Distribution of randomised controlled trials of drugs for post-operative nausea and vomiting

William P.S. McKay MD,
Raymond W. Yip MD

Purpose: Randomised controlled trials (RCTs) guide therapeutic decisions. But which RCTs are done; which omitted; and which should be done? This study illustrates a method to explore these questions applied to drugs for post-operative nausea and vomiting (PONV).

Methods: Review articles listed 18 drugs for PONV. All RCTs of these drugs for PONV were sought. The first drug mentioned in an RCT was counted and tabulated against others in all the arms of the RCT (against itself in a dose-ranging RCT). Additional drugs mentioned in these RCTs were added to the study, for a total of 40 drugs.

Results: Drugs involved in the most RCTs were: ondansetron 131 RCTs; propofol 118; droperidol 74; metoclopramide 67; granisetron 52; scopolamine 22; tropisetron 16. Drugs involved in the fewest RCTs: two drugs with 2 RCTs; twelve drugs with one; three with none. Probability that this distribution occurred by chance: $P < 0.00001$; that the distribution of dose-ranging RCTs occurred by chance: $P < 0.001$. Regression of RCT numbers on cost: $R = 0.86$, $P < 0.0001$; on year of drug introduction: $R = 0.14$. Of 1600 possible comparisons of drugs for PONV, (including dose-ranging) 97.8% have never been published.

Conclusion: Although some antiemetic drugs for PONV have been studied in large numbers of RCTs, many have not been adequately evaluated. Finding relevant RCTs and tabulating their comparison arms is useful for directing future research, and is applicable to any symptom or disorder.

Objectif : Les essais contrôlés randomisés (ECR) guident les choix thérapeutiques, mais quels ECR sont réalisés, omis ou devraient être faits ? L'étude actuelle illustre une méthode pour explorer ces questions appliquées aux médicaments contre les nausées et vomissements postopératoires (NVPO).

Méthode : L'examen des articles a donné une liste de 18 médicaments contre les NVPO. Tous les ECR de ces médicaments contre les NVPO ont été recherchés. Le premier médicament mentionné dans un ECR était pris en compte et classifié pour être comparé aux autres dans tous les groupes de l'ECR (comparé à lui-même dans un ECR portant sur un éventail de doses). Les médicaments additionnels mentionnés dans ces ECR ont été ajoutés à l'étude, amenant le total à 40 médicaments.

Résultats : Les médicaments étudiés dans le plus grand nombre d'ECR ont été : l'ondansétron, 131 ECR; le propofol, 118; le droperidol, 74; le métoclopramide, 67; le granisétron, 52; la scopolamine, 22; le tropisétron, 16. Les médicaments qu'on retrouve dans le plus petit nombre d'ECR : deux médicaments dans deux ECR; douze médicaments dans un; trois dans aucun. La probabilité que cette distribution soit l'effet du hasard : $P < 0,00001$; que la distribution d'essais sur des doses multiples soit aléatoire : $P < 0,001$. La régression du nombre des ECR selon le coût : $R = 0,86$, $P < 0,0001$; selon l'année d'introduction du médicament : $R = 0,14$. Sur 1 600 comparaisons possibles de médicaments contre les NVPO (y compris les comparaisons de doses), 97,8 % n'ont jamais été publiées.

Conclusion : Même si certains médicaments antiémétisants contre les NVPO ont été étudiés dans un grand nombre d'ECR, beaucoup n'ont pas été complètement évalués. La découverte d'ECR pertinents et le classement de leurs groupes comparatifs est utile en vue de recherches ultérieures et s'applique à tout symptôme ou tout dérèglement.

From the Department of Anaesthesia, University of Saskatchewan, Royal University Hospital, 103 Hospital Drive, Saskatoon, Saskatchewan, S7N 0W8 Canada.

Address correspondence to: Dr. W.P.S. McKay. Phone: 306-655-1183; Fax: 306-655-1279; E-mail: mckayw@sdh.sk.ca

Accepted for publication January 28, 2000.

POSTOPERATIVE nausea and vomiting (PONV) is a problem which has, in common with many symptoms and disorders in the practice of medicine, many therapies with no gold standard treatment. Three reviews (none systematic) of the subject¹⁻³ have been published in the major English-language anesthesiology journals. The last of these discusses the desirability of basing treatment on randomised controlled trials (RCTs), and the difficulty of interpreting RCTs done in the area. None of the reviews is a satisfactory source of information for evidence-based practice.

In an attempt to clarify the problem locally for our department, we reviewed RCTs pertaining to the 18 antiemetic drugs mentioned in the two most recent reviews. It soon became clear that more RCTs than we cared to read were available for a few drugs. It was equally evident that no recent RCTs existed to compare commonly used drugs. This maldistribution of information presents a dilemma in attempting to practice evidence-based medicine. A search of the pre-electronic literature confirmed that some antiemetics had apparently never been studied for PONV by published RCT, although RCTs as a scientific method have been published since 1948.⁴

The randomised controlled trial (RCT) is the source of evidence-based therapeutic decisions. The scientific quality and ethical rigor of RCTs,^{5,6} and the reliability of the information^{7,8} they provide have improved in recent years. There has been concern expressed recently by journal editors that RCTs with "negative" outcomes may not be submitted for publication.^{9,10} The present study addresses RCTs that have never been published and, perhaps, have never been conducted. We wondered if a cross-tabulation of RCTs done by comparison arm might help to clarify and quantify the problem.

Ideally, drugs used in clinical practice will have been compared in RCTs. Null-hypotheses were: 1) RCTs have been used to compare all drugs for PONV; 2) the number of RCTs published per drug: a) is equal or is distributed randomly; b) is independent of drug price; c) is independent of year of drug introduction.

Methods

The review articles cited above were used as the source of 18 putative antiemetic drugs. All RCTs ever published on these drugs, with PONV as a measured endpoint, were sought electronically in *Oldmedline* (1960 B 1965) and *Medline* (1966 B 1998);^{11,12} in *Index Medicus* (pre-1960);¹³ and in *The Cochrane Collaboration Controlled Trials Register*.⁸ Medline search-words were "<drug name>" AND ("nausea"

OR "vomiting" OR "emesis") AND ("surgery" OR "operat*") AND Randomised Controlled Trial, where "operat*" is a truncation device which retrieves citations beginning with "operat", such as operation or operative. "<drug name>" AND "nausea" AND "surgery" were used in the Cochrane Register. Drug names were sought in the Index Medicus part of the search. Where the RCT abstract did not give enough information, the original article was read. Studies of prevention and treatment of nausea and vomiting in other areas, such as cancer chemotherapy, were excluded.

Examination of the search records of the first ten drugs (alphabetically) revealed that the Cochrane Register retrieved eight RCTs not found in Medline. No RCTs before 1999 found in Medline were absent from the Cochrane Register, which is now the largest and most complete compilation of RCTs in the world. The third quarterly issue of the Cochrane Collaboration Library does not contain studies done after 1998. Thus, following this complete Medline and Oldmedline and Index Medicus search of the first ten drugs, the process was simplified for the remaining drugs. Medline searching was limited to 1999; the Cochrane Register was used exclusively for studies prior to 1999.

In any RCT, a therapy may be tested against another therapy, against placebo, against itself in various doses (a dose-finding RCT, commonly called dose-ranging in RCT parlance), against a combination of therapies (a drug-interaction RCT) or against any combination of these. Each such comparison is called an arm of the RCT. The index arm contains the sample of subjects receiving the index drug. Comparison arms comprise the subjects receiving each comparison therapy. Our method simply tabulates all arms of all relevant RCTs against the first drug mentioned in the RCT, which we designate the index drug. All additional antiemetics mentioned in articles found by this means were similarly studied. The first drug mentioned in any RCT was deemed the Aindex drug@ against which other therapies were tried.

The drugs were arrayed in alphabetical order down the side (index drugs) and along the top (comparison arms) of a computer spread sheet (Table I). Additional columns were inserted for tabulating placebo arms and for the total number of RCTs for which each drug was the index drug. Thus, horizontal rows represented index drug arms of an RCT, and vertical columns represented comparison arms. For arms containing fixed-dose combinations of drugs, an additional column was added. Combinations were listed separately.

A "double-entry book-keeping" system was used. Each RCT was tabulated by adding 1 to both the

TABLE I

index drug total RCTs	placebo arms arms	Comparison → Arms Index Arms ↓	drug a	drug b	drug ...	drug m
a as index	a vs placebo	Index drug a	a vs a	a vs b	...	a vs m
b as index	b vs placebo	Index drug b	b vs a	b vs b	...	b vs m
...	...	Index drug Y	...	...	...	...
n as index	n vs placebo	Index drug n	n vs a	n vs b	...	n vs m

TABLE II

DRUG	DATE	COST	RCTs	DRUG	DATE	COST	RCTs
Ondansetron	1987	\$10.50	138	ephedrine	1927	\$0.43	4
Propofol	1979	\$5.53	121	prochlorperidine	1963	\$0.30	4
Droperidol	1963	\$2.69	85	flunitrazepam	1973	na	3
Metoclopramide	1966	\$0.87	67	ramosetron	1999	na	3
Granisetron	1988	\$10.50	62	trimeprazine	1959	\$0.27	2
Dexamethasone	1962	\$0.40	23	betamethasone	1961	\$3.95	1
Scopolamine	1880	\$0.95	23	cimetidine	1975	\$0.05	1
Tropisetron	1984	na	16	cisapride	1984	\$0.55	1
Glycopyrrolate	1964	\$0.99	15	chlorpromazine	1952	\$0.27	1
Atropine	1867	\$0.16	14	cyclizine	1966	\$2.82	1
Dolasetron	1994	\$10.50	11	haloperidol	1958	\$1.79	1
Midazolam	1978	\$0.99	11	hydroxyzine	1956	\$0.95	1
Perphenazine	1960	\$2.46	11	phenoperidine	1968	na	1
Diazepam	1962	\$0.46	10	ranitidine	1979	\$0.12	1
Promethazine	1954	\$0.28	6	temazepam	1969	\$0.13	1
Dimenhydrinate	1950	\$0.30	5	tiapride	1976	na	1
Domperidone	1977	\$0.16	5	triflupromazine	1959	na	1
Ginger	1969	na	5	amitriptyline	1963	\$0.02	0
Lorazepam	1971	\$1.45	5	diphenhydramine	1946	\$0.86	0
Alizapride	1982	na	4	nortriptyline	1963	\$0.28	0
Dixyrazine	1967	na	4				

na = not commercially available

“Index Drug Total RCTs” for the index drug, and to the spreadsheet cell at each intersection of index drug with study arm(s). The diagonal cells from upper left to lower right corners give the counts of index-drug-against-itself, that is, of dose-ranging RCTs.

Many RCTs, especially those with propofol as the index drug, were conducted with comparison arms comprised of variants of general anesthesia. It is difficult or unsafe in such cases to Ablind@ the anesthesiologist, or to use placebo drugs. It was decided that, since both the subject (asleep) and the evaluator were blinded as to whether or not drug was given, an arm that simply gave no index drug would be treated as a placebo arm.

The year of introduction of drugs was found by seeking the earliest published citation of the drug in human studies in Medline, Oldmedline, the Cochrane Register, or Index Medicus, except for drugs available before 1950, for which standard pharmacology texts¹⁴ were consulted.

Drug costs (Table II) were obtained from the hospital pharmacy. They were the cost to the hospital per usual recommended adult dose in Canadian dollars. Price is for the parenteral form where available. Where a parenteral drug is supplied in a multidose vial with preservative, the cost of the vial is divided by the number of doses therein. If a partial ampoule must be discarded when the drug is used, the cost is that of a full ampoule.

Statistics

Distribution of numbers of RCTs by index study drug was tested for a difference from an even distribution of RCTs per drug by 1-tailed unpaired t test. Although this hypothesis test is an instance of a comparison of multiple proportions, chi-squared could not be used because the observed numbers of RCTs in many cells was zero. Fisher=s Exact test is suitable only for two by two contingency tables. Therefore, the observed distribution of RCTs was compared to a hypothetical

“expected” even distribution, that is, with each drug being involved in a number of RCT studies equal to the mean number of RCT studies-per-drug. This mean (15.1 RCTs per drug) was subtracted from the observed RCT total for each drug. Then a one-sample *t* test was applied to the absolute differences of the RCT counts from zero, that is to say, from a “zeroed flat line”.

Simple linear regression was used to compare RCTs per drug to cost and to year of introduction.

Results

A total of 268 RCTs were published. *Medline* searching in some cases yields only 20 to 80% of RCTs⁸ of a given therapy. However, the *Cochrane Collaboration Controlled Trials Register* uses a sophisticated search strategy including synonyms and words near in meaning. Thus, the simple search strategy probably retrieved almost all of the published RCTs.

Distribution of total RCTs per drug is given in Table II. Droperidol was the commonest drug used in a comparison arm. There were 369 placebo arms. Twenty nine of 40 drugs had no dose-ranging RCT. Probability that the overall distribution of RCTs occurred by chance: $P < 0.00001$; that the distribution of dose-ranging RCTs occurred by chance: $P < 0.001$. Regression of RCT numbers on cost: $R = 0.61$, $P < 0.001$; on year of drug introduction: $R = 0.14$. Of 1600 possible intersections of index drugs versus comparison arms with other drugs, 1565, or 97.8% had no RCTs.

Discussion

Distribution of RCTs

The uneven distributions of RCTs among therapies and for dose-ranging for PONV are both very unlikely to be due to chance ($P < 0.00001$ and $P < 0.001$ respectively). This raises several questions. Do some drugs have “too many” RCTs? Does it matter in a practical clinical sense that some have none? Is there a need for dose-ranging for such widely used old drugs? Is the process appropriate? Might we make more rational decisions about planning future RCT research? What about funding such research?

Drugs with “too many” RCTs

It is uncertain from this study whether there is a cause-and-effect relationship between drug price and RCT numbers. They are correlated. If there is cause-and-effect, it could lie in either direction. Pharmaceutical-company sponsored RCTs are necessary to meet regulatory demands, but are also used as a marketing tool to introduce new products by involving medical

school teachers as investigators. New drugs are thus introduced indirectly to house staff - the future physicians. The cost of RCTs done to meet regulatory requirements could drive up the price of new drugs. Alternatively, high-priced, potentially lucrative drugs could prompt a decision to sponsor more RCTs. Finally, it may be that academic anesthesiologists and their house-staff find that drugs that are expensive for reasons unrelated to the decision to do a RCT happen to be more interesting to study.

Current ethical standards for RCTs state that patients should not be recruited to take part in experiments that are unlikely to advance scientific knowledge to a degree commensurate with the risks of taking part. Further, they should not have the trouble of a decision about participating added to their worries about illness unless some probable good is served by putting them to the trouble. Although the risks of participating in PONV studies are small, in many cases it is doubtful that important scientific advancement will flow from further studies of much-tested therapies.

The process that determines the decision to launch an RCT is complex and beyond the scope of this paper. Suffice it to say that concern has been expressed about the distribution of limited resources for scientific medical research.¹⁵ Had the resources been distributed evenly among the drugs, there would have been on average 15.1 RCTs per drug. Far fewer than 15 well designed, well co-ordinated, and well-executed studies per drug would likely have produced clear winners and losers long before now.

Drugs with few or no RCTs

Meta-analysis is a powerful *retrospective* statistical technique to summarise objectively the available information from RCTs, which may have conflicting or inconclusive results.^{16,17} On the other hand, the technique of the present study is intended to foster rational *prospective* research. In the case of PONV, it draws to our attention the fact that we are using therapies with little evidence of efficacy from RCTs.

For example, dimenhydrinate is among the commonest antiemetics in Canadian hospital practice. It was no better than placebo in two of a total of five RCTs. It has been tested only against ondansetron (two RCTs), metoclopramide (one RCT), and placebo (four RCTs), with no dose-ranging study ever published. Ondansetron was superior to dimenhydrinate in one study, and no better in another.

Drugs with “enough” RCTs

Dexamethasone is a drug that has been studied by 23 RCTs in a variety of settings, all of which show it to be

safe. The RCTs show dexamethasone to be efficacious in all but two. In one of these, dexamethasone combined with ondansetron was no better than ondansetron alone. In the other study, neither dexamethasone nor granisetron performed better than placebo, but a combination of the two was very efficacious. Informal enquiries of colleagues in several Canadian cities suggest that dexamethasone is seldom used to prevent PONV. If the consistent results of its RCTs stand up to rigorous meta-analysis, its use should be actively promoted.

Dose finding

Dimenhydrinate has no dose-ranging RCTs. Are such studies useful for a drug as familiar as dimenhydrinate? An equally familiar drug, acetaminophen, has recently been found, by dose-ranging RCT,¹⁸ to be more efficacious at a dose very different from the one we learned as interns. Similarly, current doses of epidural analgesics for labour are very different from those used a few years ago because of dose-ranging RCTs.¹⁹ Dose-ranging RCTs should be done even for familiar drugs.

Placebo arms

Every drug that was involved in a study with a placebo arm performed better than placebo in at least one study. Thousands of subjects have received suboptimal care by being randomised to the 370 placebo arms of PONV RCTs. It is difficult ethically to justify placebo arms in future studies of PONV.

Drug-interaction RCTs

The interactions of drugs with similar principal effects, such as antinausea, are of several kinds. They may interact pharmacokinetically to speed or slow disposal. Pharmacodynamically, they may produce effects which are, compared with those of either drug given alone, either diminished, additive, synergistic (super-additive), or orthogonally unrelated to the principal effect.

Drugs for prevention and treatment of PONV fall into several categories, depending upon taxonomy: 1) prokinetics; 2) histamine-2 receptor blockers; central nervous system blockers of: 3) dopamine-2, 4) acetylcholine, 5) histamine-1, 6) serotonin-3 receptors, and 7) others of uncertain mechanism (such as dexamethasone, benzodiazepines, or ginger).

Drugs with additive or synergistic effects may be given together in smaller than usual doses, often with decreased side-effects. Of the few RCTs that study combinations of drugs, most demonstrate a benefit from combining drugs of different categories.^{20,21} More such studies should be done.

Habits and marketing

Several of the untested drugs (amitriptyline, nortriptyline, diphenhydramine, and haloperidol) are neither marketed nor conventionally thought of as antiemetics. They appear in the reviews of PONV because animal studies of their pharmacodynamic properties suggest that they should be effective in reducing PONV. They deserve study by RCT.

Implications for research funding

Many of the untested older drugs are inexpensive. Therefore, agencies which directly fund health care, and which have not traditionally funded research in this area, should consider investing in a co-ordinated program of RCTs combined with associated cost-effectiveness studies⁷ in this area. Such studies might quickly point to considerable cost savings.

A suggested systematic approach

It is unnecessary and impractical to plan RCTs to compare all 1565 index-arm intersections for which there are no studies. Calculation of all possible drug-interaction arms yields astronomical numbers of possible studies. However, awareness of critical gaps in our knowledge in this field can foster informed discussion of the rational allocation of future research resources. It can highlight which trials need be done to complete the essential knowledge base of medical evidence.

A suggested plan for a co-ordinated study of the field consists of three main stages. First, dose-ranging RCTs should be done for promising drugs that lack them. Second, a set of RCTs to determine the best drug for PONV in each of the seven categories listed above might be implemented. Third, thoughtful design of a set of RCTs to look at drug combinations of those best-of-category drugs, perhaps using smaller doses than would be used for single drugs used alone, could be planned. All of this presupposes rigorously conducted studies and authoritative central co-ordination.

The first stage, dose-ranging, allows optimal doses to be used in stage two. For some drugs, the second stage has already been done and need not be repeated. For example, there is a large body of non-PONV literature comparing cimetidine and ranitidine. They are equally efficacious, but ranitidine has far fewer undesirable side-effects.

Using the technique

This simple exploratory tool is a first step in reviewing therapies for any disorder. It should be thought of as a rapid screening technique that can produce a useful discussion piece for those with research interests. It can clarify what has been done and what has not. It

can help to decide which should be the next step for a therapy: meta-analysis or more RCTs? It can suggest useful RCTs to pursue. At that point, deeper and more extensive searching, as outlined in the Cochrane Collaboration Library,⁸ should be done before beginning to design a study.

Conclusions

Cross-tabulating all arms of all RCTs of all therapies for a symptom or disease is a simple but potentially enlightening preliminary step to clarify and quantify the question: which RCTs and meta-analyses ought to be done in the area of interest? It is applicable to any symptom or disorder.

Acknowledgments

The authors wish to thank Jackie Lutze and Lorene Penner of the Saskatoon District Health Pharmacy for their helpful contribution.

References

- 1 *Riding JE*. The prevention of postoperative vomiting. *Br J Anaesth* 1963; 35: 180–8.
- 2 *Palazzo MGA, Strunin L*. Anaesthesia and emesis. I: etiology. *Can Anaesth Soc J* 1984; 31: 178–87. II: Prevention and management. *ibid.*: 407–15.
- 3 *Watcha MF, White PF*. Postoperative nausea and vomiting. Its etiology, treatment, and prevention. *Anesthesiology* 1992; 77: 162–84.
- 4 Medical Research Council Investigation. Streptomycin treatment of pulmonary tuberculosis. *BMJ* 1948; 2: 769–82.
- 5 *Cochrane AL*. Effectiveness and efficiency. Random reflections on health services. London: The Nuffield Provincial Hospital Trust, 1972.
- 6 *Bulpitt CJ*. Randomised Controlled Clinical Trials, 2nd ed. Boston: Kluwer Academic Publishers, 1996.
- 7 *Petitti DB*. Meta-analysis, Decision Analysis, and Cost-Effectiveness Analysis. New York. Oxford University Press, 1994.
- 8 Cochrane Library 1999, Issue 3, CD-ROM, Update Software Limited, Oxford, England.
- 9 *Chalmers I, Dickerson K, Chalmers TC*. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992; 305: 786–8.
- 10 *Smith R, Roberts I*. An amnesty for unpublished trials (Editorial). *BMJ* 1997; 315: 622.
- 11 National Library of Medicine. MEDLINE; OLDMEDLINE. Washington DC: U.S. Dept of Health, Education, and Welfare. Public Health Service.
- 12 Internet Grateful Med. Internet Grateful Med Development Team. National Library of Medicine. Bethesda, MD, USA. <http://igm.nlm.nih.gov>
- 13 Quarterly Cumulative Index Medicus. Chicago: American Medical Association.
- 14 *Gilman AG, Goodman LS*. Goodman and Gilman's The Pharmacologic Basis of Therapeutics. New York: MacMillan Publishing Co.
- 15 Committee on the NIH Research Priority-Setting Process, Institute of Medicine (US). Scientific opportunities and public needs. Improving priority setting and public input at the National Institutes of Health. Washington: National Academy Press; 1998.
- 16 *Tramèr M, Moore A, McQuay H*. Propofol anaesthesia and postoperative nausea and vomiting: quantitative systematic review of randomised controlled studies. *Br J Anaesth* 1997; 78: 247–55.
- 17 *Tramèr M, Moore A, McQuay H*. Meta-analytic comparison of prophylactic antiemetic efficacy for postoperative nausea and vomiting: propofol anaesthesia *vs* omitting nitrous oxide *vs* total i.v. anaesthesia with propofol. *Br J Anaesth* 1997; 78: 256–9.
- 18 *Birmingham PK, Tobin MJ, Henthorn TK, et al*. Twenty-four-hour pharmacokinetics of rectal acetaminophen in children. An old drug with new recommendations. *Anesthesiology* 1997; 87: 244–52.
- 19 *Campbell DC*. Epidural analgesia in labor: evolution and revolution. *Canadian Journal of CME* 1998; 10: 105–14.
- 20 *Fujii Y, Tanaka H, Toyooka H*. Granisetron-dexamethasone combination reduces postoperative nausea and vomiting. *Can J Anaesth* 1995; 42: 387–90.
- 21 *Riley TJ, McKenzie R, Tantisira BR, Hamilton DL*. Droperidol-ondansetron combination *versus* droperidol alone for postoperative control of emesis after total abdominal hysterectomy. *J Clin Anesth* 1998; 10: 6–12.