Psychosocial treatment for severe personality disorder

36-month follow-up

MARCO CHIESA and PETER FONAGY

Background In a previous report a step-down psychosocial programme for severe personality disorders was found to be more effective at expected termination of treatment than a longer in-patient treatment with no planned after-care.

Aims To evaluate the clinical effectiveness of these two psychosocial specialist programmes over a 3-year follow-up period.

Method Two samples allocated to the in-patient treatment and to the step-down programme were compared prospectively on symptom severity, social adjustment, global assessment of mental health and other clinical indicators at 6, 12, 24 and 36 months after intake.

Results Improvements were significantly greater in the step-down programme for social adjustment and global assessment of mental health. Patients in the programme were found to self-mutilate, attempt suicide and be readmitted significantly less at 24- and 36-month follow-up than patients in the in-patient group.

Conclusions Improvements associated with specialist residential treatment continued 2 years after discharge. A step-down model has significant advantages over a purely in-patient model.

Declaration of interest Grants from the Henry Smith and Welton Foundations.

Only a few prospective studies of short-term (<5 years) follow-up of psychodynamically based treatments for personality disorders have been carried out (Bateman & Fonagy, 2000). Yet for a chronic condition with a variable course such as personality disorder it is essential to know if treatment effects are maintained after discharge from treatment (Mehlum *et al*, 1991). This is particularly pertinent in psychodynamically informed treatments, which by addressing the underlying issues of overt behavioural disturbance aim to bring about durable and stable rehabilitative change (Howard *et al*, 1993). This claim is often made but it remains substantially untested. In a previous publication (Chiesa & Fonagy, 2000), results concerning treatment effects (12 months after intake) of specialist psychosocial intervention for personality disorder were presented. The present paper outlines the main findings of a 36-month follow-up in two cohorts of patients with personality disorder admitted to different models of psychosocial treatment at the Cassel Hospital, UK, a centre renowned for the treatment of severe personality disorder. The study started in January 1993 and recruitment ended in July 1997.

METHOD

Aims, treatment allocation and hypothesis

The clinical effectiveness of a long-term purely in-patient model of treatment with no planned out-patient follow-up treatment (one-stage model) is compared with a combined medium-stay in-patient and out-patient model (step-down model). Patients residing outside the Greater London area (who are unable to attend the out-patient phase of the step-down programme) are admitted for psychosocial treatment at the Cassel Hospital for approximately 1 year; after-care follow-up is passed back to the local referring agencies. Patients admitted from within the Greater London area are allocated to the step-down programme, which consists of a medium-term in-patient stay (6 months) followed by 12–18 months of out-patient group psychotherapy and 6 months of concurrent community outreach nursing, both under the auspices of the Cassel Hospital. Our original hypothesis was that the phased and longer-term step-down model was more effective than the one-stage model for patients with severe personality disorder (Chiesa & Fonagy, 2000).

Clinical programmes

The ingredients of psychosocial residential treatment are based on a combination of a sociotherapeutic programme (daily unit meetings, community meetings, structured activities, co-responsibility planning of the running of the therapeutic community, dance therapy, etc.) and formal psychoanalytical psychotherapy (individual and in small groups) delivered by medical and non-medical psychotherapists.

In the outreach stage of the step-down programme patients are offered twice-weekly small-group psychotherapy, once-weekly meetings with the community outreach nurse and regular review with the consultant psychiatrist. Patients are actively supported in networking with other agencies within their community setting.

Subjects and design

A total of 135 patients consecutively admitted to the Cassel Hospital adult unit between 1993 and 1997 were screened for inclusion (age 19–55 years, IQ above 90 and meeting the diagnostic criteria for at least one personality disorder) and exclusion criteria (diagnosis of schizophrenia or paranoid psychosis or evidence of organic brain disorder). Two patients were excluded from the study on account of organic brain pathology (epilepsy). Of the 74 patients allocated to the one-stage condition, 28 either did not give consent to the research participation or dropped out after completing the baseline battery only. Among the 59 patients allocated to the step-down programme, 15 either did not consent or failed to comply after completing baseline measures. Between 12 and 36 months, only six and four patients dropped out of the study in the one-stage and step-down programme, respectively. This left 80 patients (40 in each sample) for whom results are outlined and discussed.

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press

Four patients who gave consent for the study and were allocated to the one-stage group subsequently died by suicide: two were excluded from the analysis because they did not complete the baseline assessment; the other two killed themselves by 6 and 12 months after admission, respectively, and were included in the multivariate analysis after we carried over the values of the standardised measures from the last assessment point but were excluded from the analysis of clinical outcome variables. The mean duration of in-patient treatment was 8.1 months (s.d.=4.4, median=9.8) for the one-stage group and 6.2 months (s.d.=1.7, median=6.4) for patients in the step-down group. Average duration of out-patient continuation treatment for step-down patients was 10.6 months (s.d.=6.9, median=13.4). Twenty-three (57.5%) of the patients in the one-stage programme dropped out of treatment before the expected termination date, whereas in the step-down group eight (20%) patients prematurely terminated the in-patient stage of treatment and a further eight (20.0%) prematurely left the out-patient group. Because we adopted an intent-to-treat design, all subjects were recalled or traced for assessments.

Measures

Forty-eight socio-demographic and clinical variables were collected at intake. The National Adult Reading Test (Nelson, 1982), which consists of a list of 50 words printed in order of increasing difficulty, was used to obtain IQ equivalents. The Structured Clinical Interview for the DSM–III–R, Version 1.0 (Spitzer *et al*, 1990) yielded full diagnostic Axis I and II profiles. Interrater reliability testing with a second researcher reviewing taped interviews showed that κ values for each Axis I diagnosis yielded a median of 0.85 (range 0.73–1.00). On Axis II, reliability of diagnosis varied between 0.61 for cluster A, 0.67 for cluster B and 1.00 for cluster C.

The Symptom Checklist–90–R (SCL–90–R; Derogatis, 1983), a four-point self-report clinical rating scale, identifies symptoms in nine major areas of the patient's psychosomatic and interpersonal functioning. The SCL–90–R General Severity Index (GSI) was the total score used in the study to report changes in degree of symptomatic distress. The interviewer-based version of the Social Adjustment Scale (SAS; Weissman, 1975) yields ratings on a four-point scale of adjustment in the areas of work, family of origin, marriage, gender and social leisure. A total social adjustment score is derived from the mean values of the subcategories. An interclass correlation coefficient of 0.78 obtained for the total score showed satisfactory inter-rater agreement. The Global Assessment Scale (GAS; Endicott *et al*, 1976) is an anchored rating scale that allows evaluation of patient's general outcome in accordance with his or her level of functioning assessed during a specified time period (4 weeks in the present study). The rating is on a continuous scale from 0 (successful suicide) to 100 (perfect functioning). Good interrater reliability was found (interclass correlation coefficient=0.79).

A team of research psychologists and psychiatrists independent of the clinical teams was trained in the reliability criteria on all measures through the use of original training videotapes. Each rater had a fortnightly supervision meeting with a senior psychiatrist experienced in the delivery of the instruments used in the study. Coding and data entry were monitored regularly and adherence to protocol was checked using audiotapes and physical records. Outcome measures were applied longitudinally at intake and at 6, 12, 24 and 36 months.

Details of self-mutilation episodes, suicide attempts and number and length of psychiatric in-patient episodes over the year prior to the assessment were obtained through a structured interview applied at intake and at 12, 24 and 36 months. The conservative data gathering and coding protocol described by Bateman & Fonagy (1999) was used to define incidence of self-mutilation, parasuicide and readmission to hospital. A random sample of the interviews was cross-checked against the records of the patients' general practitioners and a second sample was subjected to test–retest reliability checks.

Data analysis

All analyses were performed using the Statistical Package for the Social Sciences, version 11 (SPSS, 2001). Three separate analyses taking a multivariate approach to repeated-measures analysis of variance were used to test the significance of changes in mean scores on the three standardised measures (GAS, GSI and SAS) and the differences between the treatment conditions. Estimated marginal means for the two groups were contrasted at each follow-up point, with α levels adjusted for the number of significance tests performed using Bonferroni corrections.

Improvement was also examined as a categorical variable. A reliable change index[1] was calculated for all three variables using the formula provided by Jacobson *et al* and later amended by Christensen & Mendoza (1986). Kendall's τ_b test was used to assess the significance of differences in reliable improvement and deterioration rates between the two samples. Patients also were allocated to improved (if they showed reliable change on at least two measures with no concomitant deterioration on the third measure) and non-improved categories.

Differences between groups on dichotomised clinical variables (self-mutilation, suicide attempts and hospital readmissions) at 12, 24 and 36 months were examined using separate hierarchical logistic regressions, with group membership as predictor and baseline status on each clinical variable as covariate.

RESULTS

The demographic and clinical features of the sample, more extensively reported previously (Chiesa & Fonagy, 2000), are displayed in Table 1. In addition to the variables listed in the table the two groups were compared on variables known to influence the outcome, including duration of illness, past history of treatment and service utilisation, comorbidities (including antisocial behaviour), employment history and living arrangements at time of admission. No significant differences were found in any of the baseline variables between the two groups. The means and standard deviations of three standardised measures (GAS, GSI and SAS total scores) at intake and at 6, 12, 24 and 36 months for the two groups are displayed in Table 2.

Severity of psychiatric symptoms

The repeated-measures analysis of variance showed a significant effect of time (Wilks' $\lambda=0.51$, $F=17.98$, d.f.=4,75, $P<0.001$) for the GSI. At 36 months, 11 (27.5%) of

1. Reliable change takes into account measurement error according to the formula x_2-x_1/S_{diff}, where x_1 is the pre-test score, x_2 is the post-test score and S_{diff} is the standard error of difference between the two test scores. A reliable change index value below 1.96 represents a change ascribable to the imprecision of the measurement.

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press

the one-stage group and 19 (47.5%) of the step-down group were below the cut-off point for caseness (Derogatis, 1983). However, no significant group × time interaction (Wilks' λ=0.92, F=1.56, d.f.=4,75, NS) was found, indicating that the means displayed in Table 2 are not significantly different for the two groups.

Both one-stage and step-down groups achieved relatively high rates of reliable change by 6 months (38% and 45%, respectively), with progressive increases through to the 36-month follow-up; at that assessment point, 80% and 70% of the subjects showed reliable improvement (Table 3). No significant differences between the two groups were found overall.

Social adjustment

A significant time effect (Wilks' λ=0.61, F=11.93, d.f.=4,75, $P<0.001$) was revealed, with amelioration in the degree of social adaptation over time in both samples reflected in the significant linear effects (F=44.20, d.f.=4,75, $P<0.001$). Pairwise comparisons showed significant differences between the two groups both at 24 (P=0.04) and 36 months (P=0.04) in favour of the step-down model. However, no significant group × time interaction was found (Wilks' λ=0.98, F=0.31, d.f.=1,78, NS).

The two samples did not differ significantly in their respective rates of reliable change at any of the four assessment points (Table 3). Improvement in social adaptation appeared to be slower than in the psychiatric symptoms dimension. However, 40% of the subjects in both groups were found to have improved reliably by 24 months. This improvement declined at 36 months to 35% and 38% in the one-stage and step-down condition, respectively.

Global assessment of functioning

Changes in GAS mean scores over the five assessment points showed a significant effect of time (Wilks' λ=0.57, F=14.12, d.f.=4,75, $P<0.001$). Although the group × time interaction was not significant (Wilks' λ=0.96, F=1.18, d.f.=4,75, NS), greater improvement was shown by the step-down group relative to the one-stage group at 12 months (P=0.02) and this was marginally significant at 36 months ($P<0.06$) on pairwise comparisons.

Significant differences in rates of reliable change on the GAS between the two groups were found at 12 months (Kendall's τ_b=0.34, d.f.=2, $P<0.001$), 24 months (Kendall's τ_b=0.27, d.f.=2, $P<0.02$) and 36 months (Kendall's τ_b=0.37, d.f.=2, $P<0.001$). Patients in the step-down group were more likely to meet the reliable improvement criteria for GAS than patients in the one-stage condition (Table 3).

Table 1 Comparison of socio-demographic characteristics of the treatment samples

Variable	One-stage group (n=40)	Step-down group (n=40)
Age (years), mean (s.d.)	31.5 (7.3)	32.9 (8.7)
National Adult Reading Test, mean (s.d.)	109.2 (26.7)	111.6 (19.6)
Female, *n* (%)	31 (77.5)	30 (75.0)
Single, divorced, widowed or separated, *n* (%)	32 (80.0)	34 (85.0)
Employed, *n* (%)	5 (12.5)	6 (15.0)
College education, *n* (%)	31 (77.5)	29 (72.5)
Living alone, *n* (%)	19 (48.7)	26 (65.0)
Self-reported sexual abuse, *n* (%)	19 (47.5)	19 (47.5)
Self-reported physical abuse by carers, *n* (%)	15 (37.5)	15 (37.5)
Early loss, *n* (%)	20 (50.0)	21 (52.5)
Axis I DSM–III–R diagnosis, *n* (%)		
Major depression	12 (30.0)	12 (30.0)
Dysthymia	4 (10.0)	7 (17.5)
Phobic disorders	19 (47.5)	16 (40.0)
Other anxiety disorders	13 (32.5)	17 (38.6)
Eating disorders	8 (20.0)	7 (17.5)
Substance use disorders	21 (52.5)	20 (50.0)
Axis II DSM–III–R diagnosis, *n* (%)		
Paranoid	22 (55.0)	17 (42.5)
Schizotypal	7 (17.5)	7 (17.5)
Antisocial	1 (2.5)	2 (5.0)
Borderline	30 (75.0)	25 (62.5)
Dependent	12 (40.0)	16 (40.0)
Avoidant	23 (57.5)	22 (55.0)
PDNOS (self-defeating, passive–aggressive)	20 (50.0)	23 (57.5)

PDNOS, personality disorder not otherwise specified.

Overall improvement

By 12 months, 33% of patients in the step-down model *v.* 13% of the one-stage group were reliably improved on two out of the three standardised measures (GSI, SAS and GAS); the difference was significant (χ^2=4.59, d.f.=1, P=0.03). However, the difference was no longer significant at 24 (45% *v.* 35%) and 36 months (50% *v.* 38%).

Clinical change

Table 4 summarises figures for the two groups with regard to clinical variables. We found a marked decrease in the number of patients in the step-down programme who committed acts of self-mutilation by 12, 24 and 36 months. In contrast, after an increase at 12 months, the number of patients allocated to the one-stage programme who carried out self-mutilation remained constant at 24 and 36 months. Hierarchical logistic regression showed that belonging to the one-stage group was predictive of self-mutilation at 12 (B=2.28, s.e.=0.68, d.f.=1, $P<0.001$), 24 (B=1.38, s.e.=0.57, d.f.=1, $P<0.02$) and 36 (B=1.14, s.e.=0.58, d.f.=1, $P<0.05$) months. The odds ratio revealed that patients allocated to the one-stage programme were ten (95% CI 2.58–36.96), four (95% CI 1.29–12.22) and three (95% CI 1.01–9.79) times more likely to mutilate themselves at 12, 24 and 36 months, respectively, than patients in the step-down condition.

Group status also predicted attempted suicide, with the one-stage membership more likely to attempt suicide at 12 (B=2.03, s.e.=0.63, $P<0.002$) and 36

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press

Table 2 Means, standard deviations, and confidence intervals for GSI, SAS and GAS in the two samples

Variable	One-stage group (n=40) Mean (s.d.)	95% CI	Step-down group (n=40) Mean (s.d.)	95% CI
GSI				
Intake	2.08 (0.63)	1.86–2.32	1.83 (0.82)	1.60–2.06
6 months	1.85 (0.57)	1.63–2.08	1.56 (0.83)	1.34–1.78
12 months	1.65 (0.69)	1.40–1.91	1.46 (0.92)	1.20–1.71
24 months	1.66 (0.77)	1.38–1.94	1.23 (0.98)	0.95–1.51
36 months	1.38 (0.68)	1.14–1.62	1.12 (0.77)	0.89–1.36
SAS				
Intake	2.73 (0.48)	2.57–2.90	2.58 (0.56)	2.42–2.75
6 months	2.56 (0.34)	2.41–2.64	2.39 (0.47)	2.31–2.54
12 months	2.47 (0.44)	2.28–2.59	2.22 (0.59)	2.11–2.41
24 months	2.32 (0.40)	2.14–2.49	2.06 (0.55)	1.93–2.23
36 months	2.33 (0.53)	2.14–2.49	2.04 (0.60)	1.87–2.23
GAS				
Intake	46.20 (7.67)	43.97–48.43	46.53 (6.44)	44.29–48.76
6 months	49.33 (8.00)	46.61–52.18	53.16 (10.24)	50.31–55.89
12 months	50.71 (10.2)	47.14–54.47	57.19 (14.11)	53.42–60.76
24 months	54.23 (13.01)	50.37–58.33	59.25 (14.25)	55.15–63.11
36 months	53.27 (12.82)	49.72–57.08	58.56 (13.14)	54.75–62.11

GSI, Symptom Checklist–90–R General Severity Index; SAS, Social Adjustment Scale; GAS, Global Assessment Scale.

(B=1.74, s.e.=0.72, d.f.=1, P<0.02) months; the odds ratio showed that patients in the one-stage programme were eight (95% CI 2.22–26.11) and six (95% CI 1.40–23.15) times more likely to attempt suicide than step-down patients.

Significant group differences were found in psychiatric readmission rates. By the 24- and 36-month follow-ups, patients treated in the step-down condition were less likely to be readmitted to psychiatric units than patients in the one-stage programme. Logistic regression showed that the one-stage group was significantly different from the other group at 24 (B=2.09, s.e.=0.64, d.f.=1, P<0.002) and 36 (B=1.69, s.e.=0.60, d.f.=1, P<0.005) months. One-stage patients were eight (95% CI 2.29–28.25) and five (95% CI 1.69–17.55) times more likely to have at least one acute admission to a psychiatric unit at 24 and 36 months, respectively, than patients who were treated with the step-down model.

DISCUSSION

Improvement and specialist programme

This study shows that significant improvement is maintained 2 years after expected discharge in a group of patients with severe personality disorder treated in a specialist residential setting in the dimensions of symptom distress, social adjustment and global assessment of outcome. Improvements either continue to occur or become stable by the 24- and 36-month follow-ups, depending on the measure. It is notable that these substantial gains in the three key outcome measures were obtained in a relatively short period of time (3 years after admission), in contrast to early studies that reported a lack of short-term improvements in personality disorder (McGlashan, 1986; Stone, 1990). Our results are consistent with a number of more recent studies showing that psychosocial programmes specific for personality disorder applied in different settings may be demonstrably effective in the treatment of personality disorder (Najavitis & Gunderson, 1995; Dolan *et al*, 1997; Bateman & Fonagy, 2001). Improvement of the prognostic outlook of this condition, still widely considered as difficult to treat, may in turn contribute to lessening the stigma attached to this diagnostic entity (Holmes *et al*, 2001).

Reliable change

Improvement in the two samples cannot be ascribable to the unreliability of the measures, but rather is a reflection of true improvement that met the stringent reliable change index criteria. Although change was more marked in psychiatric symptoms (75% of patients reliably improved by 36 months), improvement in the area of social adjustment stabilised by 24 months (40% of patients reliably improved); improvement in global assessment of mental health continued through to 24 months, after which a slight decrease at 36 months occurred in both samples.

The relatively lower rate of reliable improvement in social adjustment compared with the higher symptom severity improvement in the samples leads to a number of considerations, because this dimension represents the main target of therapeutic community work:

(a) it confirms that social adaptation is a more complex dimension than symptom severity and hence is less easy to modify than the symptoms;

(b) impairment in social behaviour and social functioning is intrinsically linked to the concept of personality disorder, with the notion of durability and persistence therefore being notoriously slower to achieve even if treatment was effective (Gunn, 2000);

(c) social adjustment sub-dimensions such as work record may be heavily affected by environmental and social conditions (i.e. the state of the economy) and therefore less dependent on individual psychopathology;

(d) instruments used to assess social adjustment were devised for different clinical presentations (i.e. depression) and some sub-dimensions, such as one's own family or domestic functioning, may not easily apply to a sample with borderline personality disorder.

Outcome differences and treatment models

Although patients treated according to both models maintained improvements from discharge to follow-up, on a number of measures the improvements observed in the step-down group are more substantial than in the one-stage group. Thus, the differences observed at discharge reported in our previous communication are generally maintained at 36 months. This is particularly notable for clinical variables. Patients exposed to shorter residential stay followed by long-term psychosocial therapy in the external community showed significantly greater improvement in self-mutilation,

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press

Table 3 Reliable change in the two samples

Variable	One-stage group (n=40) n (%)	Step-down group (n=40) n (%)	Kendall's τ_b
GSI at 6 months			
Improved	15 (37.5)	18 (45.0)	0.019
Deteriorated	4 (10.0)	7 (17.5)	
GSI at 12 months			
Improved	19 (47.5)	22 (55.0)	−0.007
Deteriorated	2 (5.0)	8 (20.0)	
GSI at 24 months			
Improved	23 (57.5)	25 (62.5)	0.049
Deteriorated	7 (17.5)	6 (13.6)	
GSI at 36 months			
Improved	32 (80.0)	28 (70.0)	−0.092
Deteriorated	5 (12.5)	4 (10.0)	
SAS at 6 months			
Improved	6 (15.0)	6 (15.0)	−0.027
Deteriorated	1 (2.5)	2 (5.0)	
SAS at 12 months			
Improved	6 (15.0)	15 (37.5)	0.172
Deteriorated	0 (0)	3 (7.5)	
SAS at 24 months			
Improved	16 (40.0)	16 (40.0)	0.000
Deteriorated	1 (2.5)	1 (2.5)	
SAS at 36 months			
Improved	14 (35.0)	15 (37.5)	0.025
Deteriorated	1 (2.5)	1 (2.5)	
GAS at 6 months			
Improved	6 (15.0)	7 (17.5)	0.005
Deteriorated	0 (0)	1 (2.5)	
GAS at 12 months			
Improved	5 (12.5)	17 (42.5)	0.344*
Deteriorated	1 (2.5)	0 (0.0)	
GAS at 24 months			
Improved	9 (22.5)	19 (47.5)	0.271†
Deteriorated	1 (2.5)	0 (0.0)	
GAS at 36 months			
Improved	7 (17.5)	21 (52.5)	0.367*
Deteriorated	0 (0.0)	0 (0.0)	

GSI, Symptom Checklist–90–R General Severity Index; SAS, Social Adjustment Scale; GAS, Global Assessment Scale. *$P<0.01$; †$P<0.02$.

attempted suicide and readmission to hospital than patients allocated to the long-term residential treatment with no planned after-care. The considerable difference found in these variables between the two groups suggests that follow-up psychotherapy treatment after a period of hospitalisation is important to the stabilisation of these patients back into their community, and reduces the number of cases of acute distress or breakdown that may lead to acute psychiatric admission. The planned post-discharge ongoing support provided by the outreach psychosocial team engenders a sense of belonging and being contained that compensates for the deep-rooted sense of aloneness typical of these patients (Gunderson, 1996).

The absence of supportive after-care in the one-stage group seems to undermine the continuation of the healing process initiated while in in-patient treatment. In contrast, the provision of a phased long-term (albeit less intensive) treatment seems more effective in meeting the severe anxieties connected with relational and socialising difficulties presented by personality disorder than a purely residential approach. These results are consistent with the rationale of the step-down programme as enhancing patients' resilience to psychosocial stressors and maintaining satisfactory functioning in the community. In contrast, although improved on some dimensions, the one-stage group did not show any significant improvement in self-mutilation, parasuicide and acute re-admission compared with the year prior to admission to the Cassel Hospital, which is a sign of continuing vulnerability to acute decompensations.

Improvement shown by the step-down group compares well with that found in a group of patients with borderline personality disorder treated in a psychotherapeutic partial hospitalisation programme (Bateman & Fonagy, 2001) on similar dimensions. Psychosocial residential programmes have been questioned as realistic treatment options for personality disorder, because third-party payers regard the cost associated with in-patient admission as excessive. In a previous study we investigated if the cost of in-patient admission relative to that of treatment as usual reduces health and social care cost in the year after treatment termination (Chiesa *et al*, 2002). The results showed that the two specialist programmes might be justifiable in economic terms, because savings in service use would offset the cost of delivering treatment in the long run.

Methodological considerations

The non-randomised allocation of subjects to the two treatment programmes and the absence of an untreated or treatment-as-usual control condition limit claims concerning the effectiveness of the two specialist models, because possible geographical factors may have accounted for some or all of the observed differences. In the event, a remarkable similarity between the two groups was found in terms of demographic, diagnostic and other clinical variables: no significant difference was found in any of the 37 variables (including severity of presentation) that were measured. It is likely that referral for the same specialist treatment created moderate differences that would be expected from the geographical criteria used for patients' allocation to treatment condition.

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press

Although over 80% of patients met the criteria for at least one diagnosis belonging to the dramatic personality disorder cluster (or cluster B), thus making it a relatively homogeneous cohort, comorbidity with other Axis I and Axis II diagnoses was invariably present in both samples. The presence of multiple diagnoses raises the question of the impact of comorbidity as a moderator of outcome. In addition, the crucial issue of outcome for different categories of individual personality disorder (Tyrer & Seivewright, 2000) was not addressed here. However, this study was not powered for subgroup analyses and disregarding the risk of type II errors would compromise the analysis of data.

The study did not address the dimension of treatment process and leaves unanswered the question of what aspect of the intervention in the two samples was the effective component in bringing about improvement. It remains unclear whether the residential aspect of psychosocial treatment is in fact a necessary condition to effectiveness for this group of patients. These outcome results have to take into account recent studies demonstrating that day hospital (Bateman & Fonagy, 2001) and specialist out-patient programmes (Clarkin *et al*, 2001) are a promising and effective alternative to hospitalisation for borderline personality disorder.

The incremental loss of subjects to the 36-month follow-up also limits the inferences that can be drawn about treatment effectiveness, and reduces the power and sensitivity of statistical analysis to detect differences between the two groups in other variables. Although we tried to limit the impact of clinical drop-out by employing an intent-to-treat analysis, the extent of attrition over the course of the study raises the legitimate question of its interaction with the treatment in the results that were produced (Kazdin, 1994).

An additional potential source of bias was the lack of independent rater blindness with regard to subject treatment allocation.

CLINICAL IMPLICATIONS

- Marked and reliable improvement in symptomatology, social adjustment and global assessment of mental health over a 36-month follow-up period is associated with specialist psychosocial treatment for personality disorder.
- Improvements continue after discharge: a proportion of patients showed stable and durable change 2 years after termination of treatment.
- A phased programme that includes a community-based stage of treatment was found to yield more stable improvement than the purely hospital-based programme, as shown by the greater reduction in self-mutilation, attemps at suicide and readmission rates.

LIMITATIONS

- The absence of a control group limits claims concerning efficacy.
- Attrition rates may be a threat to internal and external validity.
- Generalisation to other settings is limited.

MARCO CHIESA, MRCPsych, The Cassel Hospital, Richmond, UK; PETER FONAGY, FBA, Sub-department of Clinical Health Psychology, University College London, UK

Marco Chiesa, Consultant Psychiatrist and Head of Research Unit, The Cassel Hospital, Richmond TW10 7JF, UK. Tel: 020 8237 2902; fax: 020 8332 6424; e-mail: m.chiesa@ucl.ac.uk

(First received 27 November 2002, final revision 12 May 2003, accepted 14 May 2003)

Table 4 Clinical outcomes in the one-stage and step-down models

Variable	One-stage (n=38) n (%)	Step-down (n=40) n (%)
Self-mutilation, yes		
Intake	17 (44.7)	20 (50.0)
12 months	25 (65.8)	12 (30.0)*
24 months	18 (47.4)	10 (25.0)**
36 months	15 (39.5)	9 (22.5)**
Parasuicide, yes		
Intake	17 (44.7)	21 (52.5)
12 months	17 (44.7)	5 (12.5)*
24 months	11 (28.9)	7 (17.5)
36 months	11 (28.9)	3 (7.5)**
Hospital admission, yes		
Intake	19 (50.0)	22 (55.0)
24 months	17 (44.7)	5 (12.5)*
36 months	17 (44.7)	7 (17.5)*

*P < 0.01; **P < 0.05.

Strength of the study

Despite these limitations we feel that our design addressed some of the methodological weaknesses present in previous studies of personality disorder. The use of operational diagnostic criteria, a full characterisation of the sample in terms of demographic and clinical features, the use of standardised outcome measures, the prospective nature of the investigation, the use of trained independent raters and the presence of interrater reliability checks strengthen the internal validity of the study. The adoption of stringent and conservative criteria of improvement based on reliable change and the use of multivariate and regression statistics on all measures

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press

improve the validity of the outcome results found in the samples.

ACKNOWLEDGEMENTS

We thank Carla Drahorad, Eduardo Iacoponi, Alice Levinson and Ilan Diamant for help with data collection.

REFERENCES

Bateman, A. & Fonagy, P. (1999) Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *American Journal of Psychiatry*, **156**, 1563–1569.

__ & __ (2000) Effectiveness of psychotherapeutic treatment of personality disorder. *British Journal of Psychiatry*, **177**, 138–143.

__ & __ (2001) Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization; an 18-month follow-up. *American Journal of Psychiatry*, **158**, 36–42.

Chiesa, M. & Fonagy, P. (2000) Cassel Personality Disorder Study. Methodology and treatment effects. *British Journal of Psychiatry*, **176**, 485–491.

__, Fonagy, P., Holmes, J., *et al* (2002) Health service use costs by personality disorder following specialist and non-specialist treatment: a comparative study. *Journal of Personality Disorders*, **16**, 160–173.

Christensen, L. & Mendoza, J. L. (1986) A method of assessing change in a single subject: an alteration of the RC index. *Behavior Therapy*, **17**, 305–308.

Clarkin, J. F., Foelsch, P. A., Levy, K. N., *et al* (2001) The development of a psychodynamic treatment for patients with borderline personality disorder: a preliminary study of behavioral change. *Journal of Personality Disorders*, **15**, 487–495.

Derogatis, L. R. (1983) *SCL–90–R. Administration, Scoring and Procedures. Manual–II.* Towson, MD: Clinical Psychometric Research.

Dolan, B., Warren, F. & Norton, K. (1997) Change in borderline symptoms one year after therapeutic community treatment for severe personality disorder. *British Journal of Psychiatry*, **171**, 274–279.

Endicott, J., Spitzer, R. L., Fleiss, J. L., *et al* (1976) The Global Assessment Scale. *Archives of General Psychiatry*, **33**, 766–771.

Gunderson, J. G. (1996) The borderline patient intolerance of aloneness: insecure attachment and therapist availability. *American Journal of Psychiatry*, **153**, 752–758.

Gunn, J. (2000) Personality disorder: a clinical suggestion. In *Personality Disorders, Diagnosis, Management and Course* (ed. P. Tyrer), pp. 44–50. Oxford: Butterworth Heinemann.

Holmes, J., Montgommery, C., Chiesa, M., *et al* (2001) Personality disorder and quality of life. In *Every Family in the Land* (ed. A. Crisp). http://www.stigma.org/everyfamily/.

Howard, K. I., Lueger, R. J., Maling, M. S., *et al* (1993) A phase model of psychotherapy outcome: causal mediation of outcome. *Journal of Consulting and Clinical Psychology*, **61**, 678–685.

Kazdin, A. E. (1994) Methodology, design, and evaluation in psychotherapy research. In *Handbook of Psychotherapy and Behavior Change* (eds A. E. Bergin & S. L. Garfield), pp. 19–71. New York: John Wiley.

McGlashan, T. (1986) Long-term outcome of borderline patients. *Archives of General Psychiatry*, **40**, 20–30.

Mehlum, L., Friis, S., Irion, T., *et al* (1991) Personality disorders 2–5 years after treatment. A prospective follow-up study. *Acta Psychiatrica Scandinavica*, **84**, 72–77.

Najavitis, L. M. & Gunderson, J. G. (1995) Better than expected: improvements in borderline personality disorder in a 3 year prospective study. *Comprehensive Psychiatry*, **36**, 296–302.

Nelson, H. E. (1982) *National Adult Reading Test (NART) – Test Manual.* Windsor: NFER–Nelson.

Spitzer, R. L., Williams, J., Gibbon, M., *et al* (1990) *Structured Clinical Interview for DSM–III–R.* Washington, DC: American Psychiatric Press.

SPSS (2001) *SPSS for Windows.* Chicago, IL: SPSS Inc.

Stone, M. H. (1990) *The Fate of Borderline Patients: Successful Outcome and Psychiatric Practice.* New York: Guilford Press.

Tyrer, P. & Seivewright, H. (2000) Outcome of personality disorder. In *Personality Disorders, Diagnosis, Management and Course* (ed. P. Tyrer), pp. 105–125. Oxford: Butterworth Heinemann.

Weissman, M. M. (1975) The assessment of social adjustment. *Archives of General Psychiatry*, **32**, 357–365.

https://doi.org/10.1192/bjp.183.4.356 Published online by Cambridge University Press