Problems in conducting economic evaluations alongside clinical trials

Lessons from a study of case management for people with mental disorders

A. M. GRAY, M. MARSHALL, A. LOCKWOOD and J. MORRIS

Background Case management has become the statutory basis of community care in the UK for people with long-term mental disorders, although a randomised controlled trial found no important improvements over standard care. Here we compare the costs and cost consequences of this intervention with standard care.

Method Resource-use data were collected over a six-month baseline period and for 14 months after randomisation on all patients in the trial.

Results At 14 months the ratio of control group to treatment group weekly costs was 1.09 (95% CI 0.86–1.38) for total costs; 1.12 (0.76–1.65) for state benefits, and 1.21 (0.61–2.42) for health care costs. Costs were thus lower in the treatment group, but these differences were not significant.

Conclusions Retrospective power calculations indicated that the trial could have detected differences of 30% in total cost, but would have required 700 patients per arm to detect a 20% difference in health care costs. Hence this study, which had adequate power to detect clinically meaningful differences, was found to be far too small to detect large differences in costs. Funding agencies increasingly request that clinical trials include economic alongside clinical end-points: these findings may have important lessons for that policy.

Funding agencies increasingly request that economic evaluations are performed alongside clinical trials (Advisory Committee on Health Technology Assessment, 1992). The economic evaluation reported here was conducted alongside a randomised controlled trial of social services case management, a cornerstone of community care for people with mental disorders in the UK. The trial found that social services case management was largely ineffective (Marshall *et al*, 1995*a*). Case management was introduced partly to obtain "cost improvement" (Griffiths, 1988), defined as "the most cost-effective package of services to meet the wishes of those being helped". However, some researchers have suggested that costs may rise as case managers seek to maximse their clients' welfare benefits (Holloway *et al*, 1991), and one American study found evidence of considerable increases in direct care costs (Franklin *et al*, 1987).

The present economic evaluation therefore set out to compare the costs of the case management and control (standard care) groups in the trial, to determine whether or not case management added significantly to: the total direct costs of all care provided; the amount of benefits received; and the costs of health care provided. During the study, some important questions arose about the power of such a study to detect meaningful differences in costs. These questions have not been alluded to in previous studies in this area, but raise doubts about the feasibility of performing economic evaluations alongside clinical trials that have been designed using power calculations based on clinical rather than economic outcomes.

METHOD

Trial design

Details of the trial design have been published (Marshall *et al*, 1995*a*,*b*). In brief, subjects were referred from hostels for the homeless; night shelters; a general practice clinic for the homeless; the City Council homelessness unit; and local voluntary sector group homes. Subjects were suffering from a severe, persistent mental disorder, and were either homeless, at risk of homelessness, or living in temporary, supported, or poor-quality accommodation, where they were coping poorly, experiencing social isolation, or causing disturbances. At the baseline assessment, subjects had high levels of psychopathology and social disability. Subjects were assessed before allocation to treatment or control groups, seven months after entering the study, and 14 months after entering the study. Subjects were randomised to treatment (case management, n=40).

Subjects allocated to the case management group received, as a minimum: an assessment of need from a case manager; a discussion of the findings of this assessment with the subject's carer; intervention from the case manager to meet needs that were identified; monitoring of the subject's progress by the case manager; and further assistance from the case manager should further needs arise. Hence, case managers in this trial were acting as providers of care, focusing mainly on direct work with clients and the coordination of individual 'packages' of care, rather than as purchasers using their own devolved budgets to purchase suitable care. Subjects allocated to the control group continued to receive any assistance that they had been receiving before the start of the study.

Use of resources

Resource use data were collected for each individual enrolled in the study on a prospective basis. Information on resources used was collected by means of a data collection schedule for each subject enrolled in the trial, which covered:

(1) non-psychiatric health care (for example, general practitioner (GP) consultations, prescribed medications, hospital attendances);

(2) psychiatric health care (for example, occupational therapy, psychiatric day care, hospital care, domiciliary visits);

(3) local authority welfare services (for example, local authority day centre, social worker visits, child in local authority care, travel concessions);

(4) state benefits (for example, income support, disability premium, housing benefit);

(5) accommodation;

(6) contacts with law enforcement agencies;

(7) employment services in rehabilitation and other subsidised work settings; and

https://doi.org/10.1192/bjp.170.1.47 Published online by Cambridge University Press

(8) resources associated with untoward and non-routine events involving subjects in the study.

The resource input from the case managers was measured in terms of each case manager's time. This was recorded in diaries kept by the case managers throughout the study, in which they noted the duration and frequency of work on behalf of each subject. These data were supplemented by information from hospital notes, carers and GP records.

Costing resources

Resources used were valued on the basis of their opportunity cost (Drummond, 1980; Knapp, 1993). Wherever possible, cost data provided by specific institutions were used. For example, all agencies and organisations providing accommodation were approached with requests for costs, and almost all were able to provide this information. Where specific cost information was not available, regional or national unit costs were used.

Where possible, all costs incorporate a component reflecting the capital employed, although different institutions may have used slightly different procedures in calculating these capital costs. Costs were expressed in units appropriate to the resource use measures: visits, attendances, in-patient episodes, case management minutes, etc. All costs are expressed in 1993–1994 prices. Where it was necessary to convert prices to this standard, the Hospital and Community Health Services Pay and Prices Index was used.

Details of some of the most important unit costs used in the study, together with sources, are given in Table 1. Some unit costs relating to accommodation and other services have not been given in detail as they were provided to the investigators on a confidential basis. In total, the study made use of unit costs for 91 separate items of service, and specific rates for all relevant state benefits.

Collecting data on an individual basis made it possible to explore cost variations between individuals in the study and to analyse apparent differences between the control and treatment group using formal statistical tests (advantages which are increasingly acknowledged but are realised in a very small proportion of all economic evaluations; see Drummond & O'Brien *et al*, 1994).

Analysis

Analysis was performed on an intention-to-treat basis. Cost analysis is based on recorded resources used expressed as weekly costs averaged over the first seven and the first 14 months following randomisation. In all, data were collected on 80 subjects. Some or all information on resource use was missing for 10 patients at the seven-month assessment point and for 19 patients by the 14-month assessment point. Weekly costs over the baseline period for completers and non-completers were compared, and no significant differences were found. Two commonly used methods for dealing with drop-outs – completers' analysis, and analysis based on last observation carried forward – were compared (Heyting *et al*, 1992). Results including substituted values for non-completers did not differ significantly from the completers-only results. The primary results reported below are based on completers.

All resource and cost data were assembled in linked spreadsheets using EXCEL, version 5.0 (Microsoft Corp.), and analysed using SPSS for Windows.

RESULTS

Variation in costs

All cost data were found to have a large standard deviation and to be highly positively skewed. The coefficient of variation (standard deviation/mean) was 0.58 for total costs, 0.41 for state benefits and 2.0 for health care costs. A negligible proportion of the overall cost variation was contributed by untoward events (only one untoward event occurred; a non-accidental fire caused by a member of the control group led to damage estimated at £1200). Skewing was most pronounced in the health care costs. All cost data were therefore transformed to natural logarithms before further analyses.

Costs per subject

Table 2 shows total costs per subject, expressed as the (geometrical) mean weekly cost per subject in the case management and control groups, at baseline assessment and over the trial period. Small differences in the mean costs of the control and case management groups existed at baseline, but these were not significant. In the control group total costs averaged approximately £275 per week throughout the trial, and in the case management group these costs had fallen to under £250 per week by month 14 of the trial. However, none of the differences observed, either between groups or over time, was significant, as indicated by the wide and overlapping confidence intervals.

Figure 1 shows these between-group differences in weekly total costs, benefits and health care costs at the seven- and 14-month follow-up points, expressed as ratio of control to case management group, so that a ratio above one indicates that costs are higher in the control group (Gardner & Altman, 1989). As the figure indicates, at 14 months the ratio of control group to treatment group weekly costs (with 95% confidence interval) was 1.09 (0.86–1.38) for total costs; 1.12 (0.76–1.65) for state benefits and 1.21 (0.61–2.42) for health care costs. These ratios indicate that each category of cost was lower in the treatment group than in the control group, but the 95% confidence intervals include one in all instances. The null hypothesis, that social service case management does not significantly add to the costs of care (in terms of total costs, state benefits or health care costs) therefore cannot be rejected.

Adjusting for the differences in the costs at the baseline assessment by performing an analysis of covariance with baseline costs as covariates also did not find any significant differences in costs between the two groups, over either the seven- or the 14-month follow-up period.

DISCUSSION

Interpretation of results

Substantial resources are involved in maintaining psychiatric patients in community care. Over the period of the baseline assessment and trial, the total costs incurred by the 80 subjects entered into the trial amounted to approximately £1.8 million, almost 20% of which was attributable to health care and 30% to other services (the remaining 53% being accounted for by state benefits). Psychiatric community care is therefore an important area for research into costs and effectiveness.

There is no consensus among economists on what might constitute a cost difference of practical importance: Drummond & O'Brien (1993) argue that there is no *a priori* reason to suppose that the 10 or 15% differences commonly applied to clinical situations will be appropriate in economic contexts, and suggest that any difference greater than the costs of changing to a new form of practice would be worthwhile. However, very large samples may be required to demonstrate that small differences are statistically significant. If a 20% difference is taken as a conservative estimate of a practically important difference in the costs of psychiatric community care,

https://doi.org/10.1192/bjp.170.1.47 Published online by Cambridge University Press

Table 1 Unit costs and data sources for main treatment and care services

Item	Unit	Cost (£s 1993–1994)	Source
Accommodation			
Church housing (registered)	Mean cost per person per day	30.01	Provider accounts
Night shelter	Mean cost per person per day	20.00	Provider accounts
Council homeless hostel	Mean cost per person per day	57.21	Oxfordshire county average, assuming 90% occupancy
Staffed group home A	Mean cost per person per day	34.31	Provider accounts
Staffed group home B	Mean cost per person per day	43.89	Provider accounts
Employment			
Rehabilitation workshop	Mean cost per person per day	24.60	Provider accounts
Employment training	Mean cost per person per day	9.57	Provider accounts
Law and order			
Arrests	Mean cost per arrest	97.84	Estimated from Thames Valley Police statistics
Prison	Mean cost per person per day	97.73	Estimated from Annual Abstract of Statistics, numbers in custody and prison expenditure
Court appearances	Mean cost per appearance	572.00	From CIPFA Administration of Justice Statistics, Oxfordshire
Probation day centre	Mean cost per person per attendance	163.39	Provider accounts
Local welfare services			
Psychiatric day care	Mean cost per visit	29.00	PSSRU Unit Costs of Community Care 1994
Home help	Mean cost per visit	3.41	PSSRU Unit Costs of Community Care 1994
Child in care	Mean cost per day	39.71	PSSRU Unit Costs of Community Care 1994
MHA tribunal	Mean cost per case	2553.00	From Blumenthal & Wessley (1994)
Non-psychiatric health care			
GP consultation (surgery)	Mean cost per consultation	11.9	Dept of Health Departmental Report 1994
General hospital out-patient department	Mean cost per visit	49.08	Health and Personal Social Service Statistics 1989
General hospital in-patient care	Mean cost per in-patient day	153.70	Dept of Health Departmental Report 1994
Occupational therapy	Mean cost per domiciliary visit	23.00	PSSRU Unit Costs of Community Care 1994
Psychiatric health care			
Psychiatric in-patient care	Mean cost per in-patient day	93.00	PSSRU Unit Costs of Community Care 1994
Consultant domiciliary visit	Mean cost per visit	25.61	Estimated from provider accounts
Psychiatric nurse domiciliary visit	Mean cost per visit	12.93	Estimated from provider accounts
Care management			
Care management time	Mean cost per minute	0.25	Trial accounts
Miscellaneous items			
False teeth	Cost of item	45.00	Provider information
Orthopaedic boot	Cost of item	95.00	Provider information
Prosthetic leg repair	Cost of repair	270.00	Provider information
Non-accidental fire damage to accommodation	Cost of damage and brigade time	1200.00	Provider information

https://doi.org/10.1192/bjp.170.1.47 Published online by Cambridge University Press

Table 2 (Geometrical) mean weekly costs (95% CI) in case management and control groups at baseline and over the trial period (in 1994 £s; completers only)

	Total costs	State benefits	Health care costs
Baseline (−6 to 0 months)			
Control (n=30)	272 (220–337)	142 (122–167)	24 (14–42)
Case management (n=31)	256 (220–298)	126 (88–180)	16 (9–28)
First follow-up (0–7 months)			
Control (n=30)	277 (232–331)	157 (139–179)	34 (19–61)
Case management (n=31)	261 (225–303)	133 (93–190)	31 (21–47)
Second follow-up (0–14 months)			
Control (n=30)	272 (224–329)	149 (128–172)	34 (19–61)
Case management (n=31)	249 (215–288)	132 (92–189)	28 (18–42)

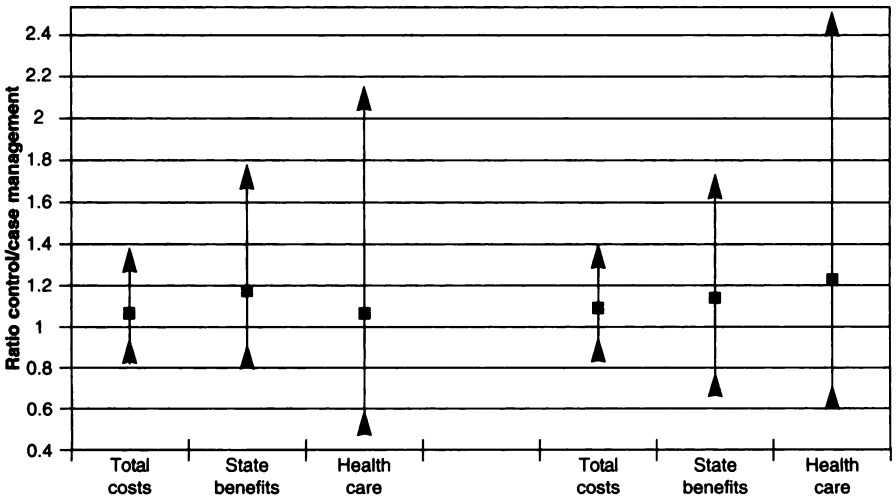

Fig. 1 Ratio of weekly costs in control group to those in case management group: geometrical mean and 95% CI (completers only) at seven-month (*left*) and 14-month (*right*) follow-up. Horizontal lines are separated by equivalent to 20% difference in costs.

the differences indicated in Figure 1 are not statistically significant, but may be large enough to be important (Armitage & Berry, 1994). For total costs, the confidence intervals suggest that the effects of treatments received by the control group are likely to lie somewhere between 13% less expensive to approximately 39% more expensive than case management. For state benefits and for health care costs, the confidence intervals are even wider.

Study power

These findings raise important questions about the power of such a study to detect differences in costs. The initial power calculations for the trial were based on the clinical outcome measures for behaviour and mental state, and the economic evaluation was designed within these power calculations: this is standard practice in clinical trials at present. The study was middle-ranking in size compared with other trials in this area that contained an economic component.

To explore the power of the study more formally, it was decided to use the log-transformed cost data to calculate the size of study that would be necessary to have an 80% chance of detecting a range of cost differences between the two groups at the 5% level of significance. To our knowledge this has not been done before in the published literature in this clinical area. The results of these retrospective power calculations are given in Figure 2. They show that the present study had sufficient power to detect a between-group difference in total cost of approximately 30%, but that sample sizes in the region of 160 subjects per arm would have been required to detect a 20% difference in state benefits, and over 700 per arm would have been necessary to detect a 20% difference in health care costs. For smaller differences in health care costs the required sample sizes would have been very much larger – over 3000 per arm for a difference of 10%, for example.

Previous research

Such power calculations were not possible before this study began, as no information was then available on the likely variation or distributional form of the cost data, and the existing literature in this particular field gave no indication of these difficulties. The results of these power calculations may throw some light on the findings of four out of five previously published UK studies, in which the costs of some form of community treatment were compared with those of a control group (Mangen *et al*, 1983; Tarrier *et al*, 1991; Burns *et al*, 1993; Knapp *et al*, 1994; Muijen *et al*, 1994). The magnitude of these five trials ranged from 19 to 94 subjects per treatment arm. Review of these studies revealed that in four out of five instances, apparently substantial cost differences between treatment arms (of 28% (Muijen *et al*, 1994), 42% (Tarrier *et al*, 1991), 47% (Burns *et al*, 1993) and 55% (Mangen *et al*, 1983)) were found not to be statistically significant. Using the health costs data from our study, sample sizes per treatment arm of up to 330 patients would have been required to demonstrate that such differences were significantly different. It is of interest that only one of the five studies (Burns *et al*, 1993) carried out a log transformation of the data before statistical analysis.

Implications for future trials

These findings suggest that more consideration needs to be given to the design of controlled trials in this area that propose to use economic data as an outcome measure. First, much more thought is required regarding what size of cost difference could be considered as worth achieving. Second, the costs and logistical complexity of mounting very large studies in order to have a chance of showing significant cost differences may sometimes be too daunting to contemplate: a controlled trial of case management involving 1500 patients might be a case in point. The cost-effectiveness of economic evaluations in such circumstances

https://doi.org/10.1192/bjp.170.1.47 Published online by Cambridge University Press

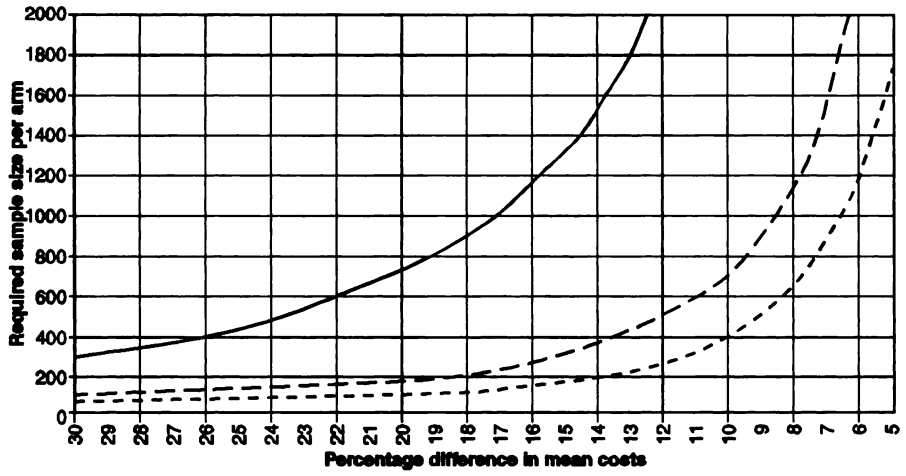

Fig. 2 Illustrative power calculations for cost data: required sample sizes for 80% power, 5% significance over a range of differences in mean costs. —, health care costs; – – –, state benefits; - - -, total costs.

may have to be considered. Third, there may be ethical implications of adding an economic component to a study if the cost end-points require a larger sample size than the clinical end-points, raising the possibility of a trial continuing for economic reasons beyond the point at which a clinical effect had been demonstrated. O'Brien *et al* (1994) have suggested that this may be acceptable in non-life-threatening circumstances. However, others may find such a situation unacceptable, for example, if there are significant differences in morbidity or health-related quality of life.

The proportion of all clinical trials containing some economic component remains low but is rising (Adams, 1992). Thus, the problems encountered in the present study are likely to be or to become familiar to other researchers conducting economic evaluations alongside clinical trials. When encouraging researchers to incorporate economic evaluations in clinical trials, funding bodies should carefully consider the practical, methodological and ethical issues that may arise.

ACKNOWLEDGEMENTS

The authors thank all those case managers, carers, nurses, doctors and subjects who kindly agreed to assist or participate in the study. M. M. was supported by the Wellcome Trust. The study was funded by the Medical Research Council.

REFERENCES

Adams, M. (1992) Economic analysis in randomized controlled trials. *Medical Care*, **30**, 231–244.

Advisory Committee on Health Technology Assessment (1992) *Assessing the Effects of Health Technologies: Principles, Practice and Proposals*. London: Her Majesty's Stationery Office, for the Department of Health.

Armitage, P. & Berry, G. (1994) *Statistical methods in medical research* (3rd edn). Oxford: Blackwell Scientific.

Burns, T., Raftery, J., Beadsmoore, A., *et al* (1993) A controlled trial of home-based acute psychiatric services. II: Treatment patterns and costs. *British Journal of Psychiatry*, **163**, 55–61.

Drummond, M. (1980) *Principles of Economic Appraisal in Health Care*. Oxford: Oxford University Press.

— & O'Brien, B. (1993) Clinical importance, statistical significance and the assessment of economic and quality-of-life outcomes. *Health Economics*, **2**, 205–212.

—, Stoddart, G. & Torrance, G. (1987) *Methods for Economic Evaluation of Health Care Programs*. Oxford: Oxford University Press.

Franklin, J., Solovitz, B., Mason, M., *et al* (1987) An evaluation of case management. *American Journal of Public Health*, **77**, 674–678.

Gardner, M. J. & Altman, D. G. (1989) *Statistics with Confidence – Confidence Intervals and Statistical Guidelines*. London: British Medical Journal.

Griffiths, R. (1988) *Community Care: Agenda for Action*. London: Her Majesty's Stationery Office.

Heyting, A., Tolboom, J. T. M. & Essers, J. G. A. (1992) Statistical handling of drop-outs in longitudinal clinical trials. *Statistics in Medicine*, **11**, 2043–2061.

Holloway, F., McLean, E. K. & Robertson, J. A. (1991) Case management. *British Journal of Psychiatry*, **159**, 142–148.

CLINICAL IMPLICATIONS

- Cost data in studies of community care are likely to have large variances, to be highly skewed and to require log transformation.
- Power calculations based on these log-transformed cost data indicate that the sample size required to detect a meaningful difference in costs may be much larger than that required to detect a clinically meaningful difference.
- Many trials in community care are inadequately powered to detect meaningful cost differences.

LIMITATIONS

- The case managers in the study were not budget-holders, as is now sometimes the case. This may affect their incentives to alter costs of care.
- The variations in costs detected in this study may be less in other care settings.
- Research-funders may be reluctant to support larger trials simply to obtain economic data

ALASTAIR GRAY, DPhil, Wolfson College, University of Oxford; MAX MARSHALL, MRCPsych, Department of Psychiatry, University of Manchester; AUSTIN LOCKWOOD, RMN, Department of Psychiatry, University of Manchester; JOAN MORRIS, Wolfson Institute of Preventive Medicine, Medical College of St Bartholomew's Hospital, University of London

Correspondence: Dr Alastair Gray, Health Economics Group, Centre for Socio-Legal Studies, Wolfson College, University of Oxford, Oxford OX2 6UD. Fax: 01865 284221; e-mail: alastair.gray@wolfson@ox.ac.uk

(First received 26 April 1996, final revision 23 July 1996, accepted 9 August 1996)

Knapp, M. (1993) Principles of applied cost research. In *Costing Community Care: Theory and Practice* (eds A. Netten & J. Beecham). Aldershot: Ashgate.

—, Beecham, J., Koutsogerogopoulou, V., *et al* (1994) Service use and costs of home-based versus hospital-based care for people with serious mental illness. *British Journal of Psychiatry*, **165**, 195–203.

Mangen, S. P., Paykel, E. S., Griffith, J. H., *et al* (1983) Cost-effectiveness of community psychiatric nurse or out-patient psychiatrist care of neurotic patients. *Psychological Medicine*, **13**, 407–416.

Marshall, M., Lockwood, A. & Gath, D. (1995*a*) How effective is social services case management for people with long-term mental disorders? A randomised controlled trial. *Lancet*, **345**, 409–412.

—, Hogg, L., Lockwood, A., *et al* (1995*b*) The Cardinal Needs Schedule: a modified version of the MRC Needs for Care Schedule. *Psychological Medicine*, **25**, 605–617.

Muijen, M., Cooney, M., Strathdee, G., *et al* (1994) Community psychiatric nurse teams: intensive support versus generic care. *British Journal of Psychiatry*, **165**, 211–217.

O'Brien, B., Drummond, M., Labelle, R. J., *et al* (1994) In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care*, **32**, 150–163.

Tarrier, N., Lowson, K. & Bow, C. (1991) Some aspects of family interventions in schizophrenia. II: Financial considerations. *British Journal of Psychiatry*, **159**, 481–484.

https://doi.org/10.1192/bjp.170.1.47 Published online by Cambridge University Press