1999;90:360–371.
- [12] 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.
- [13] Cook R, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;310:452–454.
- [14] Watcha MF. The cost-effective management of postoperative nausea and vomiting. *Anesthesiology* 2000;92:931–933.
- [15] Henzi I, Sonderegger J, Tramer MR. Efficacy, dose-response and adverse effects of droperidol for prevention of postoperative nausea and vomiting. *Can J Anaesth* 2000;47:537–551.
- [16] Gan TJ, Glass PS, Howell ST, et al. Determination of plasma concentrations of propofol associated with 50% reduction in postoperative nausea. *Anesthesiology* 1997;87:779–784.
- [17] Apfel CC, Kranke P, Katz MH, et al. Volatile anaesthetics may be the main cause for early but not delayed postoperative nausea and vomiting: a randomized controlled trial of factorial design. *Br J Anaesth* 2002;88:659–668.
- [18] Divatia JV, Vaidya JS, Badwe RA, Hawaldar RW. Omission of nitrous oxide during anesthesia reduces the incidence of postoperative nausea and vomiting. A meta-analysis. *Anesthesiology* 1996;85:1055–1062.
- [19] Splinter WM, Rhine EJ. Low-dose ondansetron with dexamethasone more effectively decreases vomiting after strabismus surgery in children than does high-dose ondansetron. *Anesthesiology* 1998;88:72–75.
- [20] Heffernan AM, Rowbotham DJ. Postoperative nausea and vomiting—time for balanced antiemesis? *Br J Anaesth* 2000;85:675–677.
- [21] Scuderi PE, James RL, Harris L, Mims GR III. Multimodal antiemetic management prevents early postoperative vomiting after outpatient laparoscopy. *Anesth Analg* 2000;91:1408–1414.
- [22] Goll V, Akca O, Greif R, et al. Ondansetron is no more effective than supplemental intraoperative oxygen for prevention of postoperative nausea and vomiting. *Anesth Analg* 2001;92:112–117.
- [23] Greif R, Lacyni S, Rapf B, Hickie RS, Sessler DI. Supplemental oxygen reduces the incidence of postoperative nausea and vomiting. *Anesthesiology* 1999;91:1246–1252.
- [24] Liu K, Hsu CC, Chia YY. The effect of dose of dexamethasone for antiemesis after major gynecological surgery. *Anesth Analg* 1999;89:1316–1318.
- [25] Wang JJ, Ho ST, Tzeng JI, Tang CS. The effect of timing of dexamethasone administration on its efficacy as a prophylactic antiemetic for postoperative nausea and vomiting. *Anesth Analg* 2000;91:136–139.
- [26] Tang J, Wang B, White PF, et al. The effect of timing of ondansetron administration on its efficacy, cost-effectiveness, and cost-benefit as a prophylactic antiemetic in the ambulatory setting. *Anesth Analg* 1998;86:274–282.
- [27] Myles PS, Hunt JO, Nightingale CE, et al. Development and psychometric testing of a quality of recovery score after general anesthesia and surgery in adults. *Anesth Analg* 1999;88:83–90.
- [28] Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719–748.
- [29] McCullagh P, Nelder JA. Generalized linear models. London: Chapman & Hall; 1983.
- [30] Armitage P, Berry G. Statistical methods in medical research. New York: Blackwell Scientific Publications; 1994.
- [31] Tramer MR, Moore A, McQuay H. Omitting nitrous oxide in general anaesthesia: meta-analysis of intraoperative awareness and postoperative emesis in randomized controlled trials. *Br J Anaesth* 1996;76:186–193.
- [32] Verhecke G, Apfel CC. Early postoperative vomiting and volatile anaesthetics or nitrous oxide. *Br J Anaesth* 2003;90:109–110.
- [33] Lubsen J, Pocock SJ. Factorial trials in cardiology: pros and cons. *Eur Heart J* 1994;15:585–588.
- [34] Byar DP, Herzberg AM, Tan WY. Incomplete factorial designs for randomized clinical trials. *Stat Med* 1993;12:1629–1641.