

BMJ

Analysis and interpretation of cost data in randomised controlled trials: review of published studies

Julie A Barber and Simon G Thompson

BMJ 1998;317;1195-1200

Updated information and services can be found at:
<http://bmj.com/cgi/content/full/317/7167/1195>

These include:

References

This article cites 29 articles, 11 of which can be accessed free at:
<http://bmj.com/cgi/content/full/317/7167/1195#BIBL>

38 online articles that cite this article can be accessed at:
<http://bmj.com/cgi/content/full/317/7167/1195#otherarticles>

Rapid responses

You can respond to this article at:
<http://bmj.com/cgi/eletter-submit/317/7167/1195>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Topic collections

Articles on similar topics can be found in the following collections

[Clinical trials \(epidemiology\)](#) (6511 articles)
[Complementary medicine](#) (805 articles)
[Internet](#) (1420 articles)

Notes

To Request Permissions go to:
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to *BMJ* go to:
<http://resources.bmj.com/bmj/subscribers>

Analysis and interpretation of cost data in randomised controlled trials: review of published studies

Julie A Barber, Simon G Thompson

Abstract

Objective To review critically the statistical methods used for health economic evaluations in randomised controlled trials where an estimate of cost is available for each patient in the study.

Design Survey of published randomised trials including an economic evaluation with cost values suitable for statistical analysis; 45 such trials published in 1995 were identified from Medline.

Main outcome measures The use of statistical methods for cost data was assessed in terms of the descriptive statistics reported, use of statistical inference, and whether the reported conclusions were justified.

Results Although all 45 trials reviewed apparently had cost data for each patient, only 9 (20%) reported adequate measures of variability for these data and only 25 (56%) gave results of statistical tests or a measure of precision for the comparison of costs between the randomised groups. Only 16 (36%) of the articles gave conclusions which were justified on the basis of results presented in the paper. No paper reported sample size calculations for costs.

Conclusions The analysis and interpretation of cost data from published trials reveal a lack of statistical awareness. Strong and potentially misleading conclusions about the relative costs of alternative therapies have often been reported in the absence of supporting statistical evidence. Improvements in the analysis and reporting of health economic assessments are urgently required. Health economic guidelines need to be revised to incorporate more detailed statistical advice.

Introduction

With the continuing development of new treatments and medical technologies, health economic evaluations have become increasingly important. To identify cost effective care, providers, purchasers, and policy makers need reliable information about the costs as well as the clinical effectiveness of alternative treatments. For clinical outcomes, randomised controlled trials are the standard and accepted approach for evaluating interventions. This design provides the most scientifically rigorous methodology and avoids the biases which limit the usefulness of alternative non-randomised designs.¹ Pragmatic randomised controlled trials provide a suitable environment not only for assessing clinical effectiveness but also for comparing costs,²⁻⁴ and an increasingly large amount of economic data is being collected within trials.^{5,6}

The costs of competing treatments are usually estimated using information about the quantities of resources used—that is, the set of cost generating items which make up the treatment and its consequences. For example, the resources used in a surgical operation may include the staff time involved, the consumables

used, and the length of a subsequent inpatient stay. To estimate the cost of treatment, this resource use information is combined with unit cost estimates, which give a fixed monetary value to each cost generating item. The total cost of treatment is then the weighted sum of the quantities of resources used, where the weights are the unit costs.

The cost associated with a treatment may be estimated as a deterministic (fixed) value by costing a typical treatment protocol. This approach requires assumptions about the usual quantities of healthcare resources that would be used during treatment. For the surgical procedure example, this would involve assumptions about the grades of staff present during the operation, the typical time taken, consumables used, and length of inpatient stay. Carrying out an economic evaluation alongside a randomised controlled trial, however, allows detailed information to be collected about the quantities of resources used by each patient in the study: a record would be kept for every patient of the actual staff present, time taken, consumables used, and inpatient stay. Such information allows an estimate of the cost of treatment to be obtained for each individual patient, producing a set of cost values, which will be referred to as “patient specific” cost data.

Availability of patient specific cost data not only allows the use of statistical inference as a basis for drawing conclusions about costs but reduces the extent to which the comparison between randomised groups is based on assumptions about resource use. In addition it allows the relation between costs and other factors such as patient characteristics and clinical outcomes to be investigated.

In trials where patient specific cost data are available, the comparison of costs between treatment groups is used to make inferences about the true cost difference in the population from which the trial sample was drawn. The evidence from the sample needs to be assessed using statistical analysis. Although several reviews of economic evaluations have been undertaken,^{5,7-15} to date none has concentrated specifically on statistical aspects of the analysis of patient specific cost data from randomised controlled trials. We therefore focused on this issue, aiming to assess the use of statistical methods in this context and whether the conclusions drawn for costs are properly justified.

Methods

Selection of study articles

Published papers included in this review are those which reported on randomised trials where patient specific cost data were available, on which statistical methods were or could have been used. The search was limited to publications in English, involving human subjects, and published during 1995 and was carried out using the Medline database as of April 1997. The search required at least one of “trial” or “interven-

Department of Medical Statistics and Evaluation, Imperial College School of Medicine, London W12 0NN

Julie A Barber, research fellow in medical statistics

Simon G Thompson, professor of medical statistics and epidemiology

Correspondence to: Ms Barber
j.barber@rps.ac.uk

BMJ 1998;317:1195-1200

tion(s)" and at least one of "health economic(s)," "economic evaluation," or "cost(s)" in the title, abstract or MeSH headings. The search identified 872 eligible articles.

Papers were excluded on the basis of their abstracts if it was clear that they were not reporting on the results of a randomised trial. The full articles were read for 111 papers. Where patient specific cost data had not been collected or when information about costing methods were insufficient to judge their suitability, the papers were excluded. For unclear cases, the articles were reread by a second reviewer and agreement reached as to their suitability. In this way 45 articles were finally included in the review.

Information collected

A data collection form was developed and was completed on reading each article in the review. This included information about the collection and calculation of costs, sample size calculations cited, summary measures reported, and statistical methods used. The final part of the assessment judged the appropriateness of any inferential conclusions drawn about costs, given the statistical results presented in the paper. These judgments did not involve consideration of design issues or methods of analysis but were simply based on cost estimates and any P values or confidence intervals reported.

Initial assessments for all papers were carried out by one assessor (JAB). Most of the information collected involved recording what was and was not explicitly stated in the paper, so that little subjective judgment was required. To examine reproducibility for these items a second investigator (SGT), unaware of the initial assessments, independently assessed a random sample of nine of the 45 trials. Agreement was complete for items reported in this paper. In the case of the potentially more subjective judgments about the appropriateness of the conclusions drawn, all 45 articles were read and categorised independently by both reviewers. There was only one disagreement, this caused by misreading of the paper by one reviewer. In five other cases, discussion was needed to determine the classification, because the reporting of results and conclusions in these was unclear.

Results

Description of papers

The 45 papers identified came from both specialist and more general journals, and covered a wide variety of clinical areas including cancer, heart disease, nursing, and psychiatry. About half (24; 53%) were primary publications for the trial which usually included both clinical and economic results. In many of these, the economic component was rather small and lacking in detail. The remaining papers (21; 47%) were "follow on" papers to the main effectiveness analyses, which reported cost results either alone or in combination with other outcomes of interest, such as quality of life. The vast majority of the studies were designed as pragmatic trials, directly relevant to clinical practice; the economic analysis thus had direct policy implications.

The economic data in these trials either came from resource use information, using some assumed unit cost values, or from data on charges for health care.

The number of resource items included in the calculation of total costs varied considerably; some used quite detailed elements while others had very few. Patient specific information was sometimes only available for a limited number of resources, while fixed cost estimates were assumed for others.

Sample size calculations

Sample size calculations were mentioned in only seven (16%) of the 45 articles in the review. None were for economic outcomes; six were based on clinical endpoints, and in the remaining case it was unclear which outcomes were being considered. In the case of health economic assessments published separately from the main effectiveness analyses, sample size calculations for clinical outcomes may have been reported elsewhere.

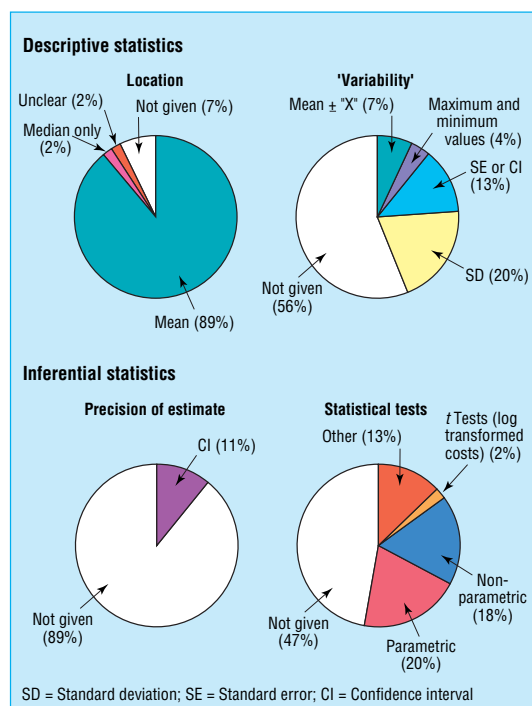
For 10 papers (22%), authors reported using a subsample of the original randomised trial for the economic analysis. Various reasons were given for this, including selection of a subset to minimise the burden on patients in the study; interest in the relative costs of only two arms of a three arm trial; and inclusion of only some centres from a multicentre trial, either because the others refused to be involved in the economic evaluation or in order to reduce data collection efforts.

Descriptive statistics

One trial in the review, which compared four three day antimicrobial regimens for treatment of acute cystitis, found mean costs (US\$) per patient of \$114 for patients treated with trimethoprim-sulpamethoxazole, \$131 for amoxicillin, \$155 for nitrofurantoin, and \$155 for cefadroxil.¹⁶ No information on the variability or ranges of costs per patient were given, so it is impossible to judge to what extent the average presented was typical for the patients studied. In a trial of whether to re-evaluate patients receiving oxygen at home at intervals of two months or six months, the mean cost and standard deviation over one year were presented for each group in the trial.¹⁷ For example, in the six month re-evaluation group the standard deviation was larger than the mean (\$11 580 and \$8870 respectively), indicating a very wide dispersion of costs between individuals. This information helps to put the mean costs observed into perspective.

Reporting of descriptive information is an important part of a statistical investigation and should precede analysis. For cost data, the crucial information is the arithmetic mean—that is, the simple average cost. This is because policy makers, purchasers, and providers need to know the total cost of implementing the treatment. This total cost is estimated as the arithmetic mean cost in the trial, multiplied by the number of patients to be treated. Measures other than the arithmetic mean (such as the median, mode, or geometric mean) cannot provide an estimate of total cost. The fact that the distribution of costs is often highly skewed does not imply that the use of the arithmetic mean is inappropriate. However, describing the variability in costs between individuals in the trial, and any peculiarities in the shape of the distribution such as skewness, is also important.

The figure shows the percentage of all the papers reviewed reporting various summary measures for the



Proportion of 45 papers reporting descriptive statistics and inferential statistics for costs. Statistical tests were *t* test or analysis of variance (parametric); Mann-Whitney U test or Kruskal-Wallis test (non-parametric); 2 regression, 4 unspecified ("other"). SD=standard deviation; SE=standard error; CI=confidence interval

cost data in each randomised group. Overall 42 papers (93%) reported measures of location, which were given as arithmetic mean or total costs in all but two articles. Six papers reported other measures of location along with the mean, five giving medians and one presenting modes in each group.

Of the 45 papers, 20 (44%) reported one or more measures which described the spread or range of the cost data across individuals in each randomised group. As shown in the figure, standard deviations were used to indicate variability between individuals in nine (20%) of the papers. The other 11 papers (24%) gave measures that do not directly or fully describe the variability in the cost data. Three gave standard errors and three gave confidence intervals for the means in each group. Two further papers reported the maximum and minimum cost values only, and the remaining three presented a mean plus or minus some quantity "X", where the authors failed to state explicitly the meaning of this quantity.

Some papers had indications that the authors were aware of the likely non-normal distribution of their cost data. For nine (20%) this was explicitly stated, and three of these represented the distribution graphically. Seven further papers (16%) indicated some awareness about distributional problems either by reporting median cost (rather than or in addition to the mean) or by using non-parametric tests or log transformations when analysing the cost data.

Inferential statistics

Inferences made about costs need to be supported by a measure of precision (standard error or confidence interval) of the difference in mean costs between

randomised groups, or at least a P value. For example, a study of induction of labour versus serial antenatal monitoring reported that the mean cost (Canadian \$) in the monitoring group was higher by \$193 (95% confidence interval \$133 to \$252, $P < 0.0001$).¹⁸ In contrast, a study of midwife team versus routine care during pregnancy and birth simply reported that the average cost (Australian \$) per delivery was "\$3324 for team care women and \$3475 for routine care women, resulting in a saving of \$151, or a 4.5% reduction in costs."¹⁹ In the latter example, no inference is justified since the precision of these cited quantities is unknown.

The inferences about the average cost difference need to be based on a comparison of arithmetic means as, for example, given by the *t* test. Analyses of log transformed costs address the differences in geometric means, while non-parametric tests address differences in both median and shape of the cost distribution between groups. These analyses do not consider the question of interest about the arithmetic mean cost difference. There may, however, be legitimate concern over the validity of the *t* test, analysis of variance, and other standard methods of comparing arithmetic means. These methods all require assumptions of normality which may be violated by the often highly skewed distribution of cost data, particularly when sample sizes are small.

Overall, only 25 of the 45 articles (56%) reported results of statistical tests or a measure of precision for the comparison of costs between the randomised groups. Only five (11%) gave a measure of precision for the estimated difference in costs (figure). All were reported as confidence intervals calculated using methods which assume normality. 24 of the studies (53%) reported a P value for a comparison of costs between the treatment groups (figure). In nine (20%) of these, P values were obtained from a two sample *t* test or from analysis of variance comparing arithmetic mean costs across more than two groups. In one paper, a *t* test was carried out on log transformed costs.

Non-parametric tests were used in eight papers (18%). Two papers reported results from regression analyses only; four papers reporting P values failed to state which test had been used, one of which reported the P value in the abstract of the report only. Three of the reviewed papers (7%) included more detailed analyses adjusting for predictors of costs by using multiple regression models of untransformed or log transformed costs.

Justification for conclusions

For the study of induction of labour versus serial antenatal monitoring mentioned at the beginning of the previous section, the authors concluded in the abstract that "a policy of managing post-term pregnancy through induction of labour...results in lower cost."¹⁸ This is an inferential conclusion that could be extrapolated from the trial results to future policy, and it is justified in terms of the confidence interval and P value for the mean cost difference presented. The trial of midwife team care versus routine care concluded that "the team approach...was associated with a reduction in costs per woman."¹⁹ This would also be likely to be interpreted as an inferential statement by readers. However, it was based simply on a comparison of mean

Classification of conclusions regarding costs. Values are number of papers with justified conclusions out of the total number in each category (percentages)

Conclusion	With statistical inference	Without statistical inference	Total
Claimed difference	13/16	0/20	13/36
Claimed no difference	1/7	0	1/7
Claimed insufficient evidence of a difference	2/2	0	2/2
Total	16/25 (64)	0/20	16/45 (36)

costs, without any information on the precision of the mean cost difference observed. It is not a justified conclusion.

The table summarises results of these assessments for all the papers in the review. All the 45 papers presented apparently inferential conclusions regarding costs. The justification of a conclusion was judged in a narrow sense in terms of whether supporting inferential statistics were cited; inadequacies of design or problems of data (such as missing values) or inappropriate analysis (such as non-parametric tests) were not considered. Hence a lenient view of the “justification for conclusions” was taken. Despite this, only 16 (36%) were judged to have been justified. This finding was identical for conclusions presented in the abstract or the main text. In a substantial number of cases (20) no statistical analysis was provided and in all cases the conclusions were not justified because they were apparently simply based on an eyeball comparison of the mean costs observed in each group. All of these papers claimed a difference in costs. Among the studies that undertook statistical analysis, the main reason that conclusions were not justified was that a claim of no difference in cost was made on the basis of a non-significant test result, without providing the necessary confidence interval for the cost difference.

Missing data, cost effectiveness, and sensitivity analyses

Information concerning the completeness of the cost data was given for only 24 studies (53%). Of these, three mentioned that their data were complete and 21 stated that some data were missing, the amount ranging up to 35% of the sample. Eleven papers apparently excluded subjects with missing cost data from the analysis without any further investigation. Five others compared characteristics of this group of patients with those whose data were complete, in order to identify any obvious biases. Four further papers dealt with missing data in other ways: one used a sensitivity analysis, another imputed values, and two used longitudinal analyses which do not require the data to be complete at all time points.

Seven (16%) of the trials reported some measure of cost effectiveness—for example, cost per quality adjusted life year, cost per year of life gained, or cost per unit change in some clinical measurement. None of these papers carried out statistical tests for the cost effectiveness estimates or used confidence intervals to report on their precision. Two, however, used the confidence intervals of the effects, and in one case costs, to consider extreme cases of the cost effectiveness ratio.

Only 11 (24%) of the 45 studies reported having carried out sensitivity analyses, and in five cases these were for the cost effectiveness results. The sensitivity analyses investigated robustness to various assumptions including unit costs, cost to charge ratios, assumed resource use values, and discount rates.

Discussion

Randomised controlled trials are not always the appropriate vehicle to address economic questions,^{20 21} and there is an important role for other methods of economic evaluation, such as modelling.²² When economic evaluations are carried out alongside randomised controlled trials, however, the cost data collected should be interpreted appropriately. This review has revealed major deficiencies in the way cost data in randomised controlled trials are summarised and analysed.

Descriptive statistics

In providing descriptive information for continuous data, such as costs, recommended practice²³ would be to present a measure of location (for example, mean or median) and variability (for example, standard deviation or interquartile range) and mention any peculiarities about the shape of the distribution (such as skewness). Cost data are typically highly skewed, because a few patients incur particularly high costs. The arithmetic mean is then larger than the median, sometimes substantially, because it is more influenced by these high costs. Although the median can be interpreted as the most “typical” cost for individual subjects, since half of them have costs below this value and half above, it is the arithmetic mean cost that is important for policy decisions. It is only the arithmetic mean—not other measures such as the median, mode or geometric mean—that, when multiplied by the number of patients to be treated, estimates the total cost that would be incurred if the treatment were implemented. Although these other measures are commonly used for skewed data in other circumstances, the more informative arithmetic mean should always be reported for costs. This was done in nearly all the papers in our review, but statistical comparisons often used methods that did not directly compare these arithmetic means.

Summarising the distribution of costs observed in a trial can be problematic unless there is space to show the distribution as a diagram. Because of skewness, the standard deviation alone is not an ideal way to represent the spread of costs between individuals. The observations lying within two standard deviations of the mean will cover about 95% of a distribution of values only if the distribution is approximately normal. Often, for cost data, the value two standard deviations below the mean is an impossible negative quantity. It is therefore also useful to present the interquartile range, a range containing the central 50% of the cost data, or a 95% reference interval, a range that excludes 2.5% of the cost data at each extreme. The full range (minimum to maximum) is less useful because it is totally dependent on just the two most extreme observations. Standard errors and confidence intervals reflect the precision of the estimated mean and are not appropriate ways of describing how the

costs vary between individuals.²³ Most papers in our review did not describe the variability of their cost data at all, and many of the others gave only unsatisfactory summary information.

Inferential statistics

The interpretation of patient specific cost data in randomised controlled trials needs to be guided by formal methods of statistical inference—but only half of the papers reviewed presented a P value or confidence interval for cost comparisons. Conclusions regarding the evidence about cost differences cannot reliably be made without such statistical analysis. Among the papers that used statistical analysis, half used inappropriate methods (such as the non-parametric Mann-Whitney U test, or analysis of log transformed costs) that do not compare of arithmetic mean costs. Only 11% of the papers presented a confidence interval for the average cost difference, although the use of confidence intervals has repeatedly been recommended in statistical guidelines.^{23 24}

The review focused on randomised controlled trials, since the rigour of this design might be expected to be accompanied by rigour in statistical analysis and reporting. However, overall, only 36% of conclusions drawn were justified. This is a lenient view since it takes no account of problems in design and execution of trials or the use of inappropriate methods of statistical analysis. Reporting inappropriate conclusions for either clinical or economic outcomes is potentially misleading and unethical.²⁵ Economic outcomes should be evaluated with the same statistical standards that are now expected for clinical outcomes. The tendency to make strong conclusions based simply on observed mean values of costs is all the more flawed when small samples have been used for the economic evaluation.

Sample size calculations

The often large variability in costs between individuals emphasises the need to perform economic evaluations on sufficiently large samples so that precise conclusions can be drawn. The rationale for sample size calculations (having adequate power for the planned analyses and having a predetermined stopping point) are as relevant to cost outcomes as to clinical outcomes. Although cost outcomes are often regarded as “secondary,” they are still important. There may be practical reasons to base the health economic evaluation on a subset of the whole trial but statistical justification is lacking. The use of subsets and the complete absence of sample size calculations reported for costs in this review indicates the large scope for improvement in the rational planning of economic evaluations.

Completeness and relevance of the review

The review was based on papers published in 1995 accessed through Medline. Limiting the search to journals on a single database means that this may not be an exhaustive review of all relevant papers. The reporting standards of journals cited by Medline, however, are likely to be better than those of non-Medline journals, therefore producing an overly optimistic view of the use of statistical methods in economic evaluations. The results of a similar search using the Cochrane Control-

led Trials Register included 43 of the 45 papers in this review (the other two were both follow on papers to a main clinical effectiveness publication and in both cases only the clinical paper appeared in the Cochrane register). The Medline search may not have identified absolutely all randomised controlled trials with patient specific costs.²⁶ Some trials were excluded from the review because it was not clear from their methods whether patient specific cost data had been collected; however, these trials presented no measures of variability or statistical inferences for costs.

Standards may have improved since 1995 in response to general guidelines,²⁷ although these currently contain little recommendation regarding statistical aspects of economic evaluations. A recent study evaluating the BMJ guidelines²⁷ failed to show that these had had any impact on the general quality of economic evaluations submitted or published.²⁸ In addition, experience with statistical guidelines indicates that the rate of response to these is generally slow,^{29 30} since precedent is a powerful inhibitor of change.

Statistical complexities

The statistical issues in analysing cost data are not, however, all straightforward,³¹ in particular how to compare arithmetic mean costs in very skewed data. Standard methods for analysing arithmetic means such as the *t* test are known to be fairly robust to non-normality. This robustness, however, depends on several features of the data, in particular sample size and severity of skewness. There are no set criteria by which to judge whether the analysis will be robust for a particular dataset, and relying on standard methods could produce misleading results, especially if sample sizes are small. Extending simple comparisons to adjust for baseline variables may exacerbate the problems. Both simple and more complex analyses of costs can, however, be carried out or checked using bootstrapping.³² This approach allows a comparison of arithmetic means without making any assumptions about the cost distribution. Although some examples of the use of bootstrapping for cost data have recently been published,³³ this method is not yet routinely used by medical researchers.

Other statistical issues in the analysis of costs include choosing an appropriate sample size for the evaluation, placing confidence intervals on cost effectiveness ratios,³⁴ handling missing data, and providing a rational strategy for sensitivity analyses. All of these are complicated issues that are in need of further clarification.

Conclusion

This review has shown that there is an urgent need to improve the statistical analysis and interpretation of cost data in randomised controlled trials. The BMJ guidelines and other health economics guidelines need to be revised to incorporate more detailed statistical advice for researchers, editors, and reviewers when dealing with patient specific cost data from trials. These guidelines not only need to encourage the use of statistical inference but need to provide advice on dealing with some of the more complex issues mentioned above.

Key messages

- Health economic evaluations required for important healthcare policy decisions are often carried out in randomised controlled trials
- A review of such published economic evaluations assessed whether statistical methods for cost outcomes have been appropriately used and interpreted
- Few publications presented adequate descriptive information for costs or performed appropriate statistical analyses
- In at least two thirds of the papers, the main conclusions regarding costs were not justified
- The analysis and reporting of health economic assessments within randomised controlled trials urgently need improving

Contributors: JB and ST were both involved in developing the ideas and methods for the paper, and in writing the text, and are guarantors of the paper. JB took the principal role in extracting the relevant information from the papers in the review and summarising the results.

Funding: JB was funded by North Thames NHS Executive; ST was funded by HEFC London University.

Competing interests: None declared.

- Bradford Hill A. Observation and experiment. *N Engl J Med* 1953;248:995-1001.
- Drummond MF, Stoddart GL. Economic analysis and clinical trials. *Controlled Clin Trials* 1984;5:115-28.
- Drummond MF, Davies L. Economic analysis alongside clinical trials. Revisiting the methodological issues. *Int J Technol Assess Health Care* 1991;7:561-73.
- Thompson SG, Barber JA. From efficacy to cost-effectiveness. *Lancet* 1998;350:1781.
- Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized control trials. *Med Care* 1992;30:231-43.
- Elixhauser A, Luce BR, Taylor WR, Reblando J. Health care CBA/CEA: an update on the growth and composition of the literature. *Med Care* 1993;31(suppl):S1-11.
- Jefferson T, Demicheli V. Is vaccination against hepatitis B efficient? A review of world literature. *Health Econ* 1994;3:25-37.
- Briggs A, Sculpher M. Sensitivity analysis in economic evaluation: a review of published studies. *Health Econ* 1995;4:355-71.
- Ancona-Berk VA, Chalmers TC. Cost and efficacy of the substitution of ambulatory for inpatient care. *N Engl J Med* 1981;304:393-7.
- Evers SMAA, van Wijk AS, Ament AJHA. Economic evaluation of mental health care interventions. A review. *Health Econ* 1997;6:161-77.
- Mason J, Drummond M. Reporting guidelines for economic studies. *Health Econ* 1995;4:85-94.

- Udvarhelyi IS, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost-benefit analyses in the medical literature. Are the methods being used correctly? *Ann Intern Med* 1992;116:238-44.
- Ganiats TG, Wong AF. Evaluation of cost-effectiveness research: a survey of recent publications. *Fam Med* 1991;23:457-62.
- Gerard K. Cost-utility in practice: a policy maker's guide to the state of the art. *Health Policy* 1992;21:249-79.
- Zhou X, Melfi CA, Hui SL. Methods for comparison of cost data. *Ann Intern Med* 1997;127:752-6.
- Hooton TM, Winter C, Tiu F, Stamm WE. Randomized comparative trial and cost analysis of 3-day antimicrobial regimens for treatment of acute cystitis in women. *JAMA* 1995;273:41-5.
- Cottrell JJ, Openbrier D, Lave JR, Paul C, Garland JL. Home oxygen therapy. A comparison of 2- vs 6-month patient reevaluation. *Chest* 1995;107:358-61.
- Goeree R, Hannah M, Hewson S. Cost-effectiveness of induction of labour versus serial antenatal monitoring in the Canadian Multicentre Postterm Pregnancy Trial. *Can Med Assoc J* 1995;152:1445-50.
- Rowley MJ, Hensley MJ, Brinsmead MW, Wlodarczyk JH. Continuity of care by a midwife team versus routine care during pregnancy and birth: a randomised trial. *Med J Aust* 1995;163:289-93.
- Fayers PM, Hand DJ. Generalisation from phase III clinical trials: survival, quality of life, and health economics. *Lancet* 1997;350:1025-7.
- O'Brien B. Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials. *Med Care* 1996;34 (suppl 12):DS99-108.
- Buxton MJ, Drummond MF, van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in economic evaluations: an unavoidable fact of life. *Health Econ* 1997;6:217-27.
- Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *BMJ* 1983;286:1489-93.
- Guidelines for referees. *BMJ* 1996;312:41-4.
- Altman DG. *Practical statistics for medical research*. London: Chapman and Hall, 1991.
- Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994;309:1286-91.
- Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313:275-83.
- Jefferson TO, Smith R, Yee Y, Drummond M, Pratt M, Gale R. Evaluating the BMJ guidelines for economic submissions: prospective audit of economic submissions to BMJ and the *Lancet*. *JAMA* 1998;280:275-7.
- Altman DG, Goodman SN. Transfer of technology from statistical journals to the biomedical literature: past trends and future predictions. *JAMA* 1994;272:129-38.
- Gore SM, Jones G, Thompson SG. The *Lancet's* statistical review process: areas for improvement by authors. *Lancet* 1992;340:100-2.
- Coyle D. Statistical analysis in pharmacoeconomic studies: a review of current issues and standards. *Pharmacoeconomics* 1996;9:506-16.
- Efron B, Tibshirani RJ. *An introduction to the bootstrap*. New York: Chapman and Hall, 1993.
- Lambert CM, Hurst NP, Forbes JF, Lochhead A, Macleod M, Nuki G. Is day care equivalent to inpatient care for active rheumatoid arthritis? Randomised controlled clinical and economic evaluation. *BMJ* 1998;316:965-9.
- Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Stat Med* 1996;15:1447-58.

(Accepted 8 October 1998)

A patient to remember A surfeit of experts

Many years ago, after attending an international congress of cardiology in Madrid, I went on a post-congress tour, which included Seville and its cathedral. Inside the cathedral is a sacristy as big as many a church, and while sitting in it listening to a lecture about the cathedral I became conscious of a commotion behind me, and went to investigate. A very elderly American had been taken ill with chest pain. He had been laid flat on his back on one of the benches and his wife, obviously much younger and fitter, was hovering over him with a tablet of nitroglycerine in her hand wondering what to do with it. Despite being a known sufferer from angina he had, with his wife, spent a very hot and exhausting morning doing the sights, of which this vast cathedral was the climax.

After being propped up and instructed to put the tablet under his tongue he recovered from the pain and went back to his hotel in a taxi. I visited him there in the evening and found him fully restored to his usual state of health. Both he and his

wife were very grateful and offered a fee, which I declined. They then remarked how very fortunate they had been to find a cardiologist so close to them in the church. Honesty compelled me to tell them that there had been 23 other cardiologists, all at least as competent, sitting on the other benches in front of me.

H A Dewar, retired cardiologist, Wylam, Northumberland

We welcome articles of up to 600 words on topics such as *A memorable patient, A paper that changed my practice, My most unfortunate mistake*, or any other piece conveying instruction, pathos, or humour. If possible the article should be supplied on a disk. Permission is needed from the patient or a relative if an identifiable patient is referred to. We also welcome contributions for "Endpieces," consisting of quotations of up to 80 words (but most are considerably shorter) from any source, ancient or modern, which have appealed to the reader.