

Effectiveness of time-limited psychotherapy for minor psychiatric disorders

Randomised controlled trial evaluating immediate v. long-term effects

SERGIO LUÍS BLAY, JAIR S. VEL FLOCKS, MÁRCIA BARRUZI,
MÔNICA CRISTINA DI PIETRO, FÁBIO LEITE GASTAL, ALFREDO MALUF NETO,
MATEUS P. DE SOUZA, LADISLAU R. U. GLAUSIUSZ and MICHAEL DEWEY

Background Psychotherapy research rarely has studied outcome in the longer term.

Aims To evaluate the effectiveness of brief group dynamic psychotherapy (BGDP) intervention in patients with minor psychiatric disorders compared with the usual clinical management shortly after treatment termination and to investigate whether intervention would show a differential effect at 2-year follow-up.

Method Patients were allocated randomly to an experimental or control group. The General Health Questionnaire (GHQ) was used as a primary outcome measure.

Results Based on improvement in the GHQ, at termination of treatment the BGDP group showed a significant improvement in 23 out of 42 (54.8%) compared with 11 out of 41 (26.8%) in the clinical management group. The difference in the total improvement rate is 28% (95% CI 8–48) ($\chi^2=6.7$; d.f.=1; $P=0.009$). In contrast, no differential follow-up effects were found between the BGDP and clinical management groups.

Conclusions Psychotherapy appears to have beneficial effects at termination of treatment but the changes attained were not stable.

Declaration of interest None.

Recently, investigators have begun to report small or no differential benefit of psychotherapies in a longer period of follow-up, compared with assessments near treatment termination, suggesting that the initial response may bear little relationship to the efficacy at clinically meaningful follow-up intervals (Snyder *et al*, 1991; Shea *et al*, 1992; Friedli *et al*, 1997; Ward *et al*, 2000). The longer-term effects of group therapy, either dynamic or cognitive-behavioural, in adult out-patients with non-psychotic disorders are relatively unknown. The present study addresses this deficit. The two major aims were to evaluate the effectiveness of this intervention in patients with minor psychiatric disorders, shortly after treatment termination, compared with usual clinical management and to investigate whether intervention showed a differential effect at 2-year follow-up.

METHOD

Study design

This study was conducted in the Department of Psychiatry of the hospital of the Federal University of São Paulo. It was organised with the aim of comparing the therapeutic effectiveness (Brook & Lohr, 1985) of brief group dynamic psychotherapy (BGDP) for patients with minor psychiatric disorders with the usual out-patient treatment.

Comparisons were conducted shortly after treatment and at 2-year follow-up. A parallel, two-group, randomised trial research design was employed. The interval between randomisation and termination of treatment was 4–8 weeks. Randomisation was carried out using a computer-generated sequence. An assessor not involved in the treatment or initial assessment of patients allocated them to the trial groups. Treatment started within a few days of randomisation.

Outcome measures

Assessments were carried out in three waves: at intake, at termination of treatment and at follow-up (2 years after finishing treatment). Demographic data and information regarding previous psychiatric treatments, medical problems and use of medications were obtained for each subject. Four instruments were used:

- the 12-item General Health Questionnaire (GHQ-12; Goldberg, 1972) was used to assess psychological morbidity; a cut-off of 3/4 was employed for case definition (Mari & Williams, 1985);
- the Clinical Interview Schedule (CIS; Goldberg *et al*, 1970) is a semi-structured psychiatric interview that gives an overall evaluation of the clinical severity of the case (overall severity rate, OSR) on a five-point scale; a psychiatric diagnosis can be recorded and we used the criteria of DSM-IV (American Psychiatric Association, 1994);
- the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967) was used to evaluate the presence and severity of depression; the 17-item version was used with a cut-off of 12/13 for the definition of a depressive condition (Dratcu *et al*, 1987);
- patient impressions of the result of intervention were obtained using the Therapy Session Report (TSR) developed by Orlinsky & Howard (1966). We focused on one item of the TSR scored in a binary fashion: 'progress' and 'no progress'.

Primary outcome measure

The primary outcome measure was improvement in GHQ status. Response was operationally defined as moving from the clinical condition of 'case' ($\text{GHQ} \geq 4$) to 'non-case' ($\text{GHQ} \leq 3$).

Secondary outcome measure: multi-dimensional response

The CIS, HRSD and patient impressions were secondary outcome measures. 'Multi-dimensional response' was defined operationally as achieving simultaneously a final GHQ score of ≤ 3 , a CIS final score of ≤ 19 , an HRSD score of ≤ 12 , an OSR of ≤ 1 and patient impressions scored as 'progress'.

Interviewers

All evaluations were made by psychiatrists trained and experienced in the application of these instruments and masked to the treatment procedures. A reliability study was conducted at the beginning of the study with the examiners evaluating the scales. The results met high standards of agreement (intraclass correlation coefficients 0.98–1.00 for GHQ, 0.70–1.00 for CIS and 0.30–0.94 for HRSD).

The initial assessor was masked to treatment allocation. At each wave (intake, termination and follow-up) the patient was evaluated by a different examiner from the initial assessor in order to avoid any influence of preliminary contact with the patient.

Selection of patients and ethical considerations

Most patients came through the hospital's normal referral procedures, out-patient services or emergency clinics and patients' spontaneous search for psychiatric treatment.

The following inclusion criteria were used: male and female out-patients aged 20–60 years. At intake, they were required to: score at least 4 on the 12-item version of the GHQ; present one of the following types of disorder according to DSM-IV: affective (mild, moderate), anxiety, somatoform, adjustment and sexual problems; and have a duration of the current episode of ≤ 5 years. To select patients who could benefit from psychological group intervention, their perception relating to the psychological nature of the disturbances was examined, based on the criterion developed by Sifneos (1987). The presence of individual distress, interpersonal difficulties or family conflicts and the ability to engage in a productive group activity were also considered (Bloch & Aveline, 1996).

The following were excluded: patients evidencing psychosis, learning disability, severe personality disorder, alcohol and drug misuse, dementia or other organic brain syndrome, severe or unstable medical illness that may create a marked change in mental state; those who, simultaneously, participated in any other psychiatric or psychological treatment through the national health system (out-patient services), private clinics or self-help groups; and those who showed any difficulty proceeding with the treatment follow-up

(for reasons of work, inaccurate or possible change of address, or if they provided no alternative contact in case of change of address).

Eligible patients were informed in detail of the programme and provided signed consent. The Medical Ethics Committee of the Federal University of São Paulo approved this research.

Treatment regimens

Attendance at brief group dynamic psychotherapy (BGDP)

The psychodynamically oriented psychotherapy followed the principles of Sifneos (1987), aiming to help the patient to identify and elaborate problematic themes such as anxiety or interpersonal conflicts. Group techniques included questioning, timed confrontations, dealing with intense feelings aroused by discussion, repetitive patterns, transference and resistance. All patients included in this group (BGDP) were to receive eight sessions of psychotherapy, twice a week, for 1 month.

Two different therapists carried out treatment. They were residents in psychiatry in an accredited residency programme. They were in their third and fourth years of training with at least 500 hours of supervised therapy experience. Training of therapists for the trial consisted of four workshops with a senior therapist of the department staff. Therapists were instructed to treat patients using the instructions given above. This procedure constituted a general guideline and not a rigid manualised treatment (Elliott, 1998). The therapists took notes of the sessions and took part in weekly supervisions throughout the research. A senior group psychotherapist conducted the supervision.

To facilitate the adaptation and participation of the group members, each session started with one of eight video films of 20–30 min duration. Eight themes were selected in accordance with Orlinsky & Howard (1975). The videos were shown in the following order: personal relations, work, religion, alcoholism, family relations, sexuality, violence and honesty. They were chosen from an initial selection of about 50 titles and included soap opera, national films, documentaries and some cartoons. Famous and popular local artists appeared in most of these films. This procedure was intended to facilitate communication, stimulate interpersonal insights, help the patient to identify with

problematic themes and create an emotional environment set in a cultural language appropriate for the socio-economic status of this patient group. Each group therapy session lasted 90 min, including the video.

The clinical management group

These patients, from now on identified as the clinical management group, were treated individually by a psychiatrist. They received one clinical consultation at the out-patient clinic. This lasted 30–40 min, with the present and past clinical and familial history and mental state examination fully annotated. Treatment could include medication, simple problem-solving techniques, advice, manipulation of the environment, reassurance, interpretation and involvement of the patient's partner. The majority of local out-patient services have a treatment format similar to that of clinical management. After the 4-week period, all patients willing to continue treatment were referred to other services.

Follow-up

Patients completing the treatment protocol entered the follow-up phase, conducted 25–26 months after randomisation. The follow-up was naturalistic, with no control over further treatments that patients might receive. Follow-up examinations were scheduled at 24 months after the end of treatment.

Sample size determination and data analysis

The study was planned to use a binary outcome: improvement. This criterion would require demonstration of statistical superiority of the BGDP over the clinical management group by indicating the improvement of clinical status over baseline, at termination and at follow-up. We know that psychotherapy has a general success rate for a treated group that is 25–30% better than an untreated group (Lipsey & Wilson, 1993). Based on a recovery rate of 50% in one group and 20% in the other, with a two-sided α of 0.05 and a power of 0.80%, we calculated that we would require a sample size of 40 patients in each group.

Overall group comparability at baseline, demographic characteristics and diagnostic and group mean scores on the

12-item GHQ and CIS (OSR) were evaluated using appropriate summary statistics.

The statistical analysis was performed on an intention-to-treat basis. The effect of psychotherapy on the primary outcome measure – improvement in the GHQ – was tested shortly after termination and follow-up using Pearson's χ^2 or Fisher's exact test (when appropriate, two-sided). Numbers needed to treat were evaluated (Sackett *et al.*, 1997). We also used a sensitivity analysis. Statistics were re-estimated after missing cases were replaced by 'improvement' or 'non-improvement' outcome (Greenhouse & Iyengar, 1994).

Comparisons were made between groups at termination and follow-up by means of an unbalanced repeated-measures model with structured covariance matrices, using maximum likelihood as implemented in the BMDP5V program. This method is more suitable for dealing with incomplete data due to drop-outs (Everitt, 1998). The analysis incorporates any data from missing observations during the trial. The first-order autoregressive structure was used because it achieved the highest value for Akaike's information criterion. Moreover, this method produces less biased results compared with other procedures, such as 'last observation carried forward' (Diggle, 1998).

Logistic regression was used to study which variables predict the result of treatment measured as a categorical variable at

termination and at follow-up assessments. Both the descriptive analyses as well as inferential statistics were carried out using the BMDP/DYNAMIC statistical package (1993 version for DOS).

RESULTS

A total of 84 patients entered the study over 13 months from January 1992. Patient progress through the trial is shown in a CONSORT diagram (see Fig. 1).

Initial group comparability

A total of 67 women and 17 men were randomised. Sample characteristics at intake and outcome measures are displayed in Table 1.

Termination of treatment

Primary outcome measure: response

At intake, all subjects were psychiatric cases showing a GHQ of ≥ 4 . Treatment effects were calculated by comparing the scores obtained by patients at the beginning of the study with those at the end of treatment. At termination of treatment and omitting the data related to the missing BGDG patient, 23 out of 42 (54.8%) improved in the BGDG group and 11 out of 41 (26.8%) in the clinical management group ($\chi^2=6.7$; d.f.=1; $P=0.009$; 95% CI 8–48). The number

needed to treat based on this reduction was four (95% CI 2.1–11.9) (see Table 2).

Using the GHQ score at termination of treatment as a binary outcome, logistic regression with backward elimination was applied to investigate the following possible predictive factors: intervention, gender, first admission, age, education, OSR and presence of physical illness. Three items emerged as being capable of predicting the improvement: intervention (OR=4.72; 95% CI 1.66–13.4), gender (OR=0.33; 95% CI 0.09–1.19) and first admission to mental health services (OR=0.3; 95% CI 0.10–0.88).

Secondary outcome measure: multi-dimensional response

At termination of treatment and omitting the data related to the missing BGDG patient, 9 out of 42 (21.4%) improved in the BGDG group and 0 out of 41 (0.0%) in the clinical management group (Fisher's exact test: $P=0.002$; 95% CI 9.6–33.1). The number needed to treat based on this reduction was four (95% CI 3–10.2) (see Table 3).

Follow-up

Although patients gave informed consent at entry, adherence to follow-up was unpredictable. Subjects were contacted by telephone or post to conduct the follow-up protocol at 2 years. Where we realised that patients were reluctant to participate or had changed address, we carried out some interviews earlier. Other subjects required longer periods in which to be contacted and interviewed. Therefore, the follow-up interval varied: BGDG group, 15–32 months (patients evaluated according to time (months) of follow-up: $<12=0$; $12-15=4$; $15-18=4$; $18-23=6$; $\geq 23=8$); clinical management group, 9–30 months ($<12=1$; $12-15=6$; $15-18=6$; $18-23=7$; $\geq 23=5$). Of the original 42 patients in the BGDG group, 22 (52.3%) were contacted for follow-up, and of the original 41 clinical management group patients, 25 (61%) were contacted for follow-up ($\chi^2=NS$; difference = -8.7% ; 95% CI -30.2 to 12.8). All follow-up evaluations were conducted face to face.

Primary outcome measure: response

At follow-up and omitting the data related to the missing patients, 10 out of

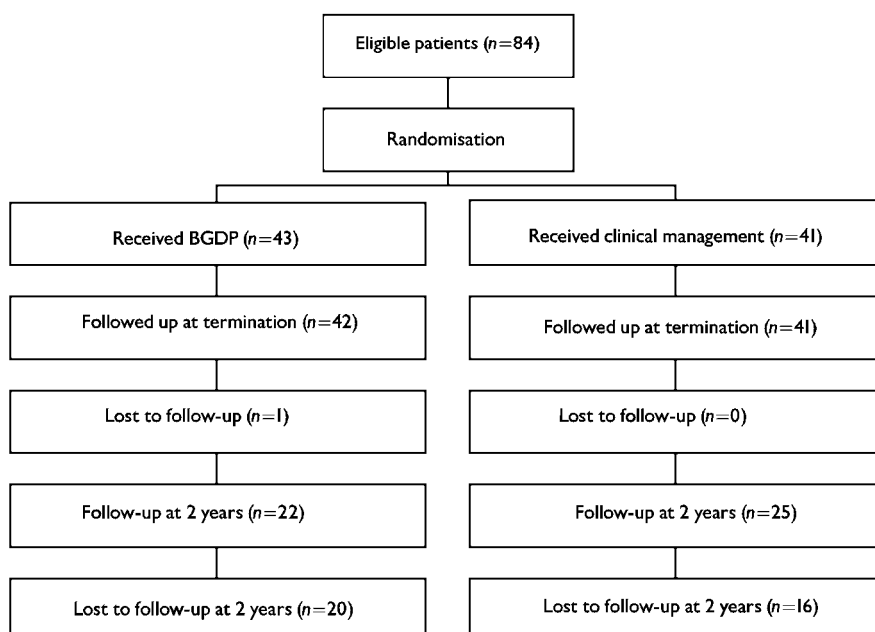


Fig. 1 CONSORT diagram (BGDP, brief group dynamic psychotherapy).

25 (40.0%) improved in the clinical management group (GHQ ≤ 3) and 7 out of 22 (31.8%) in the BGDG group ($\chi^2=NS$). Statistical estimates were unaffected by assuming that the missing cases were either a treatment success or a failure. We also considered the response rates to be similar in the missing and non-missing cases (see Table 4).

Secondary outcome measure: multi-dimensional response

At follow-up and omitting the data related to the missing patients, 6 out of 25 (24%) improved in the clinical management group and 1 out of 22 (5%) in the BGDG group (Fisher's exact test: NS). Statistical estimates were unaffected by assuming

that the missing cases were either a treatment success or a failure (see Table 5).

An unbalanced repeated-measures analysis was carried out to fit the GHQ outcome at follow-up. The main effects of intervention modality, time and the interaction of time and intervention were modelled. There was a significant group by time interaction for GHQ scores (interaction estimate=0.53; 95% CI 0.05–1.01; $P=0.02$), indicating that GHQ scores of patients in the BGDG group change over time at a different rate to those of patients in the clinical management group. Mean GHQ score at the 24-month follow-up is one point lower in the clinical management group than in the BGDG group.

Logistic regression of the probability of improvement was conducted, the model including those variables (at the beginning of the investigation) potentially involved with the outcome. All terms were excluded from the model, indicating that background items were not associated with predicting response at follow-up.

DISCUSSION

First aim: outcome shortly after treatment termination

On our primary outcome measure (GHQ status) we have shown that the BGDG group is statistically superior to the clinical management group as measured by GHQ status, with twice as many patients improved after treatment. On our secondary outcome measure where patients were classified as improved or not, using an *a priori* criterion based on a pool of clinical variables and patient self-report, we also showed evidence of a difference between the clinical management and BGDG groups. These results supplied additional evidence of the superiority of BGDG. We also analysed all the relevant predictors of outcome. With a multivariate procedure the conclusions remain rather similar. The GHQ score at termination can be predicted by intervention itself, gender and chronicity in the health system evaluated at intake. However, we must emphasise that the sample size is small and some risk factors may be missed.

Second aim: outcome at longer-term follow-up

We found that the differential benefits of BGDG over clinical management (measured

Table 1 Baseline characteristics and General Health Questionnaire (GHQ) outcome (n=84)

	BGDG group (n=43)	Clinical management group (n=41)
Gender (n female (%))	36 (83.7)	31 (75.6)
Age, years (mean (s.d.))	35.0 (9.0)	36.8 (10.0)
Educational level (n ≥ 5 years (%))	25 (58.1)	22 (53.6)
Ethnicity (n White (%))	32 (74.4)	28 (68.2)
Marital status (n married (%))	26 (56.0)	33 (80.4)
Income (n \leq US\$100.00/month (%))	26 (60.4)	26 (63.4)
Diagnosis (n anxiety and depressive states (%))	37 (86.0)	36 (87.8)
First admission for treatment (n yes (%))	25 (58.1)	28 (68.3)
Use of medication (n yes (%))	19 (44.1)	23 (56.0)
Concomitant medical problems (n yes (%))	11 (25.6)	10 (24.3)
Symptom ratings (mean (s.d.))		
GHQ	9.4 (3.5)	10.8 (3.8)
CIS	30.23 (8.2)	30.1 (11.31)
HRSD	13.8 (4.7)	13.7 (6.2)
OSR (n ≥ 2 (%))	41 (95.3)	41 (100)
GHQ at termination (mean (s.d.)) (95% CI < 0.18 to 3.08)	4.61 (3.54) ¹	6.24 (3.29)
GHQ at follow-up (mean (s.d.)) (95% CI < -0.61 to 3.19)	6.09 (3.79) ²	4.8 (3.12) ³

BGDG, brief group dynamic psychotherapy; CIS, Clinical Interview Schedule; HRSD, Hamilton Rating Scale for Depression; OSR, overall severity rate.

- 1. n=42.
- 2. n=22.
- 3. n=25.

Table 2 Outcome of intervention (BGDP or clinical management) for responders at the end of treatment

	BGDG group	Clinical management group	Difference in favour of BGDP (95% CI)	χ^2 (d.f.)	P
Omitting the drop-outs					
GHQ < 4	23	11	28% (8 to 48)	6.7 (1)	0.009
GHQ ≥ 4	19	30			
Drop-outs as treatment success					
GHQ < 4	24	11	29% (9.4 to 48.6)	7.25 (1)	0.007
GHQ ≥ 4	19	30			
Drop-outs as treatment failure					
GHQ < 4	23	11	26.7% (7.1 to 46.3)	6.19 (1)	0.012
GHQ ≥ 4	20	30			

BGDG, brief group dynamic psychotherapy; GHQ, General Health Questionnaire.

through GHQ status) were no longer seen. Compared with BGDP, patients allocated to the clinical management group showed greater reductions in GHQ total score. However, when a comparative effect of treatment

did emerge it was not a main effect but an interaction (time \times treatment). Although this difference was statistically significant, the actual difference in mean scores was small. We analysed all relevant predictors of

outcome. No term emerged as a possible predictor using multivariate procedures at follow-up.

Relevant information concerning the design

We would stress some features of this study: it is randomised in an area where there are relatively few randomised studies, particularly with extended follow-up; random allocation of patients produced well-matched samples, particularly with regard to clinical status; and standard evaluation techniques were used, which were applied by evaluators masked to treatment allocation.

Why were differences between treatments reduced?

The results obtained from the patients who were satisfactorily followed up suggest that the gains measured in the BGDP and clinical management groups were not maintained. This situation may be interpreted in several ways. First, it could be argued that the technique used would not have produced such large differences as might have been desired. Eight sessions is a relatively small number and gives little time for patients to gain insight or re-organise their defences and symptomatology. However, there is controversy over this in the literature. Short-term psychotherapy can produce important and lasting changes in the functioning and attitudes of patients (Seligman, 1995). On the other hand, data from effectiveness studies suggest that longer-term treatments, roughly 24 months, may produce more robust effects and may be useful in the treatment of patients presenting the polysymptomatic clinical profile found in everyday practice (Kopta *et al*, 1994; Seligman, 1995).

If we accept the efficacy of psychotherapies, there are ethical and methodological questions about studies with a placebo control group, but this means that the differences between the trial arms will be minimised because both receive effective treatments (Strayhorn, 1987).

Patients treated in our study not only had emotional disturbances but also substantial levels (25%) of comorbidity with medical problems: 37% were frequent users of out-patient services. Pilkonis *et al* (1984) have shown that with patients with a low income and previous

Table 3 Outcome of intervention (BGDP or clinical management) for multi-dimensional responders at the end of treatment

	BGDP group	Clinical management group	Difference in favour of BGDP (95% CI)	P
Omitting the drop-outs				
Improved	9	0	21.4% (9.6 to 33.1)	0.002
Non-improved	33	41		
Drop-outs as treatment success				
Improved	10	0	23.2% (11.5 to 34.9)	0.001
Non-improved	33	41		
Drop-outs as treatment failure				
Improved	9	0	20.9% (9.2 to 32.6)	0.002
Non-improved	34	41		

BGDP, brief group dynamic psychotherapy.

Table 4 Outcome of intervention (BGDP or clinical management) for responders at follow-up (24 months)

	BGDP group	Clinical management group	Difference in favour of BGDP (95% CI)	χ^2 (d.f.)	P
Omitting the drop-outs					
GHQ < 4	7	10	-8.2% (-34.2 to 17.8)	0.34 (1)	0.56
GHQ \geq 4	15	15			
Drop-outs as treatment success					
GHQ < 4	27	26	0.8% (-18.8 to 20.4)	0.01 (1)	0.93
GHQ \geq 4	15	15			
Drop-outs as treatment failure					
GHQ < 4	7	10	-6.7% (-19.4 to 6)	0.76 (1)	0.38
GHQ \geq 4	35	31			

BGDP, brief group dynamic psychotherapy; GHQ, General Health Questionnaire.

Table 5 Outcome of intervention (BGDP or clinical management) for multi-dimensional responders at follow-up (24 months)

	BGDP group	Clinical management group	Difference in favour of BGDP (95% CI)	χ^2 (d.f.)	P
Omitting the drop-outs					
Improved	1	6	-19% (-36.1 to -1.36)		0.10
Non-improved	21	19			
Drop-outs as treatment success					
Improved	21	21	-3.6% (-23.2 to 16)	0.11 (1)	0.74
Non-improved	21	19			
Drop-outs as treatment failure					
Improved	1	6	-12.2% (-21.4 to -2.4)		0.06
Non-improved	41	35			

BGDP, brief group dynamic psychotherapy.

interventions, therapeutic results are more limited. This may have also contributed to the reduction of differences between treatments in our trial.

Are the results generalisable?

Caution is needed with regard to generalisation of our results. First, selection of control groups for psychosocial intervention outcome studies is not an easy one (Strayhorn, 1987). The two very heterogeneous groups did not receive comparable treatment and the results can be interpreted as a consequence of attention placebo. In other words, we do not know whether it is the group itself or the treatment within it that is effective. Another group experience, as a control group, may be of help in solving this problem. However, psychotherapy already has been found to be superior to a psychological placebo in patients with somatic complaints (Guthrie *et al*, 1991). In this research we wanted to compare representative treatment procedures relevant to our health system. In addition, the exclusion on ethical grounds of a non-treatment control fitted the design of the effectiveness trial.

The 24-month response rate for psychological assessment was relatively poor. The most common reason for low response was, as expected, contact loss because of address change. This may have affected the outcome data, because it is possible that patients who were doing less well or who had a more severe clinical picture may have been more reluctant to complete the follow-up.

Criteria for empirically validated treatments recommend the use of manuals in conducting psychotherapy. Although this procedure is not free of controversy (Elliot, 1998), we must emphasise that the unstandardised group method employed in this study may reduce the power of generalisation.

Lastly, based on ethical grounds, the follow-up was naturalistic with no control over life events or any further treatments that patients might receive. Any such additional treatments would confound interpretation of results after extended periods of follow-up.

ACKNOWLEDGEMENTS

Sincere thanks to Professor Peter Fonagy, University College, London, and Professor Brian Cooper, Institute of Psychiatry, London, for their comments

CLINICAL IMPLICATIONS

- This study suggests that brief group dynamic psychotherapy (BGDP) is a superior treatment to clinical management shortly after treatment termination.
- With longer follow-up intervals, the differential effect of BGDP over clinical management was not stable, suggesting that initial response may have little relation to outcome.
- This study suggests that no clinical or socio-demographic variables at the beginning of the investigation were associated with predicting the response at 24 months.

LIMITATIONS

- The high drop-out rate from both groups at the 24-month follow-up made statistical analysis more difficult.
- The follow-up was naturalistic with no control over life events.
- The study was built to fit an effectiveness design, aiming to compare real-life treatment conditions. However, we emphasise the different treatment approach conducted with both arms.

SERGIO LUÍS BLAY, PhD, JAIR S. VEL FUCKS, MD, MÁRCIA BARRUZI, MÔNICA CRISTINA DI PIETRO, MD, FÁBIO LEITE GASTAL, PhD, ALFREDO MALUF NETO, MD, MATEUS P DE SOUZA, MD, LADISLAU R. U. GLAUSIUSZ, MD, Department of Psychiatry of the Federal University of São Paulo (UNIFESP), Brazil; MICHAEL DEWEY, PhD, University of Nottingham, UK

Correspondence: Dr Sergio Luís Blay, Department of Psychiatry, UNIFESP, R. Botucatu 740, CEP 04023-900, São Paulo, Brazil. Tel: 55-11-3816-1030; Fax: 55-11-3816-1030; E-mail: blaysl@psiquiatria.epm.br

(First received 23 May 2001, final revision 2 January 2002, accepted 14 January 2002)

on an earlier draft of this paper. Thanks to Dr Paul Crits-Christoph, University of Pennsylvania and Norman Hearst, University of California for their support and advice. Thanks also to Dr Drew Western, University of Boston, for allowing access to an unpublished manuscript. We are grateful also to Carlos D. Andrade Jr, Eliana Gonçalves and Marcelo C. Zappitelli, who carried out the testing of some cases. This work was executed with the financial support of Fundação de Amparo à Pesquisa do Estado de São Paulo (92/4152-1) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (500 525/92-7).

REFERENCES

- American Psychiatric Association (1994)** *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). Washington, DC: APA.
- Bloch, S. & Aveline, M. (1996)** Group psychotherapy. In *An Introduction to the Psychotherapies* (ed. S. Bloch), pp. 83–115. Oxford: Oxford Medical Publications.
- Brook, R. H. & Lohr, K. N. (1985)** Efficacy, effectiveness, variations, and quality. *Medical Care*, **5**, 710–722.
- Diggle, P. J. (1998)** Dealing with missing values in longitudinal studies. In *Statistical Analysis of Medical Data* (eds B. S. Everitt & G. Dunn), pp. 203–228. London: Arnold.
- Dratcu, L., Costa Ribeiro, L. & Calil, H. M. (1987)** Depression assessment in Brazil. The first application of the Montgomery–Åsberg Depression Rating Scale. *British Journal of Psychiatry*, **150**, 797–800.
- Elliot, R. (1998)** Editor's introduction: a guide to the empirically supported treatments controversy. *Psychotherapy Research*, **8**, 115–125.
- Everitt, B. S. (1998)** Analysis of longitudinal data: beyond MANOVA. *British Journal of Psychiatry*, **172**, 7–10.
- Friedli, K., King, M. B., Lloyd, M., et al (1997)** Randomised controlled assessment of nondirective psychotherapy versus routine general-practitioner care. *Lancet*, **350**, 1662–1665.
- Goldberg, D. P., Cooper, B., Eastwood, M. R., et al (1970)** A standardised psychiatric interview for use in

community surveys. *British Journal of Preventive and Social Medicine*, **24**, 18–23.

— (1972) *The Detection of Psychiatric Illness by Questionnaire* (Maudsley Monograph No. 21). Oxford: Oxford University Press.

Greenhouse, J. & Iyengar, S. (1994) Sensitivity analysis and diagnostics. In *The Handbook of Research Synthesis* (eds H. Cooper & L. Hedges), pp. 383–398. New York: Russel Sage Foundation.

Guthrie, E., Creed, F., Dawson, D., et al (1991) A controlled trial of psychological treatment for the irritable bowel syndrome. *Gastroenterology*, **100**, 450–457.

Hamilton, M. (1967) Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, **6**, 278–296.

Kopta, S., Howard, K., Lowry, J., et al (1994) Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, **62**, 1009–1016.

Lipsey, M. W. & Wilson, W. B. (1993) The efficacy of psychological, educational and behavioral treatment. *American Psychologist*, **48**, 1181–1209

Mari, J. J. & Williams, P. (1985) A comparison of the validity of two psychiatric screening questionnaires (GHQ 12 and SRQ 20) in Brazil, using relative operating characteristic (ROC) analysis. *Psychological Medicine*, **15**, 651–659.

Orlinsky, D. E. & Howard, K. I. (1966) *Therapy Session Report (Form T)*. Chicago: Institute of Juvenile Research.

— & — (1975) *Varieties of Psychotherapeutic Experience: Multivariate Analysis of Patients' and Therapists' Reports*. New York: Teachers College Press.

Pilkonis, P. A., Imber, S. D., Lewis, P., et al (1984) A comparative outcome study of individual, group, and conjoint psychotherapy. *Archives of General Psychiatry*, **41**, 431–437.

Sackett, D. L., Richardson, W. S., Rosenberg, W., et al (1997) *Evidence-based Medicine. How to Practice & Teach EBM*, p. 136. New York: Churchill Livingstone.

Seligman, M. (1995) The effectiveness of psychotherapy. *American Psychologist*, **50**, 965–974.

Shea, M., Elkin, I., Imber, S., et al (1992) Course of depressive symptoms over follow-up: findings from the National Institute of Mental Health treatment of

depression collaborative research program. *Archives of General Psychiatry*, **49**, 782–787.

Sifneos, P. (1987) *Short Term Dynamic Psychotherapy – Evaluation and Technique*. New York: Plenum.

Snyder, D. K., Willis, R. M. & Grady-Fletcher, A. (1991) Long term effectiveness of behavioral versus insight-oriented marital therapy: a 4-year follow-up study. *Journal of Consulting and Clinical Psychology*, **59**, 138–141.

Strayhorn, J. M. (1987) Control groups for psychosocial intervention outcome studies. *American Journal of Psychiatry*, **144**, 275–282.

Ward, E., King, M., Lloyd, M., et al (2000) Randomised controlled trial of non-directive counselling, cognitive behaviour therapy, and usual general practitioner care for patients with depression. I. Clinical effectiveness. *BMJ*, **321**, 1383–1388.