

Are randomised controlled trials the only gold that glitters?

MIKE SLADE and STEFAN PRIEBE

The intention of evidence-based mental health care is that every clinical decision should be underpinned by research evidence. It is therefore clearly important to agree what constitutes evidence. A hierarchy of evidence is widely used, with systematic reviews and meta-analyses being the strongest, followed by randomised controlled trials (RCTs) with definitive results, RCTs with non-definitive results, cohort studies, case-control studies, cross-sectional surveys and case reports. Thus, good quality evidence is equated with RCTs, which can be grouped using meta-analyses and systematic reviews. Can RCTs provide all the necessary evidence? Three conceptual issues will be considered: group-level research designs, generalisation and bias in the evidence base.

GROUP-LEVEL DESIGNS

Randomised controlled intervention studies involve grouping subjects, typically by diagnosis. This design is appropriate if all people with a given mental disorder are fundamentally similar, because individual differences can be addressed by controlling for other variables that are seen as relevant. This has allowed the development of a substantial evidence base regarding 'best practice' for a range of disorders, such as a deterministic flow chart describing pharmacological treatment strategies for schizophrenia (Taylor, 1996).

The danger of an evidence base using a group-based research design is that it implies that the group label (e.g. diagnosis) is a sufficient characterisation on which to make treatment decisions. Treatment protocols derived from RCT evidence have the potential to focus clinicians on diagnosis-based interventions rather than on the development of individualised formulations and intervention strategies. A practical result is that, in general, people who meet the diagnostic criteria for schizophrenia are always prescribed antipsychotic medication,

even though the evidence indicates that it will be ineffective (and, owing to side-effects, on balance will be harmful) for some patients. An alternative view of people with a particular mental disorder is that they are fundamentally different from each other, with a few similarities where they all match the operational criteria for the disorder. Such a view implies the need for individual-level research designs.

GENERALISATION

Current mental health research is dominated by inferential statistics, which involves the assumption that a result can be generalised – that it is representative of something. The use of inferential statistics only makes sense if the population from which the sample was taken can be characterised and if one can identify to which other samples, settings and times the result can be generalised to. This may not be possible. For example, recent studies investigated the effectiveness of two patterns of clinical services in London (Thornicroft *et al*, 1998) and of deinstitutionalisation in Berlin (Hoffmann *et al*, 2000). To which patients do the findings of these studies generalise? What criteria can be identified for establishing what the results are representative for? To control for context, the unit of analysis in mental health service research may have to be the service, and not the patient in a service. Involving the necessary number of services in an RCT will be impossible for many research questions. RCTs certainly have a role in the development of services, such as for evaluating which service structures lead to the provision of which treatments, but other types of evidence are also needed.

THE EVIDENCE BASE

If care is to be provided on the basis of evidence, then it follows that equal

opportunity should be available for all types of relevant research evidence to be gathered and considered. This requirement is not met for at least four reasons.

First, the methods of natural science may not be as applicable to the study of mental health as to physical health. For example, the assessment of height or electrolyte levels is relatively straightforward because they can be measured directly. The assessment of psychological characteristics, such as severity of depression or conviction in delusions, necessarily requires proxy measures. For these characteristics there cannot be a 'true' measure because they are not observable. It is tempting, therefore, to ignore them. However, as Robert McNamara (former US Secretary of State) is reported to have said: "The challenge is to make the important measurable, not the measurable important". The compelling reason to include consideration of characteristics such as 'quality of life', 'beliefs', 'motivation' and 'self-esteem' is that these are precisely what go wrong in mental disorder. Therefore, we contend that the methods of social science are as applicable as the methods of natural science.

Second, RCTs are particularly appropriate for interventions for which it can be shown that there is treatment integrity – the intervention offered is no more and no less than what is intended and the patient receives the treatment. Although haloperidol, for instance, is well defined by its chemical structure, the way in which psychological and social interventions are provided may (appropriately) vary between patients and between therapists. For service programmes and systems, the situation is even more complex. Treatment integrity is relatively easy to ensure for pharmacotherapy, relatively difficult to ensure for individual psychotherapeutic and psychosocial treatments and practically impossible where the intervention is a complex package of care or a care system.

This is illustrated by the findings from the Schizophrenia Patient Outcomes Research Team (PORT) review of outcome studies in schizophrenia (Lehman & Steinwachs, 1998), which made 30 recommendations, of which 25 were positive: 17 concerning pharmacotherapy, two concerning electroconvulsive therapy, one concerning family therapy, one concerning individual and group therapies, and four concerning services (vocational rehabilitation and assertive treatment). The use of RCTs as

the means by which evidence is gathered leads to a lot of evidence regarding pharmacotherapy, less concerning other types of intervention and little, if any, undebated positive evidence about service research. To illustrate the point, the only PORT recommendation regarding service configuration is assertive community treatment, which is a subject of active disagreement among researchers in the UK (Burns *et al*, 1999; Thornicroft *et al*, 1999). The use of RCTs therefore has not produced widely accepted evidence for mental health services. It may be that bigger and better trials – ‘mega-trials’ – will produce the desired generalisable evidence (Gilbody & Song, 2000). It may be also that conceptual shortcomings of the RCT design will mean that the lack of consensus is not solely due to under-powered trials.

A third reason for the disparity in the available evidence is bias. Researchers who undertake any research will have particular values and beliefs. In mental health research, for example, this will lead them to investigate one intervention rather than another or to present findings confirming rather than refuting their beliefs. Appraisal bias is recognised within social science research, and attempts are made to separate the roles of participant and observer. This bias is much less recognised in mental health research. There may be availability bias – a skew in the number of studies of sufficient quality for inclusion in a review. As an example, the above-mentioned PORT review (whose first author is a psychiatrist) produced 19 positive recommendations related to physical treatments and only four to psychological, social and vocational approaches, underlining the role of pharmacotherapy in schizophrenia. Another review carried out by psychologists was much more optimistic regarding the role of psychological and social interventions (Roth & Fonagy, 1996). When natural science methods are used for research into mental health, bias in the research process is unavoidable but can be reduced using the methods of the social sciences.

A fourth reason for the disparity is economic considerations. There is aggressive marketing of pharmacotherapy and of related research by pharmaceutical companies, including the use of promotional

MIKE SLADE, CPsychol, Institute of Psychiatry, Denmark Hill, London; STEFAN PRIEBE, MD, St Bartholomew's & The London School of Medicine, East Ham Memorial Hospital, London

Correspondence: Mike Slade, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK

(First received 18 September 2000, final revision 17 April 2001, accepted 23 April 2001)

material citing data that may not have been peer-reviewed (e.g. ‘data on file’) (Gilbody & Song, 2000). Furthermore, the available data may be presented selectively, such as one trial of olanzapine that has been published in various forms in 83 separate publications (Duggan *et al*, 1999). This compares with very little active marketing for psychological or social interventions. Economic factors influence the provision and availability of evidence.

CONCLUSION

Randomised controlled trials in medicine have been used for evaluating well-defined and standardised treatments. The importing of this approach into mental health service research strengthens the position of pharmacotherapy (which tends to be a standardised and well-defined intervention) compared with psychological and social interventions, and underlines the link between psychiatry and other specialities in medicine. Regarding RCTs as the gold standard in mental health care research results in evidence-based recommendations that are skewed, both in the available evidence and the weight assigned to evidence.

Mental health research needs to span both the natural and social sciences. Evidence based on RCTs has an important place, but to adopt concepts from only one body of knowledge is to neglect the contribution that other, well-established methodologies can make (Priebe & Slade, 2001). RCTs can give better evidence about some contentious research questions, but it is an illusion that the development of increasingly rigorous and sophisticated RCTs will ultimately provide a complete evidence base. If mental health researchers are to ask all possible questions, to evaluate the evidence in a disinterested fashion, and to present the results in a balanced and

non-partisan way, then there needs to be more use of established methodologies from other fields.

ACKNOWLEDGEMENTS

We are grateful to Derek Bolton, Gene Feder, Elizabeth Kuipers and James Tighe for their comments.

DECLARATION OF INTEREST

None.

REFERENCES

- Burns, T., Fahy, T., Thompson, S., *et al* (1999) Intensive case management for severe psychotic illness (letter). *Lancet*, **354**, 1385–1386.
- Duggan, L., Fenton, M., Dardennes, R. M., *et al* (1999) Olanzapine for schizophrenia. *Cochrane Library*, Issue 3. Oxford: Update Software.
- Gilbody, S. & Song, F. (2000) Publication bias and the integrity of psychiatry research. *Psychological Medicine*, **30**, 253–258.
- Hoffmann, K., Isermann, M., Kaiser, W., *et al* (2000) Quality of life in the course of deinstitutionalisation – Part IV of the Berlin deinstitutionalisation study (in German). *Psychiatrische Praxis*, **27**, 183–188.
- Lehman, A. & Steinwachs, D. (1998) Translating research into practice: the Schizophrenia Patient Outcomes Research Team (PORT) treatment recommendations. *Schizophrenia Bulletin*, **24**, 1–10.
- Priebe, S. & Slade, M. (2001) *Evidence in Mental Health Care*. London: Routledge (in press).
- Roth, A. & Fonagy, P. (1996) *What Works for Whom?* London: Guilford Press.
- Taylor, D. (1996) Through the minefield: how to decide which drug to use and when. *Schizophrenia Monitor*, **6**, 1–5.
- Thornicroft, G., Strathdee, G., Phelan, M., *et al* (1998) Rationale and design. PRISM Psychosis Study I. *British Journal of Psychiatry*, **173**, 363–370.
- , Becker, T., Holloway, F., *et al* (1999) Community mental health teams: evidence or belief? *British Journal of Psychiatry*, **175**, 508–513.