

Assessing the evidence on case management

S. J. ZIGURAS, G. W. STUART and A. C. JACKSON

Background Evidence on the impact of case management is contradictory.

Aims To discuss two different systematic reviews (one conducted by the authors and one conducted through the Cochrane collaboration) that came to contradictory conclusions about the impact of case management in mental health services.

Method We summarised the findings of the two reviews with respect to case management effectiveness, examined key methodological differences between the two approaches and discuss the impact of these on the validity of the results.

Results The differences in conclusions between the two reviews result from the differences in inclusion criteria, namely non-randomised trials, data from unpublished scales and data from variables with skewed distributions. The theoretical and empirical effects of these are discussed.

Conclusions Systematic reviewers may face a trade-off between the application of strict criteria for the inclusion of studies and the amount of data available for analysis and hence statistical power. The available research suggests that case management is generally effective.

Declaration of interest None.

The search for evidence of the effectiveness of interventions in medicine and the health sciences has a history dating back to the middle ages (Lancaster, 1994). Following the work of Fisher (1935) in theorising the design of the randomised controlled trial (RCT) as the most important method for establishing causality of interventions, RCTs have become widely used in medicine and increasingly so in the behavioural sciences and social programme evaluation fields. However, RCTs produce equivocal results due to lack of statistical power, sampling error, measurement error, different statistical techniques, heterogeneity of interventions and confounding variables (Cook & Campbell, 1979; Rosenthal, 1984; Hedges & Olkin, 1985; Cohen, 1988; Cook *et al*, 1992). In order to overcome the limitations of individual studies and the difficulties in qualitatively combining the results from many studies, the use of systematic reviews that include the statistical technique of 'meta-analysis' has become widespread. Meta-analytical techniques were developed during the 1960s and 1970s but became well known with the publication of the systematic review of the effectiveness of psychotherapy by Glass, McGaw & Smith (1981).

SYSTEMATIC REVIEWS

Systematic reviews attempt to combine research findings as objectively as possible and are used across many areas of medicine (e.g. Andrews & Harvey, 1981; Shadish, 1992; Peipert & Bracken, 1997; Goodman, 1998; Booth, 1999). Briefly, key steps involve operationalising the variables to be examined, specifying study inclusion criteria, searching for studies that meet these criteria, calculating 'effect sizes' (i.e. the size of the difference in outcomes for intervention and control groups) for the domains being examined, combining effect

sizes across studies and examining the results for possible bias (Glass *et al*, 1981; Rosenthal, 1984; Hedges & Olkin, 1985; Cooper, 1989; Cook *et al*, 1992; Cooper & Hedges, 1994).

In the early 1990s the Cochrane collaboration was established in the UK to facilitate systematic reviews of the efficacy of health interventions. The Cochrane database of systematic reviews has become the 'gold standard' of evidence for many in the health field, and these reviews have been described as 'providing the highest levels of evidence ever achieved on the efficacy of preventive, therapeutic and rehabilitative regimens' (Sackett & Rosenberg, 1995, p. 623).

THE EFFECTIVENESS OF MENTAL HEALTH CASE MANAGEMENT

Case management is the 'coordination, integration and allocation of individualised care within limited resources' (Thorncroft, 1991) and it has been widely introduced in mental health services (Onyett, 1992). Case management includes the functions of: psychosocial needs assessment; individual care planning; referral and linking to appropriate services or supports; ongoing monitoring of the care plan; advocacy; monitoring the client's mental state; compliance with medication and possible side-effects; the establishment and maintenance of a therapeutic relationship; and supportive counselling (Stein & Diamond, 1985; Dincin, 1990; Chamberlain & Rapp, 1991; Draine, 1997; Drake *et al*, 1998).

A particularly controversial and much debated Cochrane review was a meta-analysis of the effectiveness of case management in mental health services, conducted by Marshall and colleagues (Marshall *et al*, 1996). The conclusions of the 1996 Cochrane review were scathing:

'These findings have important implications for the UK government. Here the statutory introduction of case management has been triply unfortunate. First health and social services, patients, and carers have been saddled with an unproven intervention whose main effect is likely to be a considerable increase in the demand for hospital beds. Second, the obligatory nature of the intervention is likely to impede attempts to introduce superior alternatives, or to further evaluate its effectiveness. Third, the intervention has become a political policy and hence has acquired a degree of support from vested interests whose motives for continuing to support

the intervention are political rather than scientific.' (Marshall *et al*, 1996, p. 7)

Although the language was softened somewhat in an updated 1998 review, the conclusions were similar:

'In summary, therefore, case management is an intervention of questionable value, to the extent that it is doubtful whether it should be offered by community psychiatric services. It is hard to see how policy makers who subscribe to an evidence-based approach can justify retaining case management as 'the cornerstone' of community mental health care.' (Marshall *et al*, 1998, p. 1)

There have been some criticisms of Marshall's clinical case management review. One argument was that its conclusions for the ineffectiveness of clinical case management relied too greatly on increased admission being categorised as a negative outcome, with the comment that the impact on total length of hospitalisation had not been reported (Parker, 1997).

It was also argued that the case management programmes studied may not have employed skilled or competent staff, and that there was too little information about the operation of the models in practice to reach conclusions about case management generally (Parker, 1997). Although this may be true, the evidence from those studies included could be expected to provide some indication of the programme's effectiveness overall.

Not all commentaries citing the Marshall review have been dismissive. Citing that review and one other study as evidence, Tyrer was able to conclude that clinical case management was 'a profligate model which is expensive, increases bed use and separates professionals' (Tyrer, 1998, p. 2). Parker argued that a 'broader review' of elements of community psychiatry, such as case management, was necessary. In an attempt to widen the examination of the effectiveness of case management, our group conducted another systematic review that reached conclusions on a greater range of outcomes. The findings of this review (Ziguras & Stuart, 2000) compared with the Cochrane review are presented below, and some of the methodological differences between the two reviews are discussed.

COMPARISON OF SYSTEMATIC REVIEWS

Marshall and colleagues analysed the effectiveness of assertive community treatment

(ACT) and other models of case management separately (we shall refer to these collectively as 'clinical case management' because they share many common features) (Marshall *et al*, 1998; Marshall & Lockwood 1998). The meta-analysis of the effectiveness of ACT (Marshall & Lockwood, 1998) found that ACT clients were more likely than clients of standard care to remain in contact with services, less likely to be admitted, spent less time in hospital and had better outcomes on accommodation status, employment and satisfaction with services.

For clinical case management, the authors were able to reach conclusions for only two domains of outcome, using data from 11 RCTs, and found that case management increased the proportion of clients admitted (although this is also reported as increasing total admissions) but decreased drop-out rates from mental health services (Marshall *et al*, 1998).

Our review (Ziguras & Stuart, 2000) came to conclusions about 11 domains of outcome from 35 studies. We found that ACT and clinical case management were both effective in reducing symptoms of illness, improving social functioning, increasing client and family satisfaction with services and reducing client drop-out from services. Both models appeared equally effective in these areas. In contrast, ACT reduced the number of admissions and proportion of clients hospitalised, whereas clinical case management increased both. Both models reduced hospital days used but ACT was significantly more effective (Ziguras & Stuart, 2000).

Although these findings about clinical case management initially appear to contradict Marshall's results, it should be noted that the results were the same for the two domains common to both analyses. That is, both studies found that clinical case management was effective in preventing clients from dropping out from services, and also led to a greater proportion of clients being hospitalised. However, we found a range of other domains in which clinical case management was more effective than standard care and concluded that it led to small-to-moderate improvements in care provided to people with a serious mental illness.

Our overall conclusions about the effectiveness of case management were substantially different: Marshall's were primarily negative, as cited above; ours were much more positive. Given that we reviewed the same body of research, how could we come

to such different conclusions? There were three key differences between Marshall's methods and our own, which are discussed below.

Marshall *et al* included only studies with randomised control groups, whereas we included both RCTs and studies with quasi-experimental designs (i.e. control groups matched on certain characteristics but not randomly allocated). In our own analysis (Ziguras & Stuart, 2000), studies were weighted by study quality using a scale similar to that used by Glass *et al* (1981). These categories were: random assignment to conditions, with attrition less than 20% (highest rating); random assignment, with attrition greater than 20% or differing between groups; well-designed matching studies or analysis for covariance; and weak or non-existent matching procedures (lowest). The impact of including matched trials versus RCTs (83% of included studies used a RCT design) on the effect sizes obtained was examined using a sensitivity analysis and the results showed that the inclusion of non-randomised trials had not biased the overall results.

A second difference between these reviews was that Marshall *et al* excluded domains of outcome that had not been previously reported in a peer-reviewed journal. On the other hand, we included all measures, arguing that this would increase the power of the analysis. We also believed that the inclusion of measures with lower reliability (assuming that non-reported measures had lower reliability) would lead to greater variance in the outcome scores, thus lowering the effect size found from intervention, a point also made by Cohen (1988). We examined the impact of this strategy and found that the mean effect size for the non-reported measures was lower than that for the reported measures. This meant that their inclusion had led probably to the results underestimating the effectiveness of case management, but none the less provided important evidence against the proposition that the case management is ineffective.

The third difference in method was that Marshall *et al* excluded studies that included data with skewed, non-normal distributions that had not been transformed before being analysed using standard parametric statistics (such as *t*-tests or *F*-tests), whereas these were included in our analysis.

DISCUSSION

There were three key methodological differences between the two reviews – inclusion of quasi-experimental studies, inclusion of domains using non-published scales and parametric analysis of skewed data. The effect of these differences on the results is discussed below.

Randomised versus non-randomised trials

Although RCTs are acknowledged to be a superior form of evidence, there are many reasons why they are not carried out in practice. Ideally, one would only include RCTs in a systematic review. If there are insufficient RCTs to provide adequate data, then including non-randomised studies can be justified, especially as statistical methods can be used to control for known confounding variables. Typologies of levels of evidence used by many research institutions acknowledge this point. Because Marshall *et al* were only able to reach conclusions for two domains of outcome, we decided to examine the other available evidence, at the same time being aware of the possible bias that this may introduce.

We note that a similar approach was taken by the Cochrane reviewers in relation to randomisation of staff to intervention and control groups. Because staff may be more or less motivated, experienced or competent, it is possible that the method by which staff are allocated to programmes, such as self-selection, may bias the results (e.g. because staff who choose to participate in a new case management programme may be more motivated in their work). None of the studies included in either review randomly allocated staff to treatment and control conditions. However, if we were to argue that only studies using random allocation to protect against bias should be included, there would not have been any studies available for review at all. This is not a rationale *per se* for including quasi-experimental studies, but it does illustrate that the Cochrane reviewers have also had to balance inclusion criteria with the research available. A related issue regarding hierarchically structured datasets is discussed below.

Some empirical evidence shows that non-randomised trials tend to over-estimate the effect of interventions (Kunz & Oxman, 1998), suggesting that a sensitivity analysis

(comparing the results obtained by including and then excluding non-randomised studies) should be conducted to examine the effect of including studies of differing quality. A sensitivity analysis in our review showed the same trend for lower quality studies to overestimate the effect size, but also showed that, overall, their inclusion did not alter the results of the meta-analysis.

Unpublished versus published scales

Marshall *et al* excluded data from non-published instruments; this is a reasonable strategy, because measures that had not been subject to peer review may have low reliability or doubtful validity, although publication itself does not guarantee instrument quality. However, this strategy had the disadvantage of further restricting the number of studies included. Marshall *et al* (2000) showed that trials were more likely to report that an intervention was effective when unpublished scales were used, compared with the use of published scales, and that this effect was more pronounced in studies of non-pharmacological intervention. They speculated that this may be due to researchers adapting scales to find significant results and that this was more feasible when the scale was not already published. In our review, we speculated that unpublished instruments have lower reliability rates and would therefore underestimate effect sizes; in fact, we found that this was the case with the studies included. However, a sensitivity analysis showed that the inclusion of such scales did not bias the overall findings. On the face of it, these findings contradict those of Marshall *et al* (2000) and suggest that further investigation of this question is required.

Skewed data

The third major methodological difference between the two reviews was in the treatment of data with skewed distributions. Some statistics texts recommend the transformation of skewed data before analysis, but a better approach is to use statistical methods for which the assumptions are not violated by the data. Although skewed data may lead to incorrect inferences in some circumstances, the results of simulation studies show that where sample sizes are moderately large (above 30), skewed data can be analysed using parametric

statistics, without significant loss of accuracy (Sawilowsky & Blair, 1992). In our review, the median sample sizes being considered varied from a minimum of 32 for 'family satisfaction' to a maximum of 121 for 'proportion of group hospitalised' (Ziguras & Stuart, 2000). The domains of outcome most affected by skewed distributions are hospital admissions and days spent in hospital. However, given the reasonably large sample sizes and the fact that intervention and control groups will be skewed in the same direction and roughly to the same extent, analysis using parametric methods would not be expected to give misleading results.

The three exclusion criteria used by Marshall *et al* – excluding matched studies, excluding domains with non-reported measures and excluding parametric analyses of skewed data – are all defensible on theoretical grounds. However, their combined effect was to limit the number of studies to such an extent that few data remained to be analysed. The corresponding strategies used by ourselves – the inclusion of matched-group studies, more measures of outcome and studies that used parametric analyses of skewed data – could be regarded as somewhat risky. However, the first two factors were shown not to have biased the results in favour of case management (in fact, the opposite was true for the issue of measurement reliability), and the sample sizes suggest that the impact of the third was minimal.

Furthermore, the agreement between the two methods in the findings for drop-out rates and proportion of clients hospitalised suggests that the methods were comparable in accuracy, but our approach enabled the examination of a broader range of outcomes. We believe that the available evidence supports the contention that case management is effective in improving mental health services.

There are two other key issues in assessing the evidence for the effectiveness of case management that are not addressed by either of the systematic reviews referred to above: hierarchical data structures and heterogeneity of case management models.

Hierarchical data structures

In all of the case management evaluations examined, intervention of a case management programme is carried out by a relatively small group of staff. If there is a

staff effect (because of characteristics such as skill, experience, motivation, commitment), then clients sharing the same case manager are more likely to share similar outcomes. In this situation, the assumption that client outcomes are independent is violated. Differences between intervention and control groups may be due to differences in the characteristics of staff in the two programmes, which may be completely independent of the programme model. Furthermore, staff effects may be amplified when there is a reasonably large number of clients per staff member.

This problem can be dealt with by randomly allocating staff as well as clients to programmes or by using statistical methods to control for differences between staff (Kreft & de Leeuw, 1998). Such techniques, known as hierarchical or multi-level models, have been used in other areas of social programme evaluation but, to our knowledge, have not been used in case management research. The importance of such techniques is illustrated by a well-known study of schoolchildren carried out in the 1970s using traditional regression methods (Bennett, 1976), which found that children exposed to 'formal' styles of teaching reading showed better progress than those who were not. A subsequent re-analysis using hierarchical techniques (Aitkin *et al*, 1981) demonstrated that the significant differences disappeared.

Models of case management

A second issue concerns the delineation of models of case management. Although there are many possible types and dimensions of case management, a distinction is often made in mental health between 'assertive community treatment' (ACT) models and 'generic' or other models (Dincin, 1990; Chamberlain & Rapp, 1991; Draine, 1997), and this was the approach used by both our reviews.

However, many models can be conceptualised. Solomon (1992) distinguished four types of case management: assertive community treatment, strength case management, rehabilitation case management and generalist case management. Mueser *et al* (1998) described six models: broker case management, clinical case management, strength case management, rehabilitation case management, assertive community treatment and intensive case management. They pointed out that the models could be grouped into three broad

types but acknowledged that 'the differences between models within each of these broad types of community care can be difficult to establish'. Thus, there appears to be little consensus about the best way to specify models of case management.

Rather than describing discrete categories, Thornicroft (1991) proposed 12 dimensions that could be used to distinguish case management programmes. It seems likely that individual implementations of the same model (such as ACT) may vary on some of these dimensions (such as case-load size, years of experience of staff), which could affect outcomes for clients. A more productive method would be to measure each case management programme on these dimensions. This would allow us to delineate categories based on shared empirical features and, more importantly, to investigate the effects of these dimensions on effectiveness using meta-analytical linear regression techniques (Sharp, 1998). The potential importance of such analysis is illustrated by the finding of a limited meta-analysis by Gorey *et al* (1998) that the only factor influencing case management effectiveness was the size of case-loads; 80% of the studies included in that review had case-loads of less than 20.

We have discussed two different systematic reviews of the evidence regarding the effectiveness of case management. Despite methodological differences, both reached the same conclusion on the same domains of outcome. However, the reviews demonstrate that systematic reviews may involve trade-offs, in this case between the application of strict criteria for the inclusion of studies and the amount of data available for analysis and hence statistical power. We believe that meta-analysis is an important advance on simple qualitative reviews of research, but clearly it does not resolve all questions about evidence. Perhaps the most eloquent expression of these considerations is that offered in the guidelines for Cochrane reviewers:

'The guidelines provided here are intended to help reviewers to be systematic and explicit (not mechanistic!) (sic) about the questions they pose and how they derive answers to those questions. These guidelines are not a substitute for good judgement.' (Mulrow & Oxman, 1997, p. 8).

REFERENCES

Aitkin, M., Anderson, D. & Hinde, J. (1981) Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society, Part A*, **144**, 148–161.

Andrews, G. & Harvey, R. (1981) Does psychotherapy benefit neurotic patients? A reanalysis of the Smith, Glass and Miller data. *Archives of General Psychiatry*, **38**, 1203–1208.

Bennett, N. (1976) *Teaching Styles and Pupil Progress*. London: Open Books.

Booth, A. (1999) Mapping the evidence base of pathology. *Journal of Pathology*, **188**, 344–350.

Chamberlain, R. & Rapp, C. (1991) A decade of case management: a methodological review of outcome research. *Community Mental Health Journal*, **27**, 171–188.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cook, T. D. & Campbell, D. T. (1979) *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.

—, **Cooper, H., Cordray, D., et al (1992)** *Meta-analysis for Explanation: a Casebook*. New York: Russell Sage Foundation.

Cooper, H. (1989) *Integrating Research: a Guide for Literature Reviews*. Newbury Park, CA: Sage.

— & **Hedges, L. (1994)** *The Handbook of Research Synthesis*. New York: Russell Sage Foundation.

Dincin, J. (1990) Assertive case management. *Psychiatric Quarterly*, **61**, 49–55.

Draine, J. (1997) A critical review of randomized field trials of case management for individuals with serious and persistent mental illness. *Research on Social Work Practice*, **7**, 32–52.

Drake, R., McHugo, G., Clark, R., et al (1998) Assertive Community Treatment for patients with co-occurring severe mental illness and substance use disorder: a clinical trial. *American Journal of Orthopsychiatry*, **68**, 201–205.

Fisher, R. A. (1935) *The Design of Experiments*. Edinburgh: Oliver & Boyd.

Glass, G., McGaw, B. & Smith, M. (1981) *Meta-analysis in Social Research*. Beverly Hills: Sage.

Goodman, N. W. (1998) Anaesthesia and evidence-based medicine. *Anaesthesia*, **53**, 353–368.

Gorey, K. M., Leslie, D. R., Morris, T., et al (1998) Effectiveness of case management with severely and persistently mentally ill people. *Community Mental Health Journal*, **34**, 241–250.

Hedges, L. & Olkin, I. (1985) *Statistical Methods for Meta-analysis*. New York: Academic Press.

Kreft, I. & de Leeuw, J. (1998) *Introducing Multilevel Modeling*. London: Sage.

Kunz, R. & Oxman, A. D. (1998) The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *British Medical Journal*, **317**, 1185–1190.

Lancaster, H. O. (1994) *Quantitative Methods in Biological and Medical Sciences: a Historical Essay*. New York: Springer-Verlag.

Marshall, M., Gray, A., Lockwood, A., et al (1996) Case management for people with severe mental disorders. In *Schizophrenia Module of the Cochrane Database of Systematic Reviews* (eds C. Adams, C. Anderson & J. De Jesus Mari). London: The Cochrane Library, BMJ Publishing Group.

—, —, —, **et al (1998)** Case management for people with severe mental disorders. In *Schizophrenia Module of the Cochrane Database of Systematic Reviews* (eds C. Adams, C. Anderson & J. De Jesus Mari). London: The Cochrane Library, BMJ Publishing Group.

— & Lockwood, A. (1998) Assertive community treatment for people with severe mental disorders. In *Schizophrenia Module of the Cochrane Database of Systematic Reviews* (eds C. Adams, C. Anderson & J. De Jesus Mari). London: The Cochrane Library, BMJ Publishing Group.

—, —, Bradley, C., et al (2000) Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *British Journal of Psychiatry*, **176**, 249–252.

Mueser, K. T., Bond, G. R., Drake, R. E., et al (1998) Models of community care for severe mental illness: a review of research on case management. *Schizophrenia Bulletin*, **24**, 37–74.

Mulrow, C. D. & Oxman, A. D. (1997) *Cochrane Collaboration Handbook*. Oxford: Update Software.

Onyett, S. (1992) *Case Management in Mental Health*. London: Chapman & Hall.

Parker, G. (1997) Case management: an evidence-based review fails to make its case. *Current Opinions in Psychiatry*, **10**, 261–263.

Peipert, J. F. & Bracken, M. B. (1997) Systematic reviews of medical evidence: the use of meta-analysis in obstetrics and gynecology. *Obstetrics and Gynecology*, **89**, 628–633.

Rosenthal, R. (1984) *Meta-analytic Procedures for Social Research*. Newbury Park: Sage.

Sackett, D. L. & Rosenberg, W. M. C. (1995) The need for evidence-based medicine. *Journal of the Royal Society of Medicine*, **88**, 620–624.

Sawilowsky, S. S. & Blair, R. C. (1992) A more realistic look at the robustness of Type II error properties of the t-test to departures from population normality. *Psychological Bulletin*, **111**, 352–360.

Shadish, W. (1992) Do family and marital therapies change what people do? A meta-analysis of behavioral outcomes. In *Meta-analysis for Explanation: a Casebook* (eds T. Cook, H. Cooper, D. Cordray, et al), pp. 257–308. New York: Russell Sage Foundation.

Sharp, S. (1998) sbe23 Meta-analysis regression. *Stata Technical Bulletin*, **42**, 148–155.

Solomon, P. (1992) The efficacy of case management services for severely mentally disabled clients. *Community Mental Health Journal*, **28**, 163–180.

CLINICAL IMPLICATIONS

- Case management is generally effective in improving service outcomes for people with psychiatric disabilities.
- Assertive community treatment is a more effective intervention than clinical case management for people with numerous previous admissions.
- The effectiveness of clinical case management is questionable for high staff caseloads (above 30).

LIMITATIONS

- Systematic reviews face a trade-off between rigorous inclusion criteria and statistical power.
- Reviews are limited by a lack of consensus about models of case management.
- Compared to the categorical definitions of case management adopted in most reviews, a more informative approach may be to examine the impact of dimensions of case management on outcomes.

S. J. ZIGURAS, PhD, Brotherhood of St Laurence, Melbourne, Australia; G.W. STUART, PhD, Department of Psychiatry, University of Melbourne, Australia; A. C. JACKSON, PhD, School of Social Work, University of Melbourne, Australia

Correspondence: Dr Stephen Ziguras, Brotherhood of St Laurence, 67 Brunswick St, Fitzroy 3065, Victoria, Australia. Tel: 61-3-9483 1316; E-mail: sziguras@bsl.org.au

(First received 8 February 2001, final revision 13 August 2001, accepted 27 September 2001)

Stein, L. & Diamond, J. (1985) A program for difficult-to-treat patients. *New Directions for Mental Health Services*, **26**, 29–31.

Thornicroft, G. (1991) The concept of case management for long-term mental illness. *International Review of Psychiatry*, **3**, 125–132.

Tyrer, P. (1998) Cost-effective or profligate community psychiatry? *British Journal of Psychiatry*, **172**, 1–3.

Ziguras, S. & Stuart, G.W. (2000) A meta-analysis of the effectiveness of mental health case management over 20 years. *Psychiatric Services*, **51**, 1410–1421.