

A CULTURE OF ENQUIRY: RESEARCH EVIDENCE AND THE THERAPEUTIC COMMUNITY

Jan Lees, Nick Manning, Ph.D.,
and Barbara Rawlings, Ph.D.

This paper presents data from a systematic review and meta-analysis of 29 published studies of therapeutic community effectiveness using controls, including 8 randomised control trials. Meta-regressions suggest that the two types of therapeutic community, democratic and concept-based, and the age of the study, are the key sources of heterogeneity in the collection of studies analysed. Otherwise, heterogeneity is low and the meta-analysis confirms the effectiveness of therapeutic community treatment with overall summary log odds ratio for the 29 studies of -0.512 (95% ci -0.598 to -0.426).

KEY WORDS: systematic review; meta-analysis; therapeutic communities; effectiveness; evidence.

Jan Lees is Research Associate, Francis Dixon Lodge, Leicester, United Kingdom.

Nick Manning is a Professor of Social Policy and Sociology, University of Nottingham, United Kingdom.

Barbara Rawlings is Honorary Research Fellow, Department of Sociology, University of Manchester, United Kingdom.

Address correspondence to Jan Lees, School of Sociology and Social Policy, University of Nottingham, Nottingham, NG7 2RD, UK.

HISTORICAL ANTECEDENTS

The literature on therapeutic communities refers mainly to democratic psychiatric settings, originating in Britain during the second world war, or to concept-based houses, originating in the USA in the late 1950s. The early literature on British therapeutic communities was descriptive, usually single case studies, and penned by therapeutic community practitioners themselves, rather than impartial or rigorous evaluations (1–5); this sort of “practitioner research” continues today (6,7). The literature then moved to more sociological, anthropological studies of therapeutic communities, conducted by outsiders (8). Although more impartial, these studies were not always well received by practitioners (9), a negative reaction itself then studied by sociologists (10). However, Rapoport’s study of the Henderson Hospital in England became recognised as a seminal text on democratic therapeutic communities, and some of its classifications, for example the four “pillars” of democratisation, permissiveness, reality confrontation and communalism were regarded as criteria to be achieved and standards to maintain. These criteria have been reworked recently to fit in with contemporary democratic therapeutic community practice (11). Subsequently, there were some comparative, again sociological, studies (12), but more recently there has been a growing body of evaluative research in the UK on the effectiveness of the democratic therapeutic community, particularly in mental hospitals and prisons (13–19).

In the United States, a different type of therapeutic community—the “concept house”—emerged, with the founding of Synanon by Charles Dederich, and then Daytop Village, aimed at the treatment of drug addicts. Here again, the early literature was descriptive, but also anthropological and sociological, and mainly written by outsiders to the therapeutic community, more interested in organisational than practice issues (20,21). However, the focus of therapeutic community literature in the USA changed quite quickly, from the early 1970s, to concentrate on the evaluation of effectiveness as in De Leon’s long series of evaluations of therapeutic community treatment for drug abusers (22,23). Although similar studies were carried out in some UK concept-based therapeutic communities (24,25), the trend for such work there did not continue. In the USA however, evaluation research became central to the therapeutic community movement in the community (23) and in prisons (26). The positive results demonstrated through these evaluation studies have played a large part in developing the therapeutic community into a major global player in the drug treatment field.

So, research on therapeutic communities generally has moved through a series of stages, beginning with the early descriptive, informative accounts of the therapeutic community pioneers, through the qualitative, ethnographic accounts of the early outside researchers, into the quantitative, evaluative studies of more recent times. In addition, therapeutic community research has become more sophisticated in terms of research designs and methodologies, and understanding of the complexity of the issues is greater. The questions “does it work?” and “who does it work for?” have become increasingly more urgent as funding has become more and more tied to results. The question “*how* does it work?” is still largely unanswered, since it is difficult to tease out the mechanisms at work inside such a complex and multi-faceted treatment as a therapeutic community. A current study in England and Scotland, sponsored by the Association of Therapeutic Communities, is partially directed at this latter question, as well as the former two—it involves a national comparative study of the effectiveness of 19 democratic, and 2 concept-based, therapeutic communities in treating people with personality disorders, in secure and nonsecure, residential and day, settings, and will use multilevel modelling to analyse the data (27). (A more detailed overview and description of therapeutic community research can be found in Lees (28))

METHODOLOGICAL ANTECEDENTS

In the research field generally, the medical model for the hierarchy of quality of research evidence has prevailed—with initially, the randomised controlled trial being seen as the most rigorous way of obtaining hard evidence of effectiveness, and latterly meta-analyses being viewed as the apex of research and evidence-based practice. In the UK, the National Health Service hierarchy of evidence, as stated in the National Service Framework for Mental Health, is:

- *Type I evidence*—at least one good systematic review, including at least one randomized controlled trial
- *Type II evidence*—at least one good randomised controlled trial
- *Type III evidence*—at least one well designed intervention study without randomization
- *Type IV evidence*—at least one well designed observational study
- *Type V evidence*—expert opinion, including the opinion of service users and carers

NATIONAL HEALTH SERVICE CENTRE FOR REVIEWS AND DISSEMINATION—SYSTEMATIC LITERATURE REVIEW

In 1997, the English Department of Health commissioned a systematic, international literature review of therapeutic community effectiveness, focussing primarily on in- and post-treatment outcome for people with personality disorders in democratic therapeutic communities in secure and nonsecure, psychiatric and other settings (29), but also including the relevant post-treatment and in-treatment outcome studies of the effectiveness of concept-based therapeutic communities, usually for substance abusers, and particularly those in secure settings. Concept-based therapeutic communities were included because of their powerful presence in the research literature on therapeutic community effectiveness, in general, but particularly in the United States of America, Canada, and other parts of the world.

Systematic reviews differ from other types of literature review, in that “they adhere to a strict scientific design in order to make them more comprehensive, to minimise the chance of bias, and so ensure their reliability. Rather than reflecting the views of the authors or being based on only a (possibly biased) selection of the published literature, they contain a comprehensive summary of the available evidence.” (30).

FINDINGS FROM THE SYSTEMATIC LITERATURE REVIEW

The published literature search began with the Cochrane Library databases, to ensure no other similar review was being undertaken. Although there was no fixed start date for the searches, most of the research and effectiveness literature is post World War II (1946–1997). Citation searches were undertaken of a number of books on therapeutic communities, published between 1974 and 1997, and of the *International Journal of Therapeutic Communities/Therapeutic Communities*. Twenty six electronic databases were selected for inclusion in the review, and searched using predesigned search strategies, based on a preselected list of keywords (itemised in the review)—inevitably, these were the main source of information for this review. Hand-searches were also undertaken of a wide range of journals (both English- and foreign-language), and other relevant publications, including “grey literature.” Visits were made to a number of key sites (libraries, therapeutic communities and the Internet).

These searches produced 8,160 book, conference and journal article references, which included very little “grey literature.” While

considerable effort was made to include foreign-language references in the review, it was difficult to locate and gain access to these, and the preference of the majority of databases for English language journals, and a bias towards North American scholarship are factors recognised by researchers in the field of bibliometric analysis (31,32). Most references retrieved for the review are from the UK and USA, and most therapeutic communities identified are in the UK and USA. However, looking at the types of therapeutic community, the UK has produced more articles about the democratic therapeutic community than the USA, while referring to roughly the same number of therapeutic communities, whereas for the concept-based therapeutic community, the USA dominates the field, in terms of the number of articles and the number of therapeutic communities referred to, while, in turn, the articles about concept-based therapeutic communities are dominated by both a few concept-based therapeutic communities, such as “Stay’n out” and CREST (prison therapeutic communities), and by a few authors, such as DeLeon, Wexler, Inciardi and Condelli.

The 8,160 retrieved references were then catalogued by two members of the research team working independently, and according to predetermined inclusion/exclusion criteria, based both on the type of study design, and whether or not the study was evaluative, with an emphasis on post-treatment outcome studies (29, pp. 34–40). The team set an upper limit of 300 as workable, but in fact this level of sorting actually produced 294 references for the final review.

These 294 articles/books were scanned to ensure they were of sufficient quality to include in the final report, and/or that they included sufficient relevant information for data extraction. Those articles or books providing information on post-treatment or in-treatment outcome of therapeutic community treatment in secure or nonsecure democratic therapeutic community settings for people with personality disorders or mentally disordered offenders were selected for more in-depth analysis, together with those for secure concept-based therapeutic communities. These latter came to 113, of which 18 are review articles. The final number of studies included is less than the number of articles found, because some articles relate to different aspects of the same study. The overall spread of these studies is outlined in the review (29, p. 40).

The review then concentrated on the post-treatment outcome findings—there were 10 randomised controlled trial studies, 10 cross-institutional, cross-treatment or comparative studies, and a further 32 using some kind of control or comparison group. The latter were taken as the minimum level of rigour that is acceptable, which left a total of

52 acceptable studies, all of which are discussed in some detail in the review (29). Of these 52, 41 relate to democratic therapeutic communities, and 11 to concept-based therapeutic communities.

Although many of the findings are presented in narrative form, the review also provides a systematic meta-analysis of some of the studies, using odds ratios. Because of the quality of the data presented, and of the analyses of this data in the studies, the meta-analysis only included 29 studies in total, dated between 1960 and 1998, including 8 of the randomised controlled trials. Where there was a choice of outcome measures and control groups, emphasis was placed on conservative criteria, such as reconviction rates rather than psychological improvements, and on non-treated controls. The meta-analysis is discussed in more detail below.

META-ANALYSIS

Twenty-nine studies are included in the meta-analysis [see (29) for a list of these]. The standard way to combine these studies into a summary measure of the effectiveness of the treatment (the therapeutic community, in this case) is slightly more complex than simply adding up the number of studies “for and against.” This is because the importance of each study varies, depending on: the size of each sample; the confidence intervals around each result; the size and direction of the result; and, most importantly, the extent to which there is systematic variation between the studies as a group—the “heterogeneity” of the whole collection of studies being considered. These factors are systematically taken into account by calculating the summary odds ratio of the group, and running a number of tests (“meta-regressions”) to look for and account for any heterogeneity in the whole group. The summary odds ratio can be of two types—“fixed” or “random” effects—depending on the assumptions made about the representativeness of the study samples to the overall population.

The log odds ratio for each study is calculated by constructing a two by two table comparing the numbers of those with a successful outcome, and the rest, in the separate treatment and control groups in the study. A summary odds ratio combines the results of the whole group taking into account the variations in sample size, effects, and so on. This was calculated and the variance estimated for each of the 29 studies, using the Woolf method (the sum of the reciprocals of the cell counts (33)). For one study there was an observed zero, and 0.5 was added to each cell before performing the calculations. Figure 1 shows a

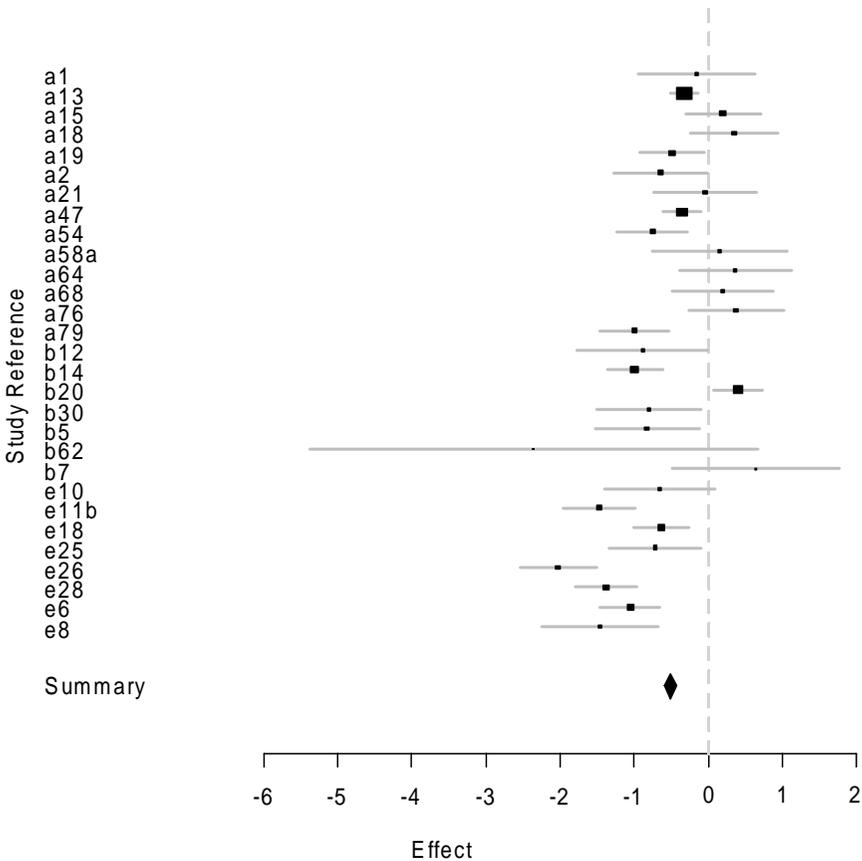


FIGURE 1. Forest plot.

“forest plot” of the study effects—each study is identified by the number [see (29) for references]. Each horizontal line represents the results of one trial; the shorter the line, the more certain the result, with the position of the black square indicating the odds ratio; the bigger the square, the more weight is given to the study, taking into account sample size, range of confidence intervals, etc.; the bottom diamond represents their combined results expressed in a summary odds ratio. The vertical line indicates the position around which the horizontal lines would cluster if the two treatments compared in the trials had similar effects; if a horizontal line touches the vertical line, it means that that particular trial found no clear difference between the treatments. The

position of the diamond to the left of the vertical line indicates that the treatment studied is beneficial. Horizontal lines or a diamond to the right of the line would show that the treatment did more harm than good. The overall summary log odds ratio is -0.512 (95% ci -0.598 to -0.426), which indicates a strong positive effect for therapeutic community treatment.

Meta-Regressions

Two key measures of the quality of any meta-analysis are: the extent to which the variation between the study results can be accounted for by factors other than the variable effectiveness of the treatment under study; and whether there is evidence that only “successful” results have been published—implying that there might be a hidden set of unpublished “unsuccessful” studies that, if included, might reduce the overall positive outcome. If there is large variation that cannot be accounted for, the implication is that the studies are too dissimilar to be combined—for example, they may not be using the same treatment method. In this particular case, there is some possibility of this, since therapeutic community treatment is multidimensional and complex, and may vary from one setting or time period to another. This is conventionally examined by looking at the extent and sources of heterogeneity (i.e. variation) between the studies. Tests can be done for both heterogeneity and publication bias through “meta-regression.”

A “fixed effects” analysis of the log odds ratios, which does not assume they are representative of any wider population, revealed considerable heterogeneity ($\chi^2 = 170.2$ with 28 df) which is significant at any reasonable level. Rather than simply trying to accommodate this extra variance with a “random effects” model, it was suspected that the heterogeneity was caused by two factors: the combination of democratic and concept TCs in the group, and age of the studies—spread across 40 years. We therefore tried to account for it with various predictor variables in a set of “meta-regressions.”

The following predictor variables were used: whether the study was a randomised controlled trial (RCT) or not (2 levels); whether the therapeutic community was a democratic or concept type, and if democratic, whether it was in a secure environment (e.g. prison) or not (all the concept-based therapeutic communities were in a secure environment) (thus a factor with 3 levels); year of publication (as a variate—in years since 1960, as the earliest study was in 1960); and the study precision (as a variate—measured by the standard error—to check for publication bias). Details of this modelling approach can be found in (34). We

TABLE 1
Effect Sizes from Model with Therapeutic Community Type

<i>Effect</i>	<i>Log odds</i>	<i>SE</i>
Intercept	-0.29	0.057
Secure, democratic	Reference category	
Non-secure, democratic	-0.04	0.122
Concept	-0.86	0.106

present a series of fixed effects meta-regressions using each predictor singly:

- a) Whether the study was an RCT or not accounted for relatively little of the heterogeneity, $\chi^2 = 0.74$ with 1 df.
- b) The therapeutic community type (democratic, concept, secure) accounted for much of the observed heterogeneity. Overall $\chi^2 = 70.2$ with 2 df. The effect sizes are shown in Table 1, indicating that while nonsecure therapeutic communities are slightly more effective than secure therapeutic communities, the key difference is that concept-based therapeutic communities are markedly more effective in this collection of studies than democratic therapeutic communities:
- c) The year of publication also accounted for much of the heterogeneity, $\chi^2 = 42.6$ with 1 df. It is of course confounded with therapeutic community type, since the concept TC studies were in general published later than the democratic TC studies. The effect sizes are shown in Table 2, indicating that later studies have a stronger effect size by about 0.03 per year:
- d) Figure 2 shows a “funnel plot” of the effect sizes, based on Egger et al (35). To check for publication bias, whereby negative outcomes had not been written up for publication, it is necessary to look for an uneven pattern, or gap, to the right hand side of the

TABLE 2
Effect Sizes from Model with Year of Publication

<i>Effect</i>	<i>Log odds</i>	<i>SE</i>
Intercept	0.30	0.132
Year of publication (per year)	-0.03	0.004

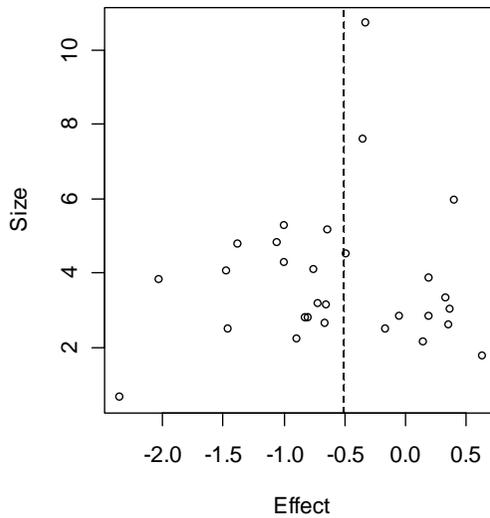


FIGURE 2. Funnel plot.

vertical dotted line, where those studies are “missing.” Visually there do not seem to be any such gaps, and formal analysis using the standard error as a predictor (which is equivalent to the Egger method) shows that there is little evidence of publication bias, overall $\chi^2 = 1.52$ with 1 df.

In conclusion, the sources of heterogeneity are largely accounted for by the differences between types of therapeutic community, with concept-based studies being generally reported in recent years. Having an explicable account of this heterogeneity allows to conclude more confidently that the overall summary odds ratio generated by this set of studies does provide evidence for the effectiveness of therapeutic community treatment for personality disordered patients. This is Type I evidence, according to the UK National Service Framework for Mental Health—at least one good systematic review, including at least one randomised controlled trial.

DISCUSSION

Systematic literature reviews like this provide very useful information on an overview of therapeutic communities research literature, and are well respected as a research tool, particularly within the NHS in

Britain. The work involved in a systematic review of this kind is considerable. The number of potential sources of material, and possible studies for inclusion, was amplified by the wide scope of the review, since it included not merely high quality randomised controlled trials, but also the many descriptive studies that have been published over the years. This ensures the integrity of the final result. Lees et al (29) did not feel that any material, including “grey literature” and unpublished work, exists that was not found for this review. However, the review was commissioned with the assumption that there were very few, if any, high quality comparison studies or randomised controlled studies to find, but the broad remit to include descriptive studies probably might not have been given, had it been suspected that there were so many high quality research studies to include.

Lees et al (29) found several trends in terms of the general scope of the literature. Although there were studies from more than 30 countries, the best evaluations have come out of the USA in recent years, many of which are focussed on concept-based therapeutic communities. Democratic therapeutic communities are more common in Europe, where the most useful studies are typically older than those in the USA. The target treatment populations are people with addictions in the USA, and personality disorders in Europe. Personality disorder itself is a rapidly expanding psychiatric and forensic category, on the back of which there is now a resurgence of interest in and government funding for therapeutic communities, both democratic and concept-based, in the UK, which work with people with personality disorders. However, the definitions and meaning of ‘personality disorder’ have changed rapidly and dramatically over the last 25 years (36, p. 12), as has the name—from character disorder and psychopath. Reed has argued that these diverse meanings have undermined the effectiveness of treatment evaluation (37, p. 34); certainly these conceptual expansions make judgements about research reports difficult, particularly if they are more than 15 years old. Manning has presented sociological analyses of the way in which these definitions have changed (38,39). In addition, in Britain, the Mental Health Acts (1959, 1983) separated psychopathy from other conditions, defined it behaviourally, and presented a pessimistic view of treatment interventions, although this view is changing.

Several methodological conclusions are also highlighted in the review (29). The difficulties presented by the randomised controlled trials reported in the literature include treatment complexity, treatment dosage and treatment integrity, population selection, dropouts, effects decay, and diagnostic shifts. Fundamentally, the nature of the treatment and the nature of the disorder need to be clearly understood and articulated.

Evaluating the effectiveness of therapeutic communities depends crucially on a clear understanding of what the therapeutic community is, the setting (whether secure or nonsecure) in which it is delivered, and at which client groups it is aimed. These elements are all evidently contestable, both within a largely sympathetic literature, and within a smaller, hostile literature.

There is a further complication for the evaluation of therapeutic communities, in that members accepted for therapeutic community treatment are normally both self-selected (in terms of application to join) and community-selected (in terms of the whole community, not just staff or consultants). This means there is no independent decision that can guarantee referral and acceptance, and consequently randomisation is difficult to set up and sustain.

The definition of therapeutic communities has also been difficult. The two main types of therapeutic communities—democratic and concept-based—emerged from quite separate origins. For some writers, these were variations on a basically common theme (21), one dealing with deeper intrapsychic change and the other with initial behavioral control; for others they have nothing in common but the name (40); while more recently, De Leon has argued that, in North America at least, the early differences between the two main variants of the therapeutic community have receded (41, p. 91). In addition, democratic type therapeutic communities developed in prisons or secure settings are inevitably modified by the requirements of prison regulations concerning security and control.

Moreover, since one of the criteria identified above for the definition of a therapeutic community is that of self-identity, there is a difficulty over treatment integrity—how do we know that the self-identity of a particular programme is not erroneous; and what happens when therapeutic communities, as they have been known to do, go through periods of disruption, or sluggishness; how pure is the therapeutic community; and how representative is it of a therapeutic community treatment modality? This question about one point at which the treatment itself is compromised is raised by Clarke & Cornish (42) in their study of an English therapeutic community. These issues are currently being addressed in the UK by the Association of Therapeutic Communities' peer-review process, known as the Community of Communities (27).

Treatment in therapeutic communities takes time—typically around six to twelve months. This heightens the possibility that patients will leave prematurely. In fact, dropouts from therapeutic community treatments are commonplace, and present difficulties for research studies. The US literature on addiction therapeutic communities contains

numerous articles on such “splittees.” In treatment terms, there is a clear association between in-treatment improvement and length of stay (43), leading to a concern with retaining patients in the program. In research, sustaining comparable dosage is an essential prerequisite for evaluation, which can be seriously compromised by dropouts.

Even if randomisation is achieved and treatment is successfully delivered and measured, the problem remains of identifying the point at which improvement should be measured. In penal research, it has been possible to follow up failures over quite long periods of time through the use of criminal records, for example for five years or more. However, many studies have looked at change while still in treatment, at the end of treatment, or at a year post-treatment. Clearly, given the likely effects of post-treatment experiences and effects, sustained effects over long periods even if smaller, may be more convincing than larger effects early on which are not sustained. The solution to this problem was felt to be the measurement of intermediate change during and soon after treatment, and the use of cross-institutional designs (see 27) to capture changes during treatment.

Although the difficulties of undertaking randomized controlled trials are widely appreciated, there is still a case for trying to undertake them for democratic therapeutic community treatment in the UK, since the most recent randomized controlled trials have been undertaken in the USA on concept-based therapeutic communities. The randomized controlled trial is undoubtedly a powerful design, where appropriate, and has in many respects become the gold standard for evidence-based medicine. The alternative, as Clarke & Cornish (42) argued, is a large scale cross-institutional study of therapeutic communities in the field, and such a study is now ongoing in the UK (see above, and 27).

CONCLUSION

In the past, therapeutic communities, at least in the UK, have been ambivalent about both the need for, and the usefulness of, research, and concerned about the impact of its execution on members of the therapeutic communities involved. Lees et al’s systematic literature review concluded that therapeutic communities have not produced the level or quality of research literature that might have been expected, given the length of time they have been in existence, and the quality of staff known to exist in therapeutic communities. This may be due partly to a lack of emphasis placed on research in the early days of therapeutic community development, and more recently to a lack of

resources, in terms of finance, staff and adequate research methodologies, designs and instruments. However, these attitudes have changed in recent years. The medical model of research has come to dominate this field of activity, because the medical research emphasis on evidence-based practice, in a time of scarce resources and funding, has made it imperative for all treatments to prove their effectiveness and efficiency, in order to survive. This has meant that evaluative research has come to take primary importance. It is clear that therapeutic communities need more, and more good quality, and comparative, research in order to counter the charge that there is not a proven case that therapeutic communities are effective, and that they are expensive. There is clinical evidence that therapeutic communities produce changes in people's mental health and functioning, but this needs to be complemented by good quality qualitative and quantitative research studies.

Lees et al's (29) systematic review furnished a substantial number of studies of sufficient quality to undertake a meta-analysis. This analysis, taking careful account of sources of heterogeneity and possible publication bias, shows a clear and positive treatment effect for therapeutic communities.

Lees et al concluded that future research on the effectiveness of therapeutic communities should include further randomised controlled trials. However, these have to overcome the difficulties posed by therapeutic communities controlling their own intake, and the multidimensional and volatile nature of the treatment intervention. Future research should also include more complex, cross-institutional studies in the field, together with further cost-offset studies to complement those few already developed (29).

REFERENCES

1. Jones M: *Social Psychiatry*. London, Tavistock, 1952.
2. Bion WR: *Experiences in Groups*. London, Tavistock, 1960.
3. Foulkes SR: *Introduction to Group Analytic Psychotherapy*. London, Heinemann, 1948.
4. Main T: The Hospital as a Therapeutic Institution. *Bulletin of the Menninger Clinic* 10:66-70, 1946. (Reprinted in *Therapeutic Communities* 17(2):77-80).
5. Harrison T: Bion, Rickman, Foulkes and the Northfield Experiments. *Advancing on a Different Front*. London, Jessica Kingsley, 2000.
6. Black T: Sound and fury: Grief and despair in the large group, in *Therapeutic Communities: Past, Present and Future*. Edited by Campling P, Haigh R. London, Jessica Kingsley, 1999. pp. 107-115.
7. Coombe P: Glimpses of a Cassel Hospital Outpatient Group. *Australian and New Zealand Journal of Psychiatry* 29:309-315, 1995.

8. Rapoport RN: *Community as Doctor*. London, Tavistock, 1960.
9. Jones M: *Beyond the Therapeutic Community*. New Haven, Yale University Press, 1968.
10. Rapoport RN, Manning N: Rejection and re-incorporation: A case study in social research utilisation. *Social Science and Medicine* 10:458–468, 1976.
11. Haigh R: The quintessence of a therapeutic environment: Five universal qualities, in *Therapeutic Communities. Past, Present and Future*. Edited by Campling P, Haigh R. London, Jessica Kingsley, 1999, pp. 246–257.
12. Bloor MJ, McKeganey NP, Fonkert JD: *One Foot in Eden*. London, Routledge, 1988.
13. Chiesa M, Fonagy P: Cassel personality disorder study: Methodology and treatment effects. *British Journal of Psychiatry* 176:485–491, 2000.
14. Chiesa M, Iacoponi E, Morris M: Changes in health service utilization by patients with severe personality disorders before and after inpatient psychosocial treatment. *British Journal of Psychotherapy* 12(4):501–512, 1996.
15. Cullen E: Grendon: The therapeutic prison that works *Therapeutic Communities* 15(4):301–311, 1994.
16. Marshall P: *A Reconviction Study of HMP Grendon Therapeutic Community (Research findings no. 53)*. London, Home Office Research and Statistics Directorate, 1997.
17. Newton M: *Changes in Test Scores During Treatment at Grendon*. Research and Development Unit, Grendon Underwood. Bucks, UK, HMP Springhill and Grendon, Unpublished research paper, 1997.
18. Shine J: *A Compilation of Grendon Research*. Aylesbury, UK, HMP Grendon, 2000.
19. Warren F, Dolan B: *Perspectives on Henderson Hospital*, 2nd ed. Sutton, Surrey, Henderson Hospital, 2001.
20. Yablonsky L: *Synanon: The Tunnel Back*. New York, Macmillan, 1965.
21. Sugarman B: *Daytop Village: A Therapeutic Community*. New York, Holt, Rinehart and Winston, 1974.
22. De Leon G: The Phoenix House Therapeutic Community: Changes in psychopathological signs. *Archives of General Psychiatry* 28:131–135, 1973.
23. De Leon G: Residential Therapeutic Communities in the Mainstream: Diversity and issues. *Journal of Psychoactive Drugs* 27:(1):3–15, 1994.
24. Ogbourne AC, Melotte C: An evaluation of a Therapeutic Community for former drug users. *British Journal of Addiction* 72(1):75–82, 1977.
25. Wilson S, Mandelbrote B: The relationship between duration of treatment in a therapeutic community for drug abusers and subsequent criminality. *British Journal of Psychiatry* 132:487–491, 1978.
26. Wexler H: *Therapeutic Communities in American prisons, in Therapeutic Community for Offenders*. Edited by Cullen E, Jones L, Woodward R. Chichester, Wiley, 1997.
27. Lees J: Practice evaluation of Therapeutic Communities, in *A Culture of Enquiry. Research Evidence and the Therapeutic Community*. Edited by Lees J, Manning N, Menzies D, Morant N. London, Jessica Kingsley, 2004, pp. 76–90.
28. Lees J: 'Research. The Importance of Asking Questions,' in *Therapeutic Communities. Past, Present and Future*. Edited by Campling P, Haigh R. London, Jessica Kingsley, 1999, pp. 207–222.
29. Lees J, Manning N, Rawlings B: *Therapeutic Community Effectiveness: A Systematic International Review of Therapeutic Community Treatment for People with Personality Disorders and Mentally Disordered Offenders (CRD Report no. 17)*. York, NHS Centre for Reviews and Dissemination, University of York, 1999.

30. CRD (Centre for Reviews and Dissemination): Undertaking Systematic reviews of Research on Effectiveness (CRD Report no. 4). York, University of York, 1996.
31. Artus HM: Science indicators derived from databases. *Scientometrics* 37:297–311, 1996.
32. Seglen PO: Why the impact factor of journals should not be used for evaluating research. *British Medical Journal* 314:498–502, 1997.
33. Kahn HA, Sempos CT: *Statistical Methods in Epidemiology*. Oxford, Oxford University Press, 1989.
34. Thompson SG, Sharp SJ: Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine* 18:2693–2708, 1999.
35. Egger M, Davey Smith G, Schneider M, Minder C: Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 315:629–634, 1997.
36. Gunderson JG: Building structure for the borderline construct. *Acta Psychiatrica Scandinavica* 89(suppl. 379):12–18, 1994.
37. DoH/HO: Report of the Department of Health/Home Office working group on psychopathic disorder. London, HMSO, 1994.
38. Manning N: Psychiatric diagnosis under conditions of uncertainty: Personality disorder, science, and professional legitimacy. *Sociology of Health and Illness* 22(5):621–639, 2000.
39. Manning N: Actor-networks, policy networks and personality disorder. *Sociology of Health and Illness* 24(5):644–666, 2002.
40. Glaser AN: Therapeutic communities and therapeutic communities: A personal perspective. *International Journal of Therapeutic Communities* 4(2):150–162, 1983.
41. De Leon G: The research context for therapeutic communities in the USA, in *A Culture of Enquiry. Research Evidence and the Therapeutic Community*. Edited by Lees J, Manning N, Menzies D, Morant N. London, Jessica Kingsley, 2004, pp. 91–108.
42. Clarke RVG, Cornish DB: *The Controlled Trial in Institutional Research—paradigm or pitfall for penal evaluators*. London, HMSO, 1972.
43. Nieminen P: *Therapeutic community research and statistical data analysis*. *Acta Universitatis Ouluensis Medica*, Oulu: University of Oulu Printing Center, 1996.