PRiSM Psychosis Study

Design limitations, questionable conclusions

MAX MARSHALL, GARY BOND, LEONARD I. STEIN, GEOFF SHEPHERD, JOHN McGREW, JOHN HOULT, ALAN ROSEN, PETER HUXLEY, RONALD J. DIAMOND, RICHARD WARNER, MELODY OLSEN, ERIC LATIMER, PAULA GOERING, TOM K. J. CRAIG, NEIL MEISLER and MARY ANN TEST

The purpose of this paper is to comment on the reports of the PRiSM Psychosis Study (*BJP*, 173, 363–427; November 1998). In particular, we have concerns about the study's rationale, methods and conclusions.

In essence, the PRiSM papers make three main points. Point one suggests that the PRiSM study is the first to examine how far specific community care interventions proven 'efficacious' in experimental studies are 'effective' in "ordinary clinical circumstances" (Thornicroft *et al*, 1998*a*, p. 363). Point two claims that the strengths of the PRiSM study make it unique and of particular relevance (Thornicroft *et al*, 1998*b*, p. 423). Point three is the claim that the PRiSM study proves that specific community care interventions are effective, but that this effectiveness is "diluted in ordinary clinical practice" (Thornicroft *et al*, 1998*b*, p. 425). As we will show below, we believe that the PRiSM researchers are wrong on all three points.

IS THE PRiSM STUDY THE FIRST STUDY OF THE 'EFFECTIVENESS' OF SPECIFIC COMMUNITY CARE INTERVENTIONS?

We take issue with the authors' suggestion that the community care literature lacks effectiveness studies. The authors assert that "Efficacy means how far a specific intervention achieves its intentions under ideal, experimental conditions...", whereas "Effectiveness . . . has been defined as . . . 'the extent to which a specific intervention, when used under ordinary clinical circumstances, does what it is intended to do' " (Thornicroft *et al*, 1998*a*, p. 363). Starting with the accompanying editorial (Tyrer, 1998) and permeating the ten papers is the repeated assertion that the existing body of assertive community treatment (ACT) research consists of efficacy research. If the authors mean to suggest that the existing ACT literature includes a great many experimental studies, then yes, it is true that there are such studies, alongside a number of very fine non-experimental studies (Mueser *et al*, 1998; Marshall & Lockwood, 1998). However, the claim that the existing literature on ACT reports work done "under ideal conditions", whereas the PRiSM study was done "under ordinary clinical circumstances", is unwarranted, as a reading of the vast ACT literature will show immediately. For one thing, ACT programmes have been widely disseminated into ordinary clinical practice in the USA, Australia and elsewhere (Teesson & Hambridge, 1992; Hambridge & Rosen, 1994; Deci *et al*, 1995; Meisler, 1997) and many of these programmes have been evaluated (Bond *et al*, 1995; Mowbray *et al*, 1998). It is also incorrect to say that the experimental studies of ACT have been done "under ideal conditions" and *not* "under ordinary clinical circumstances". A careful reading of the two-dozen published randomised controlled trials (RCTs) of ACT, starting with Stein & Test's (1980) original work, shows that, without question, these studies have been conducted squarely within the real world. Existing ACT studies were actually used as the prime example of 'effectiveness research' in a recent review of that subject (Wells, 1999). Hence, contrary to the claims of its authors, the PRiSM study is not "the first to address . . . the effectiveness rather than the efficacy of community mental health services" (Thornicroft *et al*, 1998*a*, p. 363).

Not only is the PRiSM study not the *first* study of effectiveness, it is doubtful whether it is a study of effectiveness at all. In their introduction (Thornicroft *et al*, 1998*a*, p. 363, quoted above) the PRiSM authors define an effectiveness study as a study of the implementation of "a *specific* intervention . . . under ordinary clinical circumstances" (our italics). The term 'specific' is vital here because it refers to the fact that the intervention in the 'effectiveness' study is the same as the intervention in the 'efficacy' studies. It follows from this definition that an effectiveness study must both adhere to a single, defined model of care that has already been evaluated in an efficacy study, and be compared against ordinary clinical care. There are three reasons why the PRiSM study does not meet any of these criteria.

Reason one: The intervention in the PRiSM study is not a single model of care, but a mixture of two models. In the PRiSM study, care was provided by a psychiatric acute care and emergency (PACE) team (described as a 9 am–5 pm, seven days per week crisis service) and a psychiatric assertive continuing care (PACT) team, which cared for people with long-term mental illness (Becker *et al*, 1998). The results from the Nunhead sector (one of the two geographical sectors studied) do not clearly state how many follow-up patients received 'PACE', how many received 'PACT', or how many received neither or both. This intermingling of the two models makes it impossible to make definitive statements about either.

Reason two: Neither of the two models is clearly related to models of care that have been evaluated in efficacy studies. The descriptions of the models are incomplete, so it is difficult to make definitive judgements. However, rather than following any particular model of care, the PACE team appears to have been a local innovation, and appears to be an episode-based service. It does not seem to bear much relationship to the fully integrated service approach, with a single point of responsibility, that is found in a fully functioning ACT team (Test, 1992). At first glance the acronymic title of the PACT team implies that it adhered to the carefully specified Programme of Assertive Community Treatment (Stein & Test, 1980) – an implication made explicit in the accompanying editorial (Tyrer, 1998) – but the description of the intervention provided does not appear to bear this out. For example, although the authors provide some information on staffing in the intensive sector, they do not provide certain basic information such as the actual size of the PACT team and its case load. Yet the authors must be aware that a client to staff ratio of 10:1 is accepted as optimal for an ACT team by experts

†See pp. 504–513, this issue.

(McGrew & Bond, 1995) and is nearly universally adhered to in practice (Teague *et al*, 1998). Most researchers in this area have learnt the lessons of psychotherapy research and now use established fidelity scales to ensure that when the acronym PACT is used, the designated intervention is true to the carefully defined principles of assertive community treatment (McGrew *et al*, 1994). The PRiSM researchers do not do this, but none the less seem to imply (Thornicroft *et al*, 1998*b*) – as does the accompanying editorial (Tyrer, 1998) – that ACT offers few advantages in practice. It is not possible to conclude that ACT does not work, or that its effect is 'diluted', without first demonstrating that it was properly implemented in the first place.

Reason three: The intensive intervention is not much different from the ordinary clinical care

Even if we accept the PRiSM evaluation on its own terms, is it true to say that the authors provided a fair test of the effectiveness of their intervention? A fundamental scientific principle in programme evaluation is to ensure that the experimental and control conditions are, in fact, different. In evaluating a model programme the quality and intensity of services provided to the experimental group should exceed the quality and intensity of those provided to the control group. According to their own expert ratings, the authors found that services for the intensive sector started out *worse* than those for the standard sector in 1991, one year before data collection started. By 1996, the summed expert rating score was 30 (out of 40 possible quality points) for the intensive sector, barely higher than the 26 points given to the standard sector (Becker *et al*, 1998). Hence, it is not at all clear that the intensive sector actually received more intensive or more adequate services than the standard sector. The lack of psychometric information prohibits us from drawing further conclusions, except to reiterate that, by the authors' own admission "implementation of community services, in the intensive sector, was incomplete" (Becker *et al*, 1998, p. 374).

DOES THE PRiSM STUDY HAVE UNIQUE STRENGTHS?

The authors claim that their study has six strengths (Thornicroft *et al*, 1998*b*, p. 364) and imply that these strengths make the study unique. However, five of these strengths are widely found in other studies. For example, the authors imply that it is rare for ACT programmes to serve clients with concurrent substance use problems. On the contrary, if anything, such patients tend to cluster in ACT teams (Drake *et al*, 1998). Similarly, the PRiSM study is by no means unique in examining "a routine clinical service which is designed to endure, and not . . . one especially created only for the period of the research project" (Thornicroft *et al*, 1998*b*, p. 364). The past 20 years have produced scores of studies of enduring ACT programmes in the USA and Australia. Regarding the other 'strengths', the study is not the first to evaluate a programme in an inner-city environment (Witheridge *et al*, 1982; Morse *et al*, 1992), nor is it the first to use a wide range of outcome measures (Mueser *et al*, 1998) and it is not unusual to have a two-year follow-up period – there are ten RCTs of ACT or case management that have follow up periods of two years or greater (Marshall *et al*, 1998; Marshall & Lockwood, 1998). Taken individually, therefore, each of these attributes is found in many other ACT studies, although it is also true, in support of the authors' viewpoint, that a much smaller number of prior studies satisfy all of these criteria.

In fact, only one of the six 'strengths' claimed for the study – that it is 'epidemiologically based' – is relatively novel and might differentiate the study from the existing literature on ACT and case management (Thornicroft *et al*, 1998*b*, p. 423). Unfortunately this 'strength' actually may reflect a weakness in the sampling strategy. What the "epidemiologically based" approach appears to entail is redefining the target population for the study's intensive sector to be different from that of previous ACT research. Thus, one-third of the sample had no behavioural problems, one quarter had one or fewer symptoms (Wykes *et al*, 1998, p. 385), and some had never been admitted to hospital at all. Hence, it is doubtful whether a substantial proportion of the PRiSM study sample would ever have been admitted to most ACT programmes. Stein & Test (1980) proposed two decades ago that ACT should be reserved for clients most in need, and the vast majority of ACT programmes have admission criteria consistent with this principle. Providing intensive case management to clients who have no behavioural problems, few symptoms, and no history of hospital admission makes little sense. Consequently, the intensive sector programme design, as described, appears to represent a misapplication of the knowledge base developed over the past 20 years.

HAS THE PRiSM STUDY PROVED THAT THESE SPECIFIC INTERVENTIONS ARE EFFECTIVE, BUT THAT THIS EFFECTIVENESS IS "DILUTED IN ORDINARY CLINICAL PRACTICE"?

The PRiSM researchers claim that their study has shown that the "benefits shown in experimental studies can be replicated in routine settings", but also claim that "dilution does occur" (Thornicroft *et al*, 1998*b*, p. 426). Both claims are incorrect. The first claim is incorrect because what the trial actually shows is that there were no significant differences in favour of the intensive sector on any key outcome variable. The second claim is incorrect because a non-existent effect cannot be 'diluted'.

Unfortunately, whether positive or negative, the findings of the PRiSM study are largely irrelevant because of critical failures in safeguarding internal validity. In essence, the study is a medium-sized, single-centre non-randomised trial, with a low follow-up rate (sometimes as low as 40%). Moreover, because the data were not collected contemporaneously in the two sectors ('time 1' was 1992–1993 in the intensive sector and 1993–1994 in the standard sector), the study is vulnerable to a history confound (Cook & Campbell, 1979). Even more critically, the samples in the two sites were not well matched prior to treatment; instead, the clients in the intensive sector were more disabled than those in the standard sector (Wykes *et al*, 1998). This is a major flaw for any quasi-experimental design. We would sum up the entire study in the words that the authors use to describe their findings for quality of life: "Large baseline differences between sectors make interpretation difficult" (Taylor *et al*, 1998, p. 62).

In summary, our view is that the PRiSM study is not a unique study of effectiveness; rather it appears to be an inconsistently implemented service evaluation. We disagree strongly with the authors' contention that the PRiSM study tells us more than the findings of the 30 or more RCTs in this area (Mueser *et al*, 1998). To us this claim appears to be an inversion of the principles

of evidence-based medicine. Although the authors may have shown that *their* application of the ACT model was "diluted" when implemented in ordinary clinical service, they have not demonstrated a weakness of the model as such, only an inability to implement it properly. While it is not easy to implement ACT with high fidelity, other clinical teams have done much better (Teague *et al*, 1998).

None the less, we understand the reservations in the UK about applying findings from abroad to the UK service system. For this reason, despite the weight of international evidence in favour of ACT, there is still a place for a proper RCT in the UK. However, such a trial needs to be carried out with a clear vision of what is being implemented, and close monitoring of programme fidelity. The proposed reforms of the National Health Service and social services in the UK should increase the feasibility of such a trial by removing some of the artificial barriers to coordinated care (Department of Health, 1998). Until the findings of such a trial are available, we urge UK policy-makers to rely on the findings of the numerous properly conducted ACT trials, rather than on those of the flawed PRiSM study.

REFERENCES

Becker, T., Holloway, F., McCrone, P., et al (1998) Evolving service interventions in Nunhead and Norwood. PRiSM Psychosis Study 2. *British Journal of Psychiatry*, **173**, 371–375.

Bond, G. R., McGrew, J. H. & Fekete, D. M. (1995) Assertive outreach for frequent users of psychiatric hospitals: a meta-analysis. *Journal of Mental Health Administration*, **22**, 4–16.

Cook, T. D. & Campbell, D. T. (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago, IL: Rand–McNally.

Deci, P. A., Santos, A. R., Hiatt, D. W., et al (1995) Dissemination of assertive community treatment teams. *Psychiatric Services*, **46**, 676–678.

Department of Health (1998) *Modernising Mental Health Services*. London: Department of Health.

Drake, R. E., Mercer-McFadden, C., Mueser, K. T., et al (1998) Treatment of substance abuse in patients with severe mental illness: a review of recent research. *Schizophrenia Bulletin*, **24**, 589–608.

Hambridge, J. A. & Rosen, A. (1994) Impact of a mobile community intensive case management team in suburban Sydney. *Australia and New Zealand Journal of Psychiatry*, **28**, 438–445.

Marshall, M., Lockwood, A., Green, R., et al (1998) Case management for people with severe mental disorder. In *Schizophrenia Module of the Cochrane Database Systematic Reviews* (eds C. G. Adams, L. Duggan, J. de Jesus Mari, *et al*). Oxford: Update Software.

— & — (1998) Assertive community treatment for people with severe mental disorders. In *Schizophrenia Module of the Cochrane Database Systematic Reviews* (eds C. G Adams, L. Duggan, J. de Jesus Mari, *et al*). Oxford: Update Software.

McGrew, J. H., Bond, G. R., Dietzen, L. L., et al (1994) Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology*, **62**, 670–678.

— & — (1995) Critical ingredients of assertive community treatment: judgements of the experts. *Journal of Mental Health Administration*, **22**, 113–125.

Meisler, N. (1997) Assertive community treatment initiatives: results from a survey of selected state mental health authorities. *Community Support Network News*, **11**, 3–5.

Morse, C. A., Calsyn, R. J., Allen, G., et al (1992) Experimental comparison of the effects of three treatment programs for homeless mentally ill people. *Hospital and Community Psychiatry*, **43**, 1005–1010.

Mowbray, C. T., Plum, T. R. & Masterton, T. (1998) Harbinger II: deployment and evolution of assertive community treatment in Michigan. *Administration and Policy in Mental Health*, **25**, 125–139.

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.

Stein, L. L. & Test, M. A. (1980) An alternative to mental health treatment. I: Conceptual model, treatment program, and clinical evaluation. *Archives of General Psychiatry*, **37**, 392–397.

Taylor, R. E., Leese, M., Clarkson, P., et al (1998) Quality of life outcomes for intensive *versus* standard community mental health services. PRiSM Psychosis Study 9. *British Journal of Psychiatry*, **173**, 416–422.

Teague, G. R., Bond, G. R. & Drake, R. E. (1998) Program fidelity in assertive community treatment: development and use of a measure. *American Journal of Orthopsychiatry*, **68**, 216–232.

Toesson, M. & Hambridge, J. A. (1992) Mobile community treatment in inner city and suburban Sydney. *Psychiatric Quarterly*, **63**, 119–127.

Test, M. A. (1992) Training in community living. In *Handbook of Psychiatric Rehabilitation* (ed. R. P. Liberman), pp. 153–170. New York: Macmillan.

Thornicroft, G., Strathdee, G., Phelan, M., et al (1998a) Rationale and design. PRiSM Psychosis Study 1. *British Journal of Psychiatry*, **173**, 363–370.

—, Wykes, T., Holloway, F., et al (1998b) From efficacy to effectiveness in community mental health services. PRiSM Psychosis Study 10. *British Journal of Psychiatry*, **173**, 423–427.

Tyrer, P. (1998) Whither community care? *British Journal of Psychiatry*, **173**, 359–360.

Wells, K. B. (1999) Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. *American Journal of Psychiatry*, **156**, 5–10.

Witheridge, T. F., Dincin, J. & Appleby, L. (1982) Working with the most frequent recidivists: a total team approach to assertive resource management. *Psychosocial Rehabilitation Journal*, **5**, 9–11.

Wykes, T., Leese, M., Taylor, R., et al (1998) Effects of community services on disability and symptoms. PRiSM Psychosis Study 4. *British Journal of Psychiatry*, **173**, 385–390.

MAX MARSHALL, MD, Department of Community Psychiatry, University of Manchester; GARY BOND, PhD, Indiana University–Purdue University, Indianapolis, IN, USA; LEONARD I. STEIN, MD, Department of Psychiatry, University of Wisconsin Medical School, Madison, WI, USA; GEOFF SHEPHERD, PhD, Guy's and St Thomas' Medical School, London; JOHN M McGREW, PhD, Indiana University–Purdue University, Indianapolis, IN, USA; JOHN HOULT, FRANZCP, Camden and Islington National Health Service Trust, London; ALAN ROSEN, FRANZCP, Royal North Shorte Hospital and Community Mental Health Services, Sydney, Australia and Universities of Wollongong and Sydney, Australia; PETER HUXLEY, PhD, Department of Psychiatry, University of Manchester; RONALD J. DIAMOND, MD, University of Wisconsin and Mental Health Center of Dane County, Wisconsin, Madison, WI, USA; RICHARD WARNER, MB, Mental Health Center of Boulder County and University of Colorado, CO, USA; MELODY OLSEN, MPA, formerly of Texas Department of Mental Health and Mental Retardation, Austin, TX, USA; ERIC LATIMER, PhD, Douglas Hospital Research Centre and Department of Psychiatry, McGill University, Montreal, Canada; PAULA GOERING, PhD, Department of Psychiatry, University of Toronto, Canada, and Health Systems Research Unit, Centre for Addiction and Mental Health, Clarke Division, Toronto, Canada; TOM K. J. CRAIG, PhD, Guy's, King's and St Thomas' School of Medicine, Dentistry and Biomedical Sciences, London, and Lambeth Healthcare National Health Service Trust, St Thomas' Hospital, London; NEIL MEISLER, MSW, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charlton, SC, USA; MARY ANN TEST, PhD, School of Social Work, University of Wisconsin, Madison, WI, USA

AUTHORSHIP: Max Marshall and Gary Bond wrote the original draft with assistance from Geoff Shepherd, Leonard Stein and John McGrew. Al Rosen, Neil Meisler, Melody Olsen, Peter Huxley, Dick Warner, Ron Diamond, John Hoult, Paula Goering, Eric Latimer and Tom Craig commented on the draft, suggested alterations and expressed a wish to be co-signatories. Mary Ann Test expressed a wish to be a co-signatory.

Correspondence: Dr M. Marshall, Department of Community Psychiatry, University of Manchester, Guild Academic Centre, Royal Preston Hospital, Sharoe Green Lane, Preston PR2 9HT

(First received 1 February 1999, final revision 3 March 1999, accepted 28 April 1999)