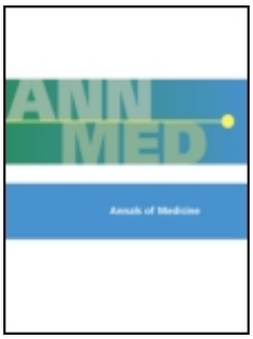

Annals of Medicine

ISSN: 0785-3890 (Print) 1365-2060 (Online) Journal homepage: informahealthcare.com/journals/iann20

Introducing economic and quality of life measurements into clinical studies

Michael Drummond

To cite this article: Michael Drummond (2001) Introducing economic and quality of life measurements into clinical studies, Annals of Medicine, 33:5, 344-349, DOI: 10.3109/07853890109002088

To link to this article: https://doi.org/10.3109/07853890109002088

Introducing economic and quality of life measurements into clinical studies

Michael Drummond

Although the collection of cost and quality of life data alongside clinical studies generates detailed patient level data in a timely fashion, it also raises practical and methodological challenges. These include the fact that the settings and patients enrolled in trials may not be typical of those found in regular clinical practice, that costs and quality of life may be influenced by the trial protocol, that the clinical alternatives compared in trials may not be the most relevant for cost-effectiveness assessments, that the length of follow-up may be too short to observe changes in cost and quality of life, and that adding these data will increase the overall measurement burden in the trial. This paper discusses these challenges and the ways in which they might be overcome, focussing particularly on preference-based measures of quality of life. In particular, recommendations are given for choosing the range of quality of life instruments, sample size calculations for quality of life measurement and the measurement of quality of life in multinational studies.

Keywords: cost-effectiveness analysis; quality-adjusted life-years; randomized controlled trials.

Ann Med 2001; 33: 344–349.

Introduction

Given the growing pressures on health care budgets, evidence on the relative value for money of health care treatments and programmes is becoming increasingly important. However, those wishing to consider value for money find that such evidence is often limited, as the majority of health care interventions have not been evaluated from a socioeconomic perspective. Specifically, information is required on the relative costs of alternative interventions and their impact on outcomes of value to the patient, such as improved quality of life.

It is attractive to consider collecting cost and quality of life data alongside clinical studies for two reasons. First, clinical studies allow us the opportunity to collect detailed, patient level data, either through the case report form (CRF) or patient diaries. The advantages of this for quality of life data are obvious. In respect of resource use and cost, patient level data enable us to capture those items of resource use that are likely to vary from patient to patient and that may not easily be extracted from routine data sources.

Secondly, clinical studies are often undertaken early in the life cycle of new health technologies before they have become widely adopted (ie, prior to the launch of new drugs). Socioeconomic data gathered at this stage may provide timely assessments of value for money, before important decisions about adoption of, and public payment for, the new treatments are taken.

Although in principle economic and quality of life measurements can take place alongside all types of clinical studies, the discussion here will focus on randomised controlled clinical trials, as these are generally regarded to be the preferred approach for producing unbiased estimates of the efficacy or effectiveness of health care interventions. (Such estimates are an important input to economic evaluations of health technologies). The discussion will also concentrate on the quality of life instruments presented in detail in this special section, which in the main are preference-weighted instruments producing a single quality of life index.

The paper is organised as follows. First, the practical and methodological challenges of introducing economic and quality of life data into clinical trials will be discussed. Then some specific issues relating to the use of quality of life instruments will be outlined. Finally, a number of conclusions will be drawn.

From the Centre for Health Economics, University of York, York, UK.

Correspondence: Michael Drummond, PhD, Centre for Health Economics, University of York, Heslington, York YO10 5DD, UK. E-mail: chedir@york.ac.uk, Fax: +44 1905 433644.

Practical and methodological challenges of introducing economic and quality of life measures into clinical trials

A number of challenges arise from the ways in which clinical trials are designed and conducted (1, 2). It is worth noting two points at the outset. First, many clinical trials are designed with a particular objective in mind, for example to provide efficacy data for a drug licensing application. Therefore the design of these trials might be constrained in various ways by the requirements of the licensing agency. Some of these constraints might lessen the suitability of the trial as a basis for undertaking socioeconomic measurements.

Secondly, one can consider trial designs to represent a spectrum, with the drug licensing study being at one end, and a trial undertaken in regular clinical practice at the other. Some use the term 'pragmatic trials' to describe those trial designs that seek, as far as possible, to mimic real life (3). Although all trials present challenges to those undertaking socioeconomic measurements, these are lessened in the case of pragmatic trials. Nevertheless, the issues we need to consider are as follows.

Selection of patients and settings

Depending on the nature of the trial, the patients and settings may be atypical of regular clinical practice. For example, many trials are conducted in specialist centres, using the latest equipment. There may also be fairly strict inclusion and exclusion criteria, with the consequence that the trial population may not be representative of the overall caseload. Finally, efforts may be made to ensure that patients (and their physicians!) comply with therapy.

Almost all of these features will have an impact on resource use, cost or quality of life. For example, treatment practices, the availability of resources and the prices/costs of resource items may be different in a specialist centre. In addition, efforts to ensure compliance may mean that it is difficult to explore the impact on cost and quality of life of a medication that has a better side-effect profile than existing treatments. In the trial, quality of life may be reduced in those patients given existing therapy because they are encouraged to continue on their medication. In real life they may cease medication, with serious consequences for efficacy and cost if their condition worsens.

In addition, the exclusion of patients with co-existing illness may mean that the changes in quality of life observed in the trial may not be seen in the broader clinical caseload. Finally, the closer monitoring and attention patients receive during the trial could affect quality of life and cost, especially if any deterioration in the patient's condition is detected earlier than it would be in regular clinical practice.

Key messages

- There are a number of advantages of collecting economic and quality of life measures in clinical trials.
- The design of many trials poses methodological challenges for economic and quality of life analysts.
- More research is required into the advantages and disadvantages of more pragmatic trial designs, sample size calculations for quality of life measures and the international transferability of tariff values in preference-based instruments.

Protocol-driven effects

The closer monitoring of patients during clinical trials is but one example of a protocol-driven effect. Other consequences of the trial protocol may include additional investigations and clinic visits. Of course, the activities solely undertaken as part of the trial protocol can be excluded from the costing, but some of the effects are quite subtle. For example, a patient experiencing a problem on (say) Day 4 may not make a visit to their physician if they have a protocol-related visit scheduled for Day 5. Therefore, exclusion of this visit from the costing may not be appropriate.

The effects of the protocol on quality of life measurements are similarly unclear. On the one hand, the extra care and attention received during the trial might improve the patient's quality of life. On the other hand, excessive clinic visits and investigations may reduce quality of life. Certainly it may not be wise to measure quality of life at the end of an intensive 4-hour clinic visit!

Selection of alternatives to be compared

In many clinical trials, especially registration trials for new pharmaceuticals, the comparator therapy is either placebo or an older, outmoded medication. Sometimes this choice is driven by the licensing agencies; sometimes it is driven by a desire to show an improvement from use of the new medication. For example, in one of the most famous quality of life studies ever conducted, Croog and co-workers (4) compared captopril an angiotensin-converting enzyme (ACE) inhibitor for treatment of hypertension with propan-

olol and methyldopa, not one of the more modern beta-blockers that would have been available at the time.

Such comparisons are not necessarily wrong or inappropriate; rather they limit the usefulness of the data to healthcare decision makers. The most relevant comparison is with current practice, or a widely-used therapy. However, whereas in principle most economic analysts would agree with this approach, 'current practice' may be a mixture of therapies which may be changing over time and may vary from setting to setting. Therefore, applying this principle in practice may sometimes be difficult.

Masking of the study

It is common for controlled clinical trials to be undertaken in a 'blinded' fashion. That is, neither the patient nor their physician is aware of the treatment assignment. The reason for this is obvious; ie, to reduce the chances of bias in the assessment of patient outcome.

In most cases one would want to preserve masking when quality of life is being measured, particularly because of the subjective nature of the assessments. However, it is virtually impossible to mask some studies (eg, comparisons between surgery and medical therapy). Also, masking can cause complications for the measurement of resource use because sometimes extra resources (eg, tests to monitor toxicity) are required in order to preserve the masking, even if one of the therapies being tested does not exhibit toxicity problems.

Other effects of masking are more subtle. For example, a given adverse event may be better tolerated by the patient if they are warned about it in advance. Also, there are occasions where knowing the treatment assignment may make the physician behave differently. For example, knowing that a sedated patient in the intensive care unit is likely to recover more quickly will enable the physician to plan for the patient's discharge to a normal ward. In a masked study this potential economic advantage could not be assessed.

Therefore, the masking of studies has both advantages and disadvantages for the measurement of socioeconomic parameters. A balance needs to be struck, and this probably depends on the stage of development of the technology and the extent of existing data from masked studies (5).

Length of follow-up

Many clinical trials are of short duration, primarily because of practical or financial considerations. In some cases (eg, treatment of perioperative infections) this might be quite appropriate, whereas in others (eg, treatment of Alzheimer's disease) it may not. This can be problematic if the changes in resource use and quality of life are only likely to be observed in the longer term. For example, if a patient is receiving nursing care and other social support at home, there might be a time lag between improvement in his or her condition and the reduction in services. Similarly, a patient suffering from epilepsy may have to wait for a period of time, after the condition has been controlled, to regain his or her driving licence. This may have an impact on quality of life, but would not be detected in a short-term study.

Another problem relates to the fact that, in many trials, the patient is considered to have ended the study when study medication is discontinued. For the researcher studying resource use or quality of life this may be a particularly interesting time, as efforts may be required to establish the patient on a new therapy or treat the consequences of an acute exacerbation. Therefore the emphasis would be on following up such patients rather than ignoring them.

Measurement burden

Finally, it is usually the case that measurement of quality of life or resource use is considered at the later stages of clinical trial design. Therefore, by the time that these additional measurements are being considered, the trial is typically over-complicated and over budget. This means that measurements of resource use and quality of life often come under intense scrutiny, with the analyst being asked to justify the range of measurements and the number of times they should be repeated.

In relation to quality of life instruments, it is important to know whether they can be self-administered and how long they take. (Most of the instruments being considered in this special section are not overly burdensome, see (6)). However, it is important to consider carefully the additional advantages and disadvantages of the more frequent administration of quality of life instruments.

With respect to measurements of resource use, the choices centre around the level of detail required. For example, is it sufficient to measure only the major items, such as days in hospital? Also, is it important to record each item of resource consumed by the patient in hospital, or is it sufficient merely to record the category of ward they are admitted to. Sometimes these choices depend on the detail of unit cost or price data available; there is little point in collecting lots of detail on resource use if it is not possible to cost it.

On occasions debates also take place about the range of resource data capture. For example, resource consumption in the hospital or clinic is usually easy to measure. Resource consumption in community care, or by the patient and their family, is often harder to

track and it may be necessary to rely on patient diaries.

Towards a resolution of these issues: ideals and compromises

The ideal solution would be to undertake resource use and quality of life measurements alongside trials with a more pragmatic design. These trials would: 1) be undertaken in 'real-life' settings; 2) enrol a broad range of patients; 3) have an unintrusive protocol; 4) compare the new intervention with a widely used alternative; 5) include a broad range of socioeconomic measurements; 6) follow patients for a reasonable amount of time, irrespective of whether they discontinue study treatment.

However, such trials are costly and time consuming to conduct and, in the case of drugs, do not remove the need to perform traditional efficacy studies for licensing purposes. Nevertheless, they should be encouraged wherever possible.

In situations where pragmatic trials are not feasible, analysts wishing to introduce economic and quality of life measurements into clinical studies may have to compromise on some of the methodological issues outlined above. Also, it is sometimes possible to adjust data collected during the trial to reflect regular clinical practice. However, it is much easier to see how adjustments could be made to the resource use data (eg, by excluding protocol-driven costs) than to the quality of life data. Therefore the artificial nature of some clinical trials is something that the quality of life analyst may have to live with.

Specific issues relating to the use of quality of life instruments

Given the particular focus of this special section on several specific quality of life instruments, it is worth exploring a number of issues in more detail. These are: choosing the range of quality of life instruments; sample size calculations for quality of life measurement and quality of life measurement in multinational studies. In exploring these issues I am grateful for information provided by analysts closely associated with the instruments concerned (see Acknowledgements).

Choosing the range of instruments

Figure 1 illustrates that potentially a number of different types of quality of life measures can be introduced into a clinical trial. The first point to note is that the different types of measure have different primary advantages. For example, a disease-specific scale may have the maximum responsiveness to change, whereas a 'utility' or preference-based measure may have the potential to influence public policy and resource allocation decisions, as it enables quality-adjusted life-years (QALYs) to be calculated.

One approach would be to include the full range of quality of life measures, on the grounds that they perform different tasks. This is the approach suggested by the Canadian Co-ordinating Office for Health Technology Assessment (8) in its guidelines for economic evaluation. However, it was mentioned above that there are sometimes concerns about the

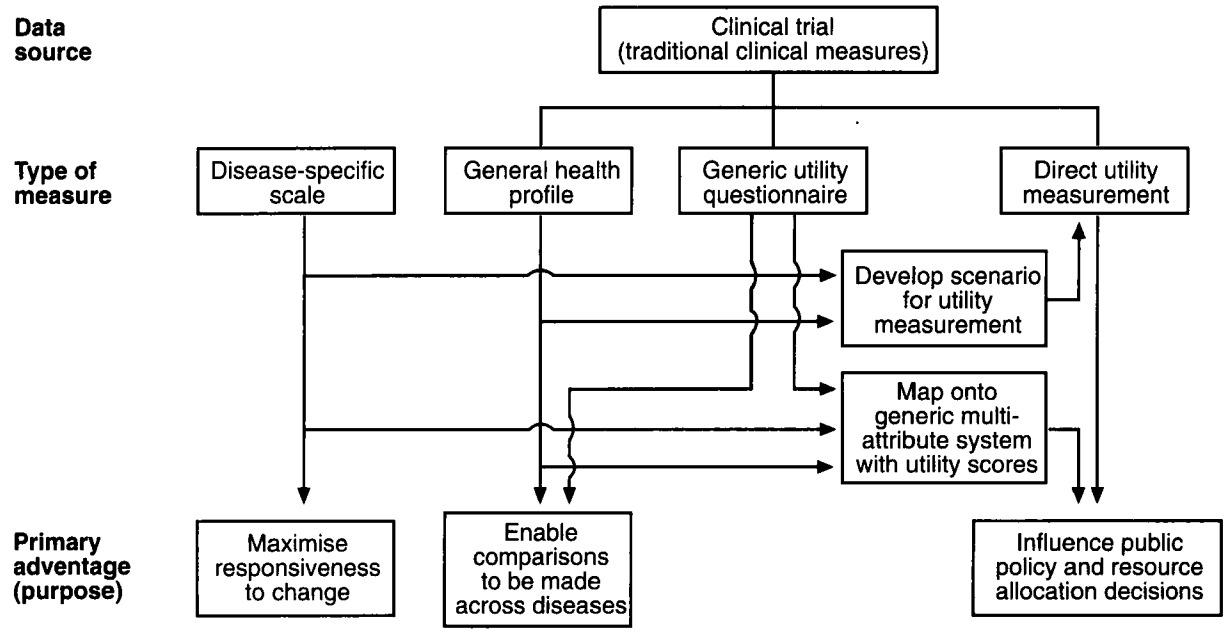

Figure 1. Interrelationships between health-related quality of life measures. (Adapted from (7) with permission).

measurement burden in clinical trials. Therefore choices among the measures will sometimes have to be made.

Of the measures being considered in this special section, all are generic measures, with all but one (the SF-36) being preference-based. (However, some work is underway to develop a preference-based measure, using the SF-36 as a starting point.) Therefore if we did choose to use one of these measures in a given clinical study, is there any guidance on other measures to include or exclude?

In general, none of the proponents of the instruments of interest would definitely exclude other quality of life measures if these are consistent with the aims of the study. However, there is little point of including two instruments of the same type unless the interest is methodological. In addition, one might question the relevance of disease-specific or general health profile measures to economic evaluation, although they can be useful for assessing efficacy or effectiveness.

Table 1. Guidance on the minimum clinically important difference (MCID) and sample size calculations.

Instrument	
15D	A difference of ± 0.03 or more (on a scale 0–1) is one that people feel is better or worse. Has been used in sample size calculations.
EQ-5D	No statement on the MCID. Strategies for sample size calculations using the three parts of EQ-5D (see (10)).
HUI	A difference of 0.03 is considered important. Some studies report differences of 0.01 to be important. Sample size calculations have been used in a number of studies.
SF-36	A threshold of one point on a given scale is not particularly meaningful. Others have suggested that the MCID is in the range of three to five points. Sample size calculations have been performed.

HUI, health utilities index.

Sample size calculations for quality of life measurement

In clinical trials a sample size calculation is usually performed for the primary clinical endpoint. Typically, a minimum clinically or quantitatively important difference (δ) is specified; that is, the size of difference or change in the measure that would cause the clinician to change his or her treatment policy. Then sample size is calculated to enable this difference to be detected with a given statistical power (usually 80%) at the conventional level of statistical significance (usually 5%).

Although all three judgements (quantitative importance, required statistical power and level of statistical significance) are all essentially value judgements made by the analyst, two of them (power and significance level) seem to follow an agreed convention. On the other hand, reviews of judgements about quantitative importance (9) show these to be very variable, following no set logic.

In clinical studies, measurement of quality of life is usually a secondary endpoint. Thus the sample size calculation is performed on the primary endpoint in the hope that there is sufficient power to detect important differences in quality of life. However, on occasions there may be a case for making the quality of life measure the primary endpoint, in which case the minimum clinically important difference needs to be specified.

It can be seen from Table 1 that many of the instruments of interest here have featured in sample size calculations. In some cases statements have been made about what constitutes a minimum clinically important difference (MCID). In general, these have been justified by reference to the size of changes in a given scale that patients believe to be important to them. This is one area where more research is required.

Finally, in the case of preference-based measures one might argue that, as the ultimate objective is to influence resource allocation decisions, it is the difference in cost-effectiveness (eg, incremental cost per QALY) that is important, not the change in quality of life. Therefore changes in the measure alone may not be of interest without also considering the cost of bringing about those changes. Thus the sample size calculation, if one were performed, would be designed such that it would be possible to assess whether the incremental cost per QALY for the new treatment, compared with the existing one, is in the acceptable range (eg, less than US$ 50 000 or US$ 100 000). Nevertheless, it may still be relevant to specify a MCID for a given quality of life measure for 'clinical', as opposed to economic, reasons.

Quality of life in multinational studies

More and more clinical trials are being undertaken on a multinational level. The motivations for this are twofold. First, a large number of clinical centres may be required to recruit a sufficient number of patients in a reasonable amount of time. Secondly, it may be of interest to demonstrate the effects of the new treatment in a range of settings.

The problems of undertaking economic evaluations on an international level have been well documented (11), and often adjustments have to be made to resource use data in order to generalise from one setting to another. (Costs definitely have to be presented in local prices).

In principle, quality of life data may be more

transferable and, in a multinational study, it may be possible to pool the data from different countries. (This is the normal approach for the clinical data after tests of heterogeneity have been performed). For the measures discussed here the main limitation on their use in international studies would be the lack of availability of questionnaires in different languages.

However, it can be seen from Table 2 that many of the questionnaires are available in a number of languages. In general, no major problems have been experienced with using any of the instruments in other languages, although there is obviously much more to cultural adaptability than merely translating the instrument word for word. There are therefore a number of examples where modifications have been made to the phrasing of questions to make them more easily understood in a different culture.

In addition, for the preference-based instruments, it would be necessary to have the preference weights (or tariff values) available for the countries of interest, or some confirmation that the values do not differ very much from country to country. Much less known about this aspect of the international transferability of instruments, although the research to date suggests that the mean values for different health states does not vary greatly from place to place (13, 14).

Table 2. Use of instruments in multinational studies.

Instrument	
15D	Available in 15 languages. The tariff values are based on a survey in Finland.
EQ-5D	Currently available in 47 languages. Used in multinational studies by most major pharmaceutical companies. Also the ICS-BPH and ISAT studies. Comparisons being made of visual analogue scores in several countries.
HUI	Available in 12 or more languages, being used in more than 12 multinational clinical studies. Tariff values available for Canada and values being generated for other settings.
SF-36	Used in many languages and many countries, (see (12)).

Conclusion

There are a number of reasons why one would wish to introduce economic and quality of life measures into clinical studies. However, the design of many clinical trials poses methodological challenges for economic and quality of life analysts.

There is already considerable experience in using the instruments (considered in this special section) in clinical studies, and this suggests that many of the methodological challenges can be overcome. Areas where more research is required include: 1) the advantages and disadvantages of more pragmatic clinical trial designs; 2) sample size calculations for quality of life measures; and 3) the international transferability of tariff values in preference-based instruments.

I am grateful to the following individuals who provided information about the various quality of life instruments discussed in this paper: Frank de Charro, Bill Furlong, Ron Hays and Harri Sintonen.

References

1. **Drummond MF.** Economic analysis alongside clinical trials: problems and potential. *J Rheumatol* 1995; 22: 1403–7.
2. **Drummond MF, O'Brien BJ.** Economic analysis alongside clinical trials: practical considerations. *J Rheumatol* 1995; 22: 1418–9.
3. **Schwartz D, Lellouch J.** Explanatory and pragmatic attitudes in therapeutic trials. *J Chron Dis* 1967; 20: 637–48.
4. **Croog SH, Levine S, Testa MA, Brown B, Bulpitt CJ, Jenkins CD, et al.** The effects of antihypertensive therapy on the quality of life. *N Engl J Med* 1986; 314: 1657–64.
5. **Freemantle N, Drummond MF.** Should clinical trials with concurrent economic analyses be blinded? *JAMA* 1996; 277: 63–4.
6. **Nord E.** *A review of synthetic health indicators.* Background paper for the OECD Directorate for Education, Employment, Labour and Social Affairs. Paris: OECD, 1997.
7. **Drummond MF, Davies LM.** Economic analysis alongside clinical trials: revisiting the methodological issues. *Int J Technol Assess Health Care* 1991; 7: 561–73.
8. **Canadian Coordinating Office for Health Technology Assessment (CCOHTA).** *Guidelines for economic evaluation of pharmaceuticals: Canada.* 2nd edn. Ottawa: CCOHTA; 1997.
9. **Burnand B, Kerman WN, Feinstein AR.** Indexes and boundaries for 'quantitative significance' in statistical decisions. *J Clin Epidemiol* 1990; 43: 1273–84.
10. **Roset M, Badia X, Mayo NE.** Sample size calculations in studies using the EuroQoL-5D. *Qual Life Res* 1999; 8: 539–49.
11. **Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW.** *Methods for the economic evaluation of healthcare programmes.* Oxford: Oxford University Press; 1997.
12. **Gandek B, Ware JE, eds.** Translating functional health and wellbeing: International Quality of Life Assessment (IQOLA) project studies of the SF-36 Health Survey. *J Clin Epidemiol* 1998; 51: Special Issue, Nov.
13. **Johnson JA, Ohinmaa A, Murti B, Sintonen H, Coons SJ.** Comparison of Finnish and US-based visual analog scale valuations of the EQ-5D measure. *Med Decis Making* 2000; 20: 281–9.
14. **Le Gales C, Buron C, Costel N, Rosman S, Slama G.** The French Health Utilities Index Mark 3 (Abstract). *Value in Health* 2000; 3: 103.