Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting

Stavros Petrou,[1] Alastair Gray[2]

[1]Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK
[2]Health Economics Research Centre, Department of Public Health, University of Oxford, Oxford, UK
Correspondence to: S Petrou
S.Petrou@warwick.ac.uk

bmj.com
▶ See other articles in the comparative effectiveness research collection at http://www.bmj.com/cgi/collection/comparative_effectiveness_research

bmj.com
Previous articles in this series
▶ Comparisons within randomised groups can be very misleading (*BMJ* 2011;342:d561)
▶ Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting (*BMJ* 2011;342:d1766)
▶ Strengthening the reporting of genetic risk prediction studies: the GRIPS statement (*BMJ* 2011;342:d631)
▶ Correlation in restricted ranges of data (*BMJ* 2011;342:d556)

Collecting economic data at the same time as evidence of effectiveness maximises the information available for analysis but requires proper consideration at the design stage

Economic evaluation involves the comparative analysis of the costs and consequences of alternative programmes or interventions.[1] It has increasingly been used to inform decision making about healthcare in the United Kingdom and other industrialised nations.[2-5] Randomised controlled trials are commonly used as a vehicle for economic evaluations. Indeed, many funders, such as the UK National Institute for Health Research Health Technology Assessment Programme, routinely request that assessments of cost effectiveness are incorporated in the design of randomised trials. This article outlines some of the key issues concerning the design, conduct, analysis, and reporting of economic evaluations based on trials with individual patient data. Economic evaluations that synthesise data from disparate sources using decision analytical models (typically using summary rather than individual patient data) are discussed in an accompanying article.[6]

What are the objectives of economic evaluation?

Economic evaluations are typically concerned with two quantities: the additional cost of a new treatment compared with the existing alternative and the additional health benefits. If all the costs and outcomes relevant to this comparison can be measured, they can then be aver-

SUMMARY POINTS

Economic evaluation is increasingly used to inform the regulatory and reimbursement decisions of government agencies

Evaluations conducted alongside randomised controlled trials provide access to data on individual patients

This enables a wide range of analytical techniques—for example, to examine the relation between events of interest and health related quality of life

Designing a rigorous trial based economic evaluation requires close collaboration between trialists and health economists from the outset of the trial

Key issues concerning the design, conduct, analysis, and reporting of economic evaluations based on randomised trial with individual patient data are outlined
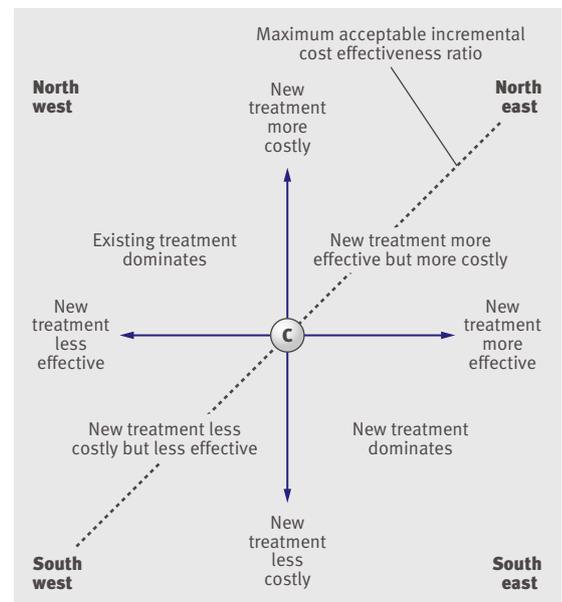
The cost effectiveness plane. The x axis shows the difference in effectiveness between the new treatment and the comparator and the y axis shows the difference in cost. The slope of the line from any point on the figure to the origin is the incremental cost effectiveness ratio

aged across all patients in the treatment (t) or the control (c) group to obtain mean cost C and mean effect E for each group. It then becomes possible to calculate a third quantity: the cost effectiveness of the new treatment compared with the alternative. The incremental cost effectiveness ratio (ICER) will simply be the difference in costs divided by the difference in effects: ICER$=C_t-C_c/E_t-E_c=\Delta C/\Delta E$.

This can be shown neatly on the cost effectiveness plane (figure).[7] In the south east quadrant of the figure the new intervention is less costly and more effective and (assuming there is no uncertainty surrounding the cost effectiveness ratio) should be adopted; equally, if the new intervention is less effective and more costly (the north west quadrant), it can readily be rejected. More controversially, new interventions may turn out to be more effective but also more costly (north east quadrant) or less effective but also less costly (south west quadrant): in either case, a trade-off then exists between effect and cost: additional health benefit can be obtained but at higher cost, or costs can be saved but only by giving up health benefit.

The question that then arises is whether the trade-off is acceptable: is the health gain (or cost saving) worth the

additional cost (or health loss)? To address this question, we can imagine a diagonal line running through the figure depicting the maximum we are willing to pay for a unit of effect (sometimes represented by the Greek letter λ). All points to the right of this line involve a trade-off between costs and health benefits that a decision maker might consider acceptable; while all points to the left require an unacceptable trade-off. The steeper the slope of this diagonal line, the more decision makers are willing to pay for a unit of effect. There are various techniques for estimating the value of λ.[8] In many jurisdictions, however, it has tended to reflect externally determined decision rules that have evolved historically and with little scientific basis.[9][10] The point to note is that economic evaluation involves two large uncertainties: where an intervention is located on the cost effectiveness plane, and how much a decision maker is willing to pay for health gains. Here we focus on the first question, and consider the use of trial based economic evaluations to obtain precise estimates of incremental cost effectiveness.

Design of trial based economic evaluations

Designing a rigorous trial based economic evaluation requires close collaboration between trialists and health economists. This collaboration should be reflected in the standard operating procedures of the coordinating clinical trials unit, including its data collection and informed consent procedures.[11] Where possible, the instruments and procedures used to collect economic data should be pilot tested for efficiency, clarity, and ease of use.[12] Pragmatic trials offer analysts an opportunity to evaluate the cost effectiveness of an intervention under real world conditions, with enrolled patients representative of typical clinical caseloads, a comparison of the intervention of interest with current practice, and follow-up under routine conditions.[1] With less naturalistic trial designs, such as explanatory trials designed primarily to answer questions about safety and efficacy, the generalisability of the economic evaluation may be impaired by stringent inclusion criteria and treatment protocols or by the absence of a "usual care" arm.[12][13] These limitations are hard to overcome, although protocol driven costs can and should be factored out of cost effectiveness calculations.[14]

Although several techniques have been proposed to estimate the appropriate sample size and statistical power for economic end points in randomised trials,[15-17] power calculations are almost invariably based on primary clinical outcomes. This is partly because of the complexity of trying to forecast the main outcome of interest to economists—the joint distribution of the difference in costs and benefits between treatment arms. Furthermore, large sample sizes may be needed to detect statistically significant differences because of the large variability in use of healthcare resources and cost measures,[18] and this may be neither financially nor ethically acceptable.[1] Economists therefore focus on estimating cost and effect differences and assessing the likelihood that an intervention is cost effective, rather than testing a particular hypothesis concerning cost effectiveness.[18]

Some people might wonder whether an economic evaluation should be included in a trial before we know whether the new treatment is more effective. However, if it is not included we risk losing the opportunity to collect information on use of resources and health related quality of life. Furthermore, as noted above, cost effectiveness is about the joint distribution of differences in cost and effect. This joint distribution could show clear cost effectiveness when neither cost nor effect differences are individually significant. Indeed, some economists have argued that reliance on traditional rules of statistical inference surrounding a single parameter, such as clinical effectiveness, is arbitrary and may result in inferior healthcare outcomes compared with decisions based on expected cost effectiveness.[1][19]

How are data measured and valued?
Use of resources
The resources used by patients, such as hospital admissions, consultations, and types and quantities of drugs administered, are normally recorded for each patient over the trial follow-up. The categories of resource use that are included in the study will be determined by the perspective of the analysis—whether it is confined to the healthcare system (sometimes referred to as the payer) or includes broader societal costs. The system perspective would typically include direct medical care, including the intervention itself, treatment of any side effects or complications, and follow-up care to the intervention or the underlying condition. It may also include medical care not directly associated with the underlying condition, although regression modelling may be required to disentangle background noise that often occurs when this is included.[20] In England and Wales, the National Institute for Health and Clinical Excellence (NICE) recommends including NHS and personal social services as a minimum.[2] The societal perspective also considers care provided by other sectors of the economy, costs incurred by patients, informal care provided by family and friends, and productivity losses from morbidity and premature death.

Use of many resources can normally be recorded on trial case report forms with little or no extra burden, but sometimes additional information will be required from medical records, patient questionnaires and diaries, and other sources.[21] A recent trial based economic evaluation of neonatal extracorporeal membrane oxygenation used observational research to estimate resource use associated with complications and parental questionnaires to document use of hospital and community health services after discharge.[22] Computerised record linkage may in future simplify the process of collecting these additional data.

When data are collected through patient or carer questionnaires, researchers have to balance recall bias against completeness of sampling information.[23] Shorter recall periods of a few days or weeks reduce the chance of a patient forgetting an episode or incorrectly recalling when it occurred, but if a study is trying to estimate resource use over a longer period, such as 12 months, sampling over a short recall period misses lots of data. It may be better to maximise completeness at the cost of some recall bias.[24]

Valuation of resource use

The total cost for an individual patient participating in a trial is the product of the quantity of each resource item they use and the unit cost of each item. Unit costs should theoretically be based on the economic notion of opportunity cost, which represents the value of the resource in its most highly valued alternative use.[1] In practice this is usually assumed to be approximated by nationally representative healthcare tariffs, such as the NHS payment by results tariffs[25] or the diagnosis related group payments in the US Medicare system.[26] However, unit costs are not always readily available from such sources and may have to be calculated using a combination of accounting data, time and motion studies, interviews with caregivers, and case note analysis.

Standard unit costs are often applied across all patients and trial centres, but non-standardised unit costs may be appropriate if the relative prices of factors such as labour and equipment vary between trial centres, especially in multinational trials.[27 28] All costs should be valued at the same price date, with adjustment using healthcare specific inflation indices when necessary.[29] Economic evaluations based on multinational trials should convert costs into a common currency by using purchasing power parity adjustments.[30]

Measurement and valuation of outcomes

Cost effectiveness may be reported in terms of many different outcome measures, ranging from biomedical markers to more final health outcomes.[12] The preferred outcome measure for many health economists, and many reimbursement agencies, remains the quality adjusted life year (QALY), a preference based measure of health outcome that combines length of life and health related quality of life.[31] For reimbursement agencies, the QALY has the advantage of allowing cost effectiveness comparisons between interventions for disparate health conditions. For economists, the QALY offers the additional advantage that it incorporates individual preferences for health outcomes, thereby moving beyond the narrow biomedical model for evaluative research.

To estimate QALYs patients typically complete at different time points a generic health related quality of life questionnaire with pre-existing preference weights that can be attached to each health state—for example EQ-5D,[32] the Health Utilities Index,[33] and the SF-6D.[34] The underpinning preference weights for these measures are generally drawn from surveys of the general population, and so descriptive data from patients are combined with health related quality of life weights (or utility scores) from the general population and survival data from the trial to generate QALY profiles. An alternative approach is to ask patients not only to describe their health status but also to value it using a complex scaling technique such as the standard gamble approach or time trade-off approach.[35] These techniques are more expensive and time consuming than using population weighting. For both approaches, the frequency and timing of assessments should be influenced by disease severity, speed of progression, and the questionnaire burden on patients.[36] When patients are too ill or do not have the cognitive competencies to complete a questionnaire, proxy measurements may be considered.[37]

If a trial has not included a quality of life assessment, mapping techniques can sometimes be used to predict preference based health related quality of life (or utility) scores based on responses to non-preference based measures.[38] This requires data from a separate study where both the preference based and non-preference based measures were completed. For example, Williamson and colleagues used data from a study that developed a utility algorithm for the OM8-30 otitis media with effusion questionnaire[39] to predict preference based health related quality of life scores, and consequently QALYs, for children participating in a randomised trial of topical intranasal corticosteroids for persistent bilateral otitis media with effusion.[40]

The QALY may be considered too restrictive or insensitive to the main outcomes of interest in some circumstances. Researchers are therefore developing instruments that try to measure broader outcomes such as attachment, security, enjoyment, role, and control that can still be used within an economic evaluation framework.[41]

Analysis and reporting of data

Discounting

Trial based economic evaluations often measure and value costs and outcomes over several years of patient follow-up. In this situation costs and outcomes that occur after the first year of follow-up are typically reduced by a discount factor so that they can be fairly compared. Some economists have argued for applying different discount rates to future costs and health outcomes,[42] but NICE recommends that economic evaluations conducted in England and Wales should discount both costs and outcomes at an annual rate of 3.5%.[2] Sensitivity analyses that test the effects of differential discount rates on costs and outcomes are recommended.[2]

Dealing with skewed, missing, and censored data

The costs and outcomes of the trial groups, and the respective differences between them, can be summarised in several ways. For costs, the crucial information is usually the arithmetic mean—that is, average cost—as this allows policy makers to estimate the total cost of implementing a programme or intervention.[43] Although cost data are often right skewed because a few patients use very high amounts of resource, producing a distribution that may violate the assumptions of standard statistical tests, it is not clear that the many alternative approaches suggested, such as bootstrapping, provide more reliable test results.[44] Indeed, simple approaches for analysing cost data that assume normal distributions may be preferable in large samples where the near-normality of sample means is assured, while relatively simple approaches, such as generalised linear models, may be sufficient in smaller samples to deal with problems such as skewness and excess zeros.[44] For clinical outcomes, it is generally most transparent to replicate the methods of the primary clinical analysis plan to summarise outcomes. Note,

however, that economists will be interested in all events, whereas trialists may be more interested in time to the first event. For QALYs, it may be important to adjust for baseline differences in health status between the trial groups.[45]

Missing data are a particular problem for economic evaluations alongside trials, as the analysis may be drawing on information from case report forms, adverse event data, medication files, health related quality of life and resource use questionnaires, and other sources at all points across the study. The problem may be mitigated, in part, by approaches such as postal or telephone reminders to patients. Nevertheless, a complete case approach might require that most patients are dropped and is seldom a realistic option. Instead, analysts increasingly favour multiple imputation, where multivariate regression techniques are used to predict missing values on the basis of existing data.[46] The precise approach will depend on the nature of the missing data: (i) missing completely at random, with no relation to the value of any other factors in the study population; (ii) missing at random, but correlated in an observable way with the mechanism that generates the outcome; and (iii) not missing at random, but dependent on unobserved variables.[46] A particular form of missing data is censoring, where information on some patients is truncated and not available for the full duration of interest. In the past, analysts often ignored the effects of censoring on the estimation of costs and outcomes, but survival analysis methods are increasingly being used to deal with this problem.[8]

Handling uncertainty

As we noted when discussing the cost effectiveness plane, trial based economic evaluations result in different types of uncertainty. Sampling (or stochastic) uncertainty, usually reported as a confidence interval, depends on variation in both the numerator (incremental cost) and the denominator (incremental effectiveness) of the incremental cost effectiveness ratio. However, ratios are difficult statistics to work with: in cost effectiveness ratios the denominator (the effect difference) may be zero, producing an intractable result. Another problem is that the same negative value might represent improved outcomes and lower costs or worse outcomes and higher costs. This makes confidence intervals hard to compute and to interpret.

An alternative is to assume that the decision maker's willingness to pay for health gain is known and then rearrange the ratio on a linear scale to see whether the intervention produces any net benefit. For example, if the actual health benefits produced by the intervention are multiplied by the assumed willingness to pay for these benefits, and the net costs are subtracted, we produce a linear scale where a negative is unambiguously bad (the costs outweigh the value placed on the health benefits) and larger benefits are unambiguously better.[47] This net benefit approach makes it straightforward to examine decision uncertainty—that is, uncertainty over the value of the willingness to pay for health gain (λ)—by assessing the probability that the new intervention is

cost effective across a range of values of λ. This is often displayed as a cost effectiveness acceptability curve.[48]

Heterogeneity in the trial population could also be explored by formulating a net benefit value for each patient from the observed costs and effects, and then constructing a regression model with a treatment variable and covariates such as age, sex, and disease severity. The magnitude and significance of the coefficients on the interaction between the covariates and the treatment variable might then provide an estimate of cost effectiveness by subgroup. Finally, methodological uncertainty, the uncertainty concerning issues such as the appropriate discount rate or cost perspective, can be explored by standard sensitivity analyses.

Longer term extrapolation

Cost effectiveness observed within a trial may be substantially different from what would have been observed with continued patient follow-up: for example, the benefits of reducing fatal outcomes typically continue well beyond the end of the trial. Consequently, extrapolation of cost effectiveness over an extended period, often a lifetime, is considered important.[36 49] Survival analysis models, such as the Cox proportional hazard model or Weibull model, are often used to estimate life expectancy with and without an intervention.[50] However, unbiased estimation of long term cost effectiveness may require more complex models of the disease process, accompanied by related information on the cost and utility of interventions and complications.

Trial based economic evaluations that use patient level data have important advantages in permitting the construction of such models and permitting them to be validated. A good example comes from the UK Prospective Diabetes Study.[51] The trial followed up patients for a median of 10 years. However, the wealth of individual patient data permitted the construction of a model consisting of a set of linked equations that predict the risks of major diabetes related complications, and this has been used to estimate the lifetime costs, utilities, and cost effectiveness of diabetes related interventions.[52] The interactions within the model between risk factors, individual characteristics, and disease history could not have been captured reliably without access to patient level data.

Advantages of trial based economic evaluations?

Economic evaluations conducted alongside randomised controlled trials provide an early opportunity to produce reliable estimates of cost effectiveness at low marginal cost. Access to individual patient data also permits a wide range of statistical and econometric techniques—for example, to examine the relation between events of interest and health related quality of life or to explore subgroup differences. Although trial based evaluations have limitations—truncated time horizons, limited comparators, restricted generalisability to different settings or countries, and the failure to incorporate all relevant evidence[53]—they are likely to continue to have an important role in producing reliable estimates of cost effectiveness.

1 Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart G. *Methods for the economic evaluation of health care programmes*, 3rd ed. Oxford University Press, 2005.
2 National Institute for Health and Clinical Excellence. *NICE guide to the methods of technology appraisal*. NICE, 2008.
3 Scottish Medicines Consortium. *Guidance to manufacturers for completion of new product assessment form (NPAF)*. SMC, 2007.
4 Pharmaceutical Benefits Advisory Committee. *Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee*. PBAC, 2002.
5 CDR Directorate at Canadian Coordinating Office for Health Technology Assessment. *Canada common drug review submission guidelines for manufacturers*. CCOHTA, 2003.
6 Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis and reporting. *BMJ* 2011;342:d1766.
7 Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;10:212-4.
8 Gray A, Clarke P, Wolstenholme J, Wordsworth S. *Applied methods of cost-effectiveness analysis in health care*. Oxford University Press, 2010.
9 Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ* 2004;13:437-52.
10 Grosse S. Assessing cost-effectiveness in healthcare: history of the $50,000 per QALY threshold. *Expert Rev Pharmacoeconomics Outcomes Res* 2008;8:165-78.
11 Edwards RT, Hounsome B, Linck P, Russell IT. Economic evaluation alongside pragmatic randomised trials: developing a standard operating procedure for clinical trials units. *Trials* 2008;9:64.
12 Glick HA, Joshi JA, Sonnad SS, Polsky D. *Economic evaluation in clinical trials*. Oxford University Press, 2007.
13 O'Sullivan AK, Thompson D, Drummond MF. Collection of health-economic data alongside clinical trials: iIs there a future for piggyback evaluations? *Value Health* 2005;8:67-79.
14 Coyle D, Lee LM. The problem of protocol driven costs in pharmacoeconomic analysis. *Pharmacoeconomics* 1998;14:357-63.
15 Briggs AH, Gray AM. Power and sample size calculations for stochastic cost-effectiveness analysis. *Med Decis Making* 1998;18:S81-92.
16 Gardiner JC, Sirbu CM, Rahbar MH. Updated on statistical power and sample size assessments for cost-effectiveness studies. *Expert Rev Pharmacoeconomics Outcomes Res* 2004;4:89-98.
17 Gafni A, Walter SD, Birch S, Sendi P. An opportunity cost approach to sample size calculation in cost-effectiveness analysis. *Health Econ* 2008;17:99-107.
18 Briggs A. Economic evaluation and clinical trials: size matters. *BMJ* 2000;321:1362.
19 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341-64.
20 Simon J, Gray A, Duley L, Magpie Trial Collaborative Group. Cost-effectiveness of prophylactic magnesium sulphate for 9996 women with pre-eclampsia from 33 countries: economic evaluation of the Magpie Trial. *BJOG* 2006;113:144-51.
21 Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. *Int J Technol Assess* 2005;21:165-71.
22 Petrou S, Bischof M, Bennett C, Elbourne D, Field D, McNally H. Cost-effectiveness of neonatal ECMO based on seven year results from the UK Collaborative ECMO Trial. *Pediatrics* 2006;117:1640-9.
23 Ridyard CH, Hughes DA. Methods for the collection of resource use data within clinical trials: a systematic review of studies funded by the UK Health Technology Assessment program. *Value Health* 2010;13:867-72.
24 Clarke PM, Fiebig DG, Gerdtham UG. Optimal recall length in survey design. *J Health Econ* 2008;27:1275-84.
25 Department of Health, England. *Payment by results*. Department of Health, 2010. www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/NHSFinancialReforms/DH_900.
26 Centers for Medicare and Medicaid Services. Acute inpatient PPS. www.cms.gov/InpatientPPS/.
27 Raikou M, Briggs A, Gray A, McGuire A. Centre-specific or average unit costs in multi-centre studies? Some theory and simulation. *Health Econ* 2000;9:191-8.
28 Torti FM, Reed SD, Schulman KA. Analytic considerations in economic evaluations of multinational cardiovascular clinical trials. *Value Health* 2006;9:281-91.
29 Curtis L. *Unit costs of health and social care 2009*. Personal Social Services Research Unit, University of Kent, 2009. www.pssru.ac.uk.
30 World Bank. *Global purchasing power parities and real expenditures 2005*. World Bank International Comparison Program, 2009. http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf.
31 Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care* 1989;5:559-75.
32 Brooks R. EQ-5D, the current state of play. *Health Pol* 1996;37:53-72.
33 Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions: Health Utilities Index. *PharmacoEcon* 1995;7:503-20.
34 Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271-92.
35 Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996;15:209-31.
36 Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health* 2005;8:521-33.
37 Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. *Measuring and valuing health benefits for economic evaluation*. Oxford University Press, 2007.
38 Brazier J, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ* 2010;11:215-25.
39 Dakin H, Petrou S, Haggard M, Benge S, Williamson I. Mapping analyses to estimate health utilities based on responses to the OM8-30 otitis media questionnaire. *Qual Life Res* 2010;19:65-80.
40 Williamson I, Benge S, Barton S, Petrou S, Letley L, Fasey N, et al. A double-blind randomised placebo controlled trial of topical intra-nasal corticosteroids in 4-11 year old children with persistent bilateral otitis media with effusion in primary care. *Health Technol Assess* 2009;13:1-274.
41 Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. *Soc Sci Med* 2008;67:874-82.
42 Brouwer WBF, van Exel NJA. Discounting in decision making: the consistency argument revisited empirically. *Health Policy* 2004;67:187-94.
43 Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ* 2000;320:1197-200.
44 Mihaylova B, Briggs A, O'Hagan A, Thompson S. Review of statistical methods for analysing healthcare resources and costs. *Health Econ* 2010, doi:10.1002/hec.1653.
45 Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;14:487-96.
46 Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
47 Zethraeus N, Johannesson M, Jonsson B, Lothgren M, Tambour M. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. *Pharmacoeconomics* 2003;21:39-48.
48 Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves: facts, fallacies and frequently asked questions. *Health Econ* 2004;13:405-15.
49 Hlatky MA, Owens DK, Sanders GD. Cost-effectiveness as an outcome in randomized clinical trials. *Clin Trials* 2006;3:543-51.
50 Mark DB, Hlatky MA, Califf RM, Naylor CD, Lee KL, Armstrong PW, et al. Cost-effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. *N Engl J Med* 1995;332:1418-24.
51 UKPDS Group. Cost effectiveness analysis of improved blood pressure control in hypertensive patients with type 2 diabetes: UKPDS 40. *BMJ* 1998;317:720-6.
52 Clarke PM, Gray AM, Briggs A, Stevens R, Matthews DR, Holman RR. Cost-utility analyses of intensive blood glucose and tight blood pressure control in type 2 diabetes (UKPDS 72). *Diabetologia* 2005;48:868-77.
53 Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Econ* 2006;15:677-87.