

Effectiveness of Different Modalities of Psychotherapeutic Treatment for Patients with Cluster C Personality Disorders: Results of a Large Prospective Multicentre Study

Anna Bartak^{a, c} Marieke D. Spreeuwenberg^{e, f} Helene Andrea^{a, g}
Lot Holleman^h Piet Rijnierse^d Bert V. Rossumⁱ Elisabeth F.M. Hamers^b
Anke M.M.A. Meerman^j Janneke Aerts^k Jan J.V. Busschbach^{a, g} Roel Verheul^{b, c}
Theo Stijnen^l Paul M.G. Emmelkamp^c

^aViersprong Institute for Studies on Personality Disorders (VISPD), and ^bCentre of Psychotherapy De Viersprong, Halsteren; ^cDepartment of Clinical Psychology, University of Amsterdam, and ^dCentre of Psychotherapy Mentrum, Amsterdam; ^eDepartment of Methodology and Statistics, Tilburg University, Tilburg; ^fVilans, Hoensbroek; ^gDepartment of Medical Psychology and Psychotherapy, Erasmus Medical Centre, Rotterdam; ^hZaans Medical Centre, Zaandam; ⁱAltrecht, Zeist; ^jCentre of Psychotherapy De Gelderse Roos, Lunteren; ^kGGZWNB, Bergen op Zoom; ^lDepartment of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands

Key Words

Personality disorders · Psychotherapy · Treatment effectiveness · Treatment outcome · Compulsive personality disorder · Dependent personality disorder · Avoidant personality disorder

Abstract

Background: No previous studies have compared the effectiveness of different modalities of psychotherapeutic treatment, as defined by different settings and durations, for patients with cluster C personality disorders. The aim of this multicentre study was to compare the effectiveness of 5 treatment modalities for patients with cluster C personality disorders in terms of psychiatric symptoms, psychosocial functioning, and quality of life. The following treatment modalities were compared: long-term outpatient (more than 6 months), short-term day hospital (up to 6 months), long-term day hospital, short-term inpatient, and long-term inpa-

tient psychotherapy. **Methods:** The study was conducted between March 2003 and June 2008 in 6 mental health care centres in the Netherlands, with a sample of 371 patients with a DSM-IV-TR axis-II cluster C diagnosis. Patients were assigned to 5 different modalities of psychotherapeutic treatment, and effectiveness was assessed at 12 months after baseline. An intention-to-treat analysis was conducted for psychiatric symptoms (Brief Symptom Inventory), psychosocial functioning (Outcome Questionnaire-45), and quality of life (EQ-5D), using multilevel statistical modelling. As the study was non-randomised, the propensity score method was used to control for initial differences. **Results:** Patients in all treatment groups had improved on all outcomes 12 months after baseline. Patients receiving short-term inpatient treatment showed more improvement than patients receiving other treatment modalities. **Conclusions:** Psychotherapeutic treatment, especially in the short-term inpatient modality, is an effective treatment for patients with cluster C personality disorders.

Copyright © 2009 S. Karger AG, Basel

Introduction

An estimated 2.6% of the general population is affected by cluster C personality disorders (PD): avoidant, dependent, and obsessive-compulsive PD [1]. This cluster of PD is associated with significant functional impairment [2–4] and a high economic burden [5], yet studies investigating treatment effectiveness in this patient population are scarce [6]. As in research on other psychological disorders [e.g. 7], the available studies on cluster C PD typically compare treatments that are identical in treatment setting and duration. Investigators have compared different outpatient treatments [8–13], different day hospital treatments [14, 15], and different inpatient treatments [16, 17]. One recent study in Norway [18] compared outpatient and day hospital treatment for patients with all forms of PD, and found no significant superiority of one treatment over another at 8 months after the start of treatment. However, so far, no study has compared the effectiveness of treatments across widely differing settings and durations. In this article, treatment modality was specified as a combination of treatment setting (i.e. outpatient, day hospital, or inpatient) and duration (i.e. short term or long term), as these are the most important aspects regarding treatment costs, a crucial aspect in times of restricted health care budgets.

It is likely that one of the reasons this comparison has not been undertaken previously is the difficulty of random assignment to different treatment modalities in clinical samples due to practical or ethical constraints [19]. Furthermore, even if researchers were successful in setting up and starting a randomised treatment modality study, its external validity would be doubtful because a high number of patients would refuse to participate [e.g. 20]. Therefore, quasi-experimental studies using statistical correction models to counter selection bias are increasingly being found in the literature [e.g. 21–24].

The aim of the present quasi-experimental study was to compare the effectiveness of different treatment modalities for patients with cluster C PD in a naturalistic setting, thereby insuring high external validity. In fact, treatment modality might be an overlooked factor in psychotherapy effectiveness research.

Methods

Participants

Participants ($n = 371$) were recruited from consecutive admissions to 6 mental health care centres in the Netherlands (Centre of Psychotherapy De Viersprong, Halsteren; Altrecht, Utrecht;

Zaans Medical Centre, Zaandam; Centre of Psychotherapy De Gelderse Roos, Lunteren; GGZWNB, Bergen op Zoom & Roosendaal; Centre of Psychotherapy Mentrum, Amsterdam). These institutions offer outpatient, day hospital, and/or inpatient psychotherapeutic treatment for patients with personality pathology. From March 2003 to March 2006, 1,379 patients completed the intake procedure and were selected for treatment (fig. 1).

Of these, 146 patients (10.6%) were excluded from the study for not meeting one of the following inclusion criteria: age between 18 and 70 ($n = 13$), significant personality pathology ($n = 34$), and referral for psychotherapeutic treatment aimed at personality problems ($n = 99$). Nine patients (0.7%) met one of the following exclusion criteria: insufficient command of the Dutch language ($n = 6$), organic cerebral impairment ($n = 1$), mental retardation ($n = 1$), and schizophrenia ($n = 1$). This left 1,224 participants, of whom 100 (8.2%) refused to participate. Another 31 patients (2.5%) could not participate due to logistic reasons (e.g. no appointment could be made to provide informed consent), and 133 patients (10.9%) were excluded due to missing or unreliable baseline data. Thirty-eight patients (3.1%) received less than 2 treatment sessions or less than 2 days of inpatient or day hospital therapy, and were therefore excluded. The remaining 922 patients were informed about the study and its procedure, provided written informed consent, and entered the study. Of those, 466 patients (50.5%) had 1 or more cluster C PD.

In the absence of explicit guidelines for treatment assignment in PD [25, 26], the selection procedure was based on the expert opinion of clinicians who used their clinical experience combined with patient data from standardised instruments. To elucidate the criteria used for the assignment process, our research group recently conducted a study with intake clinicians from the participating treatment centres. We found evidence of substantial (implicit) consensus among clinicians concerning the criteria used for treatment decision-making. For example, focality of problems (focal or broad spectrum of problems) and ego strength were found to be related to decisions about a short or long treatment duration for a substantial number of intake clinicians [25].

Patients were assigned to 1 of 6 treatment modality groups: 18 to short-term outpatient (up to 6 months), 96 to long-term outpatient (more than 6 months), 85 to short-term day-hospital, 103 to long-term day hospital, 63 to short-term inpatient, and 101 to long-term inpatient treatment. The short-term outpatient group was excluded from the analyses for 2 reasons: (1) only a minority of patients (3.9%) were assigned to this short and low-frequency treatment modality, as could be expected in a PD patient population; (2) these patients differed significantly from patients in the other treatment groups on a high number of pre-treatment variables, indicating a dissimilar and – most importantly – a structurally less ‘sick’ patient population, incomparable with the rest of the sample. A comparison with this treatment modality would most probably also fail when trying to design a randomised trial, as short-term outpatient therapy differs most from all other modalities in terms of its relatively low impact on patients’ lives compared to other treatment modalities. In the end, 448 participants were included in the study. Follow-up data were not available for 77 patients (17.2%; patients who did not respond to any follow-up assessment or patients where follow-up measurements were not yet available). There was no difference in psychiatric symptoms at baseline between patients with follow-up data and those without (this holds true for both the comparison in the total sample and

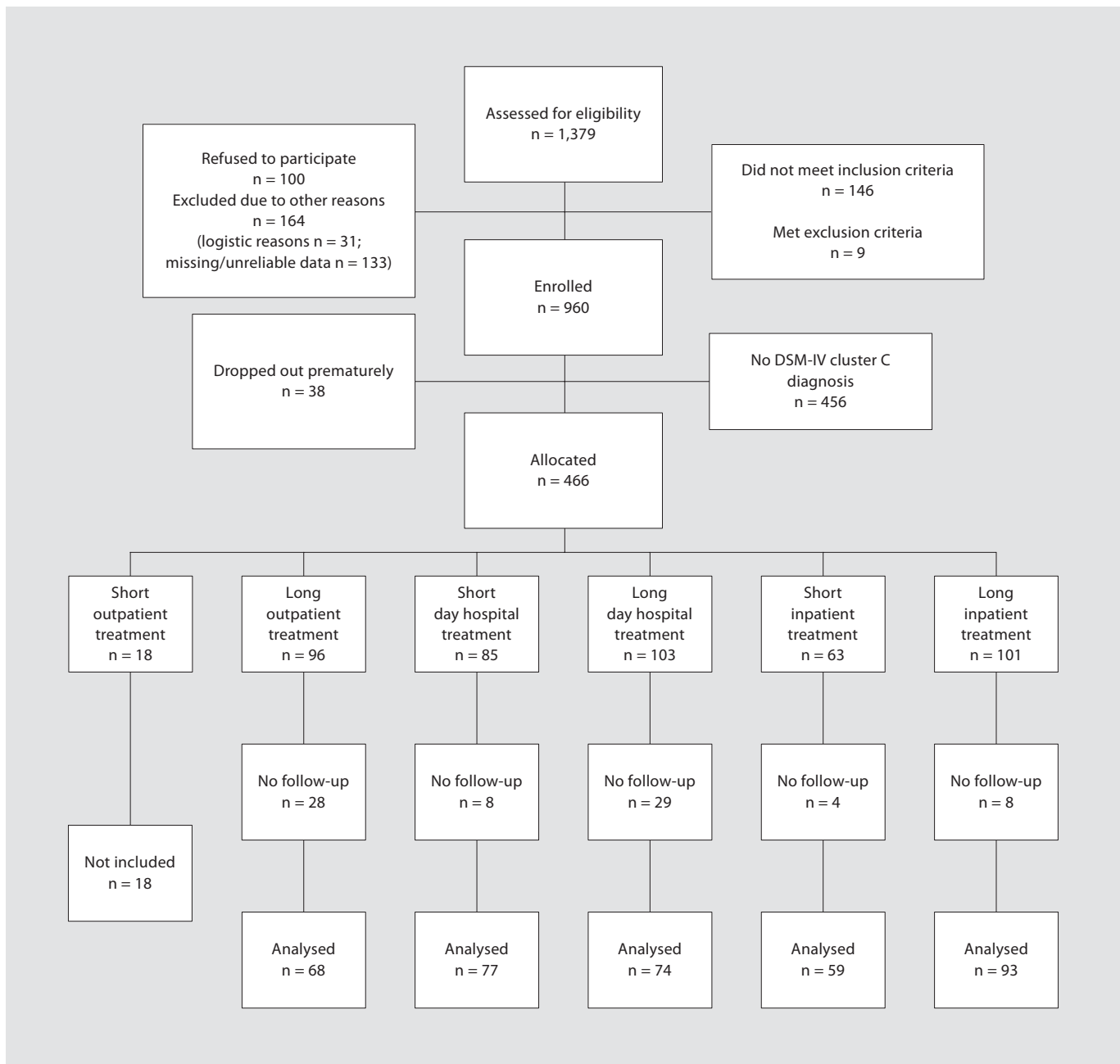


Fig. 1. Patient flow.

the comparisons within the 5 treatment groups). The final sample consisted of 371 patients to be included in the analyses.

Treatment

The 6 mental health care centres offer a variety of psychotherapeutic treatments tailored to a PD patient population. Their treatments differ according to several features. As this study focused on different treatment modalities in terms of setting and duration, the following 5 treatment groups were compared:

- Patients in long-term outpatient treatment (n = 68, 18.3% of the study sample). These patients come for individual (76.5%) or group (23.5%) psychotherapy sessions, for up to 2 sessions per week (mean 0.8 sessions/week, SD 0.51, median 0.5) for more than 6 months (mean duration 15.4 months, SD 6.36, median 12.0).
- Patients in short-term day hospital treatment (n = 77; 20.8% of the study sample). These patients come to the institutions at least 1 morning/afternoon per week (mean 3.2 days/week, SD

1.51, median 3.0) for up to 6 months (mean duration 5.4 months, SD 1.32, median 6.0) and receive different forms of psychotherapeutic and psychosocial treatment, but sleep at home.

- Patients in long-term day-hospital treatment ($n = 74$, 19.9% of the study sample). These patients come to the institutions at least 1 morning/afternoon per week (mean 3.3 days/week, SD 1.42, median 3.0) for more than 6 months (mean duration 12.1 months, SD 2.41, median 12.0) and receive different forms of psychotherapeutic and psychosocial treatment, but sleep at home.
- Patients in short-term inpatient treatment ($n = 59$, 15.9% of the study sample). These patients stay at the institutions 5 days a week for up to 6 months (mean duration 4.2 months, SD 1.48, median 3.0) and receive different forms of psychotherapeutic and psychosocial treatment.
- Patients in long-term inpatient treatment ($n = 93$, 25.1% of the study sample). These patients stay at the institutions 5 days a week for more than 6 months (mean duration 10.2 months, SD 1.98, median 10.0) and receive different forms of psychotherapeutic and psychosocial treatment.

Day hospital and inpatient programs typically consist of group psychotherapy as a core element, mostly in combination with one or more non-verbal or expressive group therapies, individual psychotherapy, sociotherapy within the therapeutic community, coaching for social problems, community meetings, and/or pharmacological treatment. The psychotherapists are all licensed psychiatrists or psychologists. On average, they had 14.9 years (SD 10.1) of postgraduate clinical experience. The treatments under study can be considered highly representative of regular clinical practice in the Netherlands, as therapists did not receive specific training for this study and treatment integrity was not monitored.

Assessments

Baseline Measures. An extensive standard assessment battery of instruments was administered to the patients before treatment assignment. PD were measured using the Dutch version of the Structured Interview for DSM-IV Personality [27, 28]. This interview covers the 11 formal DSM-IV-TR axis II diagnoses including PD not otherwise specified, 2 appendix diagnoses (i.e. depressive and negativistic PD), and self-defeating PD. Interviewers were master's level psychologists, who were trained thoroughly by one of the authors (R.V.), and who received monthly booster sessions to avoid deviation from the interviewer guidelines. Inter-rater reliability was evaluated in 25 video-taped interviews, which were rated by 3 observer-raters. Percentage of agreement between observer-raters ranged from 84 (avoidant PD) to 100% (schizoid) (median 95%). Intraclass correlation coefficients for the sum of DSM-IV PD traits present (i.e. scores '2' or '3') ranged from 0.60 (schizotypal) through 0.92 (antisocial) (median 0.74). To measure patient characteristics at baseline, the assessment battery also included 3 self-report instruments. The first of those was the Dutch version of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ), for measuring the type and degree of personality pathology [29, 30]. We used patients' scores on this questionnaire for the 4 higher-order factors: emotional dysregulation, dissociative behaviour, inhibition, and compulsivity. To measure the severity of personality pathology we used the 5 higher-order domains of the Severity Indices of Personality Problems

(SIPP): self-control, social concordance, identity integration, relational capacities, and responsibility [31]. To measure patients' motivation for treatment, we used the 2 scales of the Motivation for Treatment Questionnaire (MTQ-8): need for help and readiness to change [32].

Outcome Measures. The primary outcome measure was general psychiatric symptomatology. This was measured using the Dutch version of the Brief Symptom Inventory [33, 34], a validated self-report scale derived from the Symptom Checklist 90 - Revised [35, 36]. In this study, we used the mean score of the 53 items of the Brief Symptom Inventory, i.e. the Global Severity Index (GSI), ranging from 0 to 4. Psychosocial functioning was measured with 2 subscales of the Outcome Questionnaire-45 (OQ-45): (1) interpersonal relations and (2) social role functioning [37]. Health-related quality of life was measured using the EuroQol EQ-5D (EQ-5D) [38]. All 4 outcome measures, the GSI, OQ-45 interpersonal relations, OQ-45 social role, and EQ-5D, were assessed at baseline and several follow-up points. Three treatment centres conducted their follow-up at approximately 12, 24, and 36 months after baseline; the other 3 treatment centres conducted their follow-up at the end of treatment, approximately 6 and 12 months afterwards, and again at 36 months after baseline. The use of different assessment points was due to logistic reasons, and was taken into account by choosing multilevel modelling as the statistical method for the analyses.

Statistical Analyses

We first examined the uncorrected results on all 4 outcome measures at 12 months after baseline. We used multilevel modelling to deal with: (1) the dependency of repeated measures on the same subject in time and (2) longitudinal data with observations unequally spaced in time (see 'Outcome measures'). To estimate the uncorrected treatment effect at 12 months after baseline, we used a random intercept and random slope model with time as level I and patient number as level II. This resulted in a final best-fitting model with the following independent variables: dummy variables indicating group membership, time, and interaction between group membership and time. Subsequently, we calculated within-group effect sizes (Cohen's d) [39] to describe change from baseline to 12 months in each group.

However, since this is a non-randomised study, the comparison of the groups had to be corrected for the influence of confounders, i.e. initial patient differences. To adjust for these differences and avoid bias in effect estimation, we included a 'multiple propensity score' in our analyses. The classic propensity score is defined as the conditional probability of assignment to 1 of 2 treatment groups given a set of observed pre-treatment variables [40]. The multiple propensity score is an extension of the classic propensity score to more than 2 treatment groups [42]. Statistical inclusion of possible confounders in the outcome analyses controls selection bias due to known confounders while comparing multiple groups. To identify relevant confounders, we considered a long list of social, economic, and diagnostic variables carefully selected by both clinicians and researchers, based on the literature and clinical knowledge [43]. All variables significantly related to a specific outcome were used to estimate the multiple propensity scores in a multinomial regression analysis, with group membership as a dependent variable (see table 1 for the variables included in the GSI propensity score; complete list of potential/identified confounders for all outcome variables available upon request).

Table 1. Variables used for propensity score estimation, outcome GSI

Variable	Content
Age	patient's age
DAPP-BQ Emotional dysregulation	unstable affective responding, interpersonal problems
DAPP-BQ Inhibition	deriving little enjoyment from intimate relationships
MTQ-8 Need for help	patient's expressed desire for external help
MTQ-8 Readiness to change	willingness for treatment-seeking behaviour
EQ-5D	quality of life
SIPP Self-control	capacity to tolerate, use and control one's own emotions and impulses
SIPP Identity integration	coherence of identity; the ability to see oneself and one's own life as stable, integrated and purposive
SIPP Relational capacities	capacity to genuinely care about others as well as feeling cared for by them, to be able to communicate personal experiences, and to hear and engage with the experiences of others often but not necessarily in the context of a long-term intimate relationship
SIPP Responsibility	capacity to set realistic goals, and to achieve these goals in line with the expectations generated in others
GSI	level of psychiatric symptoms
OQ-45 Symptom distress	level of symptom distress
OQ-45 Relational functioning	level of interpersonal functioning
OQ-45 Social role functioning	level of social and work functioning
Dimensional score cluster C PD	dimensional score of cluster C PD characteristics
Total dimensional score all PD	dimensional score of all PD characteristics
Avoidant PD	diagnosis of avoidant PD
Dependent PD	diagnosis of dependent PD
Obsessive-compulsive PD	diagnosis of obsessive-compulsive PD

One major advantage of the propensity score method, as compared to other correction techniques, is the fact that the overlap in propensity score distributions (and thus the overlap in relevant variables) between treatment groups can be easily judged and visualised. From looking at the overlap between the 5 treatment groups, it appeared that in spite of some differences these groups were readily comparable. For a detailed description of this method and its use in psychotherapy research, see Bartak et al. [43].

A more sophisticated multilevel model, now including multiple propensity scores, was used to compare change in outcome variables across treatment groups. Dependent variables were the change scores (from baseline) observed during follow-up for each of the outcome measures. Independent variables were dummy variables indicating group membership, time, interaction between group membership and time, and the multiple propensity scores (with their mutual interactions). This model estimated differences in change scores at 12 months after baseline in pairwise comparisons of the 5 treatment groups. If significant differences in change scores were found, we calculated between-group effect sizes.

To render the outcome estimates at 12 months more reliable, we made optimum use of the potential of our data-set by including all available data collected up to 800 days after baseline. Data collected after that point was not used in order to prevent bias of the 12-month data due to changes much later in the process. The number of available follow-up measures was as follows: up to 800 days, 30.5% of the total sample had 1 follow-up measure, 36.7% had 2 follow-up measures, and 32.9% had 3 follow-up measures. The analyses were performed using SPSS 15.0 for data preparation and Proc Mixed of SAS 9.1.3 for multilevel modelling (SAS Institute, Cary, N.C., USA).

Results

Sample Characteristics

Of the 371 patients, 29.6% were male and 70.4% were female. The mean age was 33.5 years (SD 9.5). The highest level of education was low for 22.9%, medium for 19.4%, and high for 57.7%. Furthermore, 70.4% were unmarried, 21.3% were married, and 8.4% were divorced or widowed. The majority, 66.6%, had 'pure' cluster C PD (i.e. no comorbid cluster A or B PD), 23.7% had a combination of cluster C PD and cluster B PD, 4.0% had a combination of cluster C PD and cluster A PD, and 5.7% had a combination of cluster C PD and both cluster A and B PD. A majority (63.3%) had a diagnosis of avoidant PD, 49.3% had a diagnosis of obsessive-compulsive PD, and 22.6% a diagnosis of dependent PD.

Uncorrected Outcome

One year after baseline, patients in all treatment groups showed improvement in terms of psychiatric symptoms (GSI), the primary outcome measure. This is shown in table 2 and figure 2. Within-group effect sizes

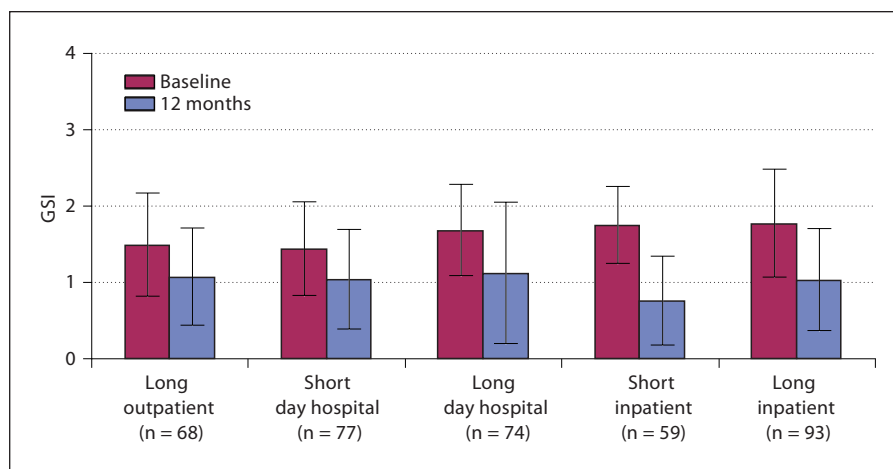


Fig. 2. GSI uncorrected mean scores at baseline and 12-month follow-up.

Table 2. Uncorrected outcomes (mean \pm SD) and effect sizes in 5 treatment groups for all outcome variables

Variable	Treatment group	Baseline	12 months	Within-group effect size, Cohen's d
GSI	long, outpatient (n = 68)	1.49 \pm 0.69	1.07 \pm 0.65	0.63
	short, day hospital (n = 77)	1.44 \pm 0.63	1.04 \pm 0.67	0.62
	long, day hospital (n = 74)	1.68 \pm 0.61	1.12 \pm 0.94	0.71
	short, inpatient (n = 59)	1.75 \pm 0.52	0.76 \pm 0.60	1.78
	long, inpatient (n = 93)	1.77 \pm 0.72	1.03 \pm 0.68	1.06
OQ-45 Social role	long, outpatient (n = 68)	15.84 \pm 4.27	12.98 \pm 4.42	0.66
	short, day hospital (n = 77)	15.20 \pm 4.52	13.59 \pm 4.53	0.36
	long, day hospital (n = 74)	16.79 \pm 4.75	13.39 \pm 5.29	0.68
	short, inpatient (n = 59)	17.78 \pm 3.84	12.41 \pm 4.83	1.24
	long, inpatient (n = 93)	16.97 \pm 4.64	12.42 \pm 5.31	0.92
OQ-45 Interpersonal relations	long, outpatient (n = 68)	22.22 \pm 5.98	19.37 \pm 6.43	0.46
	short, day hospital (n = 77)	20.93 \pm 5.24	18.17 \pm 5.90	0.50
	long, day hospital (n = 74)	22.89 \pm 6.41	18.41 \pm 8.05	0.62
	short, inpatient (n = 59)	23.97 \pm 5.63	17.54 \pm 6.77	1.04
	long, inpatient (n = 93)	24.09 \pm 5.24	18.38 \pm 6.59	0.96
EQ-5D	long, outpatient (n = 68)	0.58 \pm 0.24	0.73 \pm 0.16	0.74
	short, day hospital (n = 77)	0.60 \pm 0.25	0.69 \pm 0.24	0.37
	long, day hospital (n = 74)	0.50 \pm 0.27	0.72 \pm 0.22	0.90
	short, inpatient (n = 59)	0.49 \pm 0.27	0.78 \pm 0.21	1.21
	long, inpatient (n = 93)	0.51 \pm 0.26	0.68 \pm 0.25	0.67

of the uncorrected scores ranged from 0.62 (medium effect, short-term day hospital group) to 1.78 (huge effect, short-term inpatient group).

Improvements were also seen in terms of psychosocial functioning and quality of life (table 2). Effect sizes for these outcome measures were somewhat lower compared to psychiatric symptoms, but a positive change in psychosocial functioning and quality of life was evident.

Corrected Comparison

After correction for all relevant pre-treatment differences, improvement between baseline and assessment at 12 months proved to be significant for patients in all treatment groups on all 4 outcome measures ($p < 0.001$).

The short-term inpatient group showed significantly more improvement in psychiatric symptoms (GSI) than 3 other groups: the short-term day hospital group ($\beta =$

Table 3. Difference scores (β) of 5 treatment groups 12 months after baseline, corrected for propensity score (all outcome variables)

Variable	Treatment group	n	β			
			long, outpatient	short, day hospital	long, day hospital	short, inpatient
GSI	long, outpatient	68				
	short, day hospital	77	-0.078			
	long, day hospital	74	-0.128	-0.050		
	short, inpatient	59	0.302	0.380**	0.430**	
	long, inpatient	93	-0.004	0.075	0.124	-0.306*
OQ-45 Social role	long, outpatient	68				
	short, day hospital	77	-1.632			
	long, day hospital	74	-1.123	0.460		
	short, inpatient	59	0.876	2.508**	2.048*	
	long, inpatient	93	-0.169	1.463	1.003	-1.045
OQ-45 Interpersonal relations	long, outpatient	68				
	short, day hospital	77	-0.836			
	long, day hospital	74	-0.611	0.225		
	short, inpatient	59	1.704	2.540*	2.315	
	long, inpatient	93	-0.084	0.752	0.527	-1.788
EQ-5D	long, outpatient	68				
	short, day hospital	77	-0.060			
	long, day hospital	74	0.001	0.061		
	short, inpatient	59	0.089	0.149***	0.088	
	long, inpatient	93	-0.021	0.039	-0.022	-0.110*

Positive coefficients indicate that the treatment group shown in the left column is superior, negative coefficients indicate that the treatment group in the above row is superior. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

0.38, $p = 0.0059$, 95% CI 0.11–0.65), the long-term day-hospital group ($\beta = 0.43$, $p = 0.0032$, 95% CI 0.15–0.71), and the long-term inpatient group ($\beta = 0.31$, $p = 0.0248$, 95% CI 0.04–0.57) (table 3). Between-group effect sizes (Cohen's d) were 0.54, 0.57, and 0.40, respectively. This indicates medium effect sizes for the between-group comparisons of short-term inpatient treatment versus other treatment groups.

In terms of social role functioning, the short-term inpatient group improved significantly more than 2 other groups – the short-term day hospital group ($\beta = 2.51$, $p = 0.0067$, 95% CI 0.71–4.31) and the long-term day hospital group ($\beta = 2.05$, $p = 0.0476$, 95% CI 0.02–4.07) – with between-group effect sizes of 0.49 and 0.38, respectively. The improvement in interpersonal functioning was significantly higher in the short-term inpatient group than in one other group – the short-term day hospital group ($\beta = 2.54$, $p = 0.0319$, 95% CI 0.22–4.86) – with a between-group effect size of 0.39. Quality of life improved significantly more in the short-term inpatient group than in 2 other groups: the short-term day-hospital group ($\beta =$

0.15, $p = 0.0009$, 95% CI 0.06–0.23) and the long-term inpatient group ($\beta = 0.11$, $p = 0.0113$, 95% CI 0.03–0.19). Between-group effect sizes were 0.6 and 0.42, respectively.

All results were based on intention-to-treat analyses (ITT), whereby ITT is defined as assignment and a minimal exposure to the intended treatment modality. The analyses were repeated with the treatment completers, i.e. those who actually stayed in the intended treatment modality group during their treatment ($n = 298$, 80.3% of the ITT sample, ranging from 66.2% for short-term day hospital to 89.7% for long-term outpatient treatment). These results followed the same pattern as the results from the ITT analyses: significant change within all treatment groups and a superiority of short-term inpatient treatment across all outcome measures (data available on request).

Discussion

This is the first study comparing the effectiveness of 5 modalities of psychotherapeutic treatment in a large population of patients with cluster C PD, as a contribution to the search for effective treatments for this patient group. Patients in all treatment groups had improved psychiatric symptoms, psychosocial functioning, and quality of life after 12 months. Most improvement was observed in the short-term inpatient group. This finding held when pre-treatment differences were controlled for with the propensity score.

Strengths and Limitations

A clear strength of the present study is its external validity and clinical utility: it was conducted in regular clinical practice, not under experimental conditions [44]. A second strength is the rigorous statistical control of potential confounders, using the multiple propensity score methodology. Finally, a major asset of this study is its large number of patients. All this enabled the comparison of different psychotherapeutic treatment modalities while keeping sufficient statistical power.

Despite these strengths, the present findings have to be interpreted considering several limitations. First, even though we controlled for all observed pre-treatment differences, it cannot be ruled out that results have been influenced by unobserved confounders. To diminish this constraint as much as possible, a broad range of possible confounders was carefully selected and measured, based on both clinical and empirical knowledge [43], including variables identified in the literature as significant predictors of therapy outcome or process, such as severity of baseline psychopathology, previous hospitalisation, and substance misuse [e.g. 45–50]. In line with these earlier findings, previous hospitalisation and substance misuse for example were significantly related to one of the secondary outcome measures, interpersonal functioning, and were therefore included in the propensity score for this measure. However, even when considerably reducing the possibility of important confounders being overlooked, not all possible variables could be covered in interviews and questionnaires at baseline, and therefore several variables, such as self-harm [51], were not measured.

Second, for ethical reasons, a control group receiving no treatment at all was not included. Yet, several previous studies showed that specialised psychotherapeutic treatment yields better outcomes than various control conditions (for example waiting list controls) [8, 9, 13].

Third, research compliance differed between the treatment groups compared with most missing follow-up observations in the long-term treatment groups (fig. 1). This might cause a problem of internal validity if non-response is not random, but related to systematic bias in effect estimation (positive or negative). However, there are 2 reasons why systematic bias seems unlikely: (1) responders and non-responders did not differ in psychiatric symptoms at baseline, and therefore it seems that they do not represent 2 structurally different groups of patients; (2) during the frequent telephone contact the authors had with non-responding patients to remind them to send back their questionnaires, these patients reported both negative and positive outcomes as reasons why they did not respond: some of them argued that their problems had worsened and that therefore they felt they did not have enough energy to fill in the questionnaires, others argued that their life had changed in a positive way and that therefore they did not want to be reminded of their time in therapy by filling in the questionnaires. Keeping this in mind, it seems unlikely that non-response was related to systematic negative or positive bias.

Fourth, this study does not rule out the possibility that treatment characteristics other than setting and duration played a role in the differential effectiveness of the 5 treatment modalities, e.g. frequency of sessions or theoretical orientation of treatment. This might represent a potential threat to internal validity. This is especially true for the role of theoretical orientation as a possible factor in the superiority of short-term inpatient treatment: most short-term inpatient programs were based on psychodynamic principles. This concern is somewhat mitigated by previous studies comparing different theoretical orientations where no differences were found [e.g. 12]. However, to test the differential effect of modality and other treatment characteristics, a combined research design combining all these factors is needed.

Future Directions and Implications

What are the implications of the present results for future research, for practice guidelines, and for everyday clinical practice?

For patients with cluster C personality pathology, short-term inpatient treatment clearly was associated with the highest improvement within 12 months. For this patient group, this modality of therapy seems to be the treatment backed up by the best available evidence – in absence of long-term follow-up data. Replication of these results in a long(er)-term follow-up study is of vital importance to draw final conclusions. There might be a bias

in favour of short-term treatment because patients in the long-term treatment groups might still be in therapy at 12 months. Long-term follow-up after termination of all treatment programs is therefore warranted. Another question is whether the benefit in terms of effectiveness is worth the potential cost differences when evaluated with recently upcoming state-of-the-art cost-effectiveness analyses [e.g. 52, 53]. From these analyses within our study sample, it appeared that the mean direct treatment costs of the 5 treatment modalities were EUR 10,005 (SE 1,134) for long-term outpatient treatment, EUR 16,813 (SE 1,361) for short-term day hospital treatment, EUR 27,648 (SE 2,654) for long-term day hospital treatment, EUR 25,933 (SE 859) for short-term inpatient treatment, and EUR 49,260 (SE 2,435) for long-term inpatient treatment [Soeteman, unpubl. data]. It would be interesting to compare the cost-effectiveness of short-term inpatient psychotherapeutic treatment with that of manual-based outpatient treatments, such as cognitive-behavioural therapy [9]. A state-of-the-art cost-effectiveness analysis would include medical costs incurred outside the treatment institution, productivity costs, and other indirect costs. This kind of analysis and its economic interpretation is beyond the range of this study and needs considerable research in the future.

If the superiority of short-term inpatient psychotherapeutic treatment holds at long-term follow-up, in cost-effectiveness analyses, and in comparison with other evidence-based manual-based treatments, this treatment modality might be considered as the treatment of choice for this patient group. This would be a thought-provoking finding, as previous studies in cluster B PD patients have found outpatient [54–56] and day hospital [57] treatments to be very effective in this population. Even though no study compared one of these modalities directly with inpatient therapy, one might speculate that different therapy modalities are effective for different groups of patients. It could be that the success of short-term inpatient treatment in a cluster C PD sample is embedded in the combination of only short hospitalisation – thereby preventing iatrogenic effects – and a high level of therapeutic intensity and pressure. Patients with cluster C personality pathology might be able to handle the high pressure of this treatment modality better than (pure) cluster B PD patients, who probably have a lower tolerance for therapeutic pressure, resulting in more early dropouts and thus a less effective treatment. They might instead need less pressure with a longer treatment duration [57, 58]. Future studies may verify this hypothesis. However, even when superiority of short-term inpatient treatment for

cluster C PD patients has been confirmed in the literature, patients caring for children might still not be assigned to inpatient treatment. Also, patients with a high severity of psychiatric symptoms or a low level of ego strength might not be able to handle the pressure of intensive inpatient treatment. It is recommended to investigate these potential matching factors further as this would enable clinicians to make specific treatment recommendations for different subgroups of cluster C PD patients and to develop new clinical practice guidelines.

In conclusion, this study suggests that psychotherapy, especially in a short-term inpatient modality, is an effective treatment for patients with cluster C PD. This makes inpatient psychotherapeutic treatment an interesting option for patients with avoidant, dependent, and obsessive-compulsive PD. The present findings can contribute to more adequate and tailored health care for this vulnerable patient group, as implementing effective treatments may reduce the considerable burden to individuals and society as a whole.

Acknowledgements

The authors report no financial or other relationship relevant to the subject of this article. We are very grateful to Els Havermans who helped with the data collection. We also thank all the patients and therapists who took part in the study.

References

- 1 Coid J, Yang M, Tyrer P, Roberts A, Ullrich S: Prevalence and correlates of personality disorder in Great Britain. *Br J Psychiatry* 2006;188:423–431.
- 2 Skodol AE, Gunderson JG, McGlashan TH, Dyck IR, Stout RL, Bender DS, Grilo CM, Shea MT, Zanarini MC, Morey LC, Sanislow CA, Oldham JM: Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *Am J Psychiatry* 2002;159:276–283.
- 3 Skodol AWE, Johnson JG, Cohen P, Sneed JR, Crawford TN: Personality disorder and impaired functioning from adolescence to adulthood. *Br J Psychiatry* 2007;190:415–420.
- 4 Grant BF, Hasin DS, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP: Prevalence, correlates, and disability of personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry* 2004;65:948–958.
- 5 Soeteman DI, Hakkaart-van Roijen L, Verheul R, Busschbach JJV: The economic burden of personality disorders in mental health care. *J Clin Psychiatry* 2008;69:259–265.

- 6 Duggan C, Huband N, Smailagic N, Ferriter M, Adams C: The use of psychological treatments for people with personality disorder: a systematic review of randomized controlled trials. *Pers Mental Health* 2007;1:95–180.
- 7 van Emmerik AAP, Kamphuis JH, Emmelkamp PMG: Treating acute stress disorder and posttraumatic stress disorder with cognitive behavioral therapy or structured writing therapy: a randomized controlled trial. *Psychother Psychosom* 2008;77:93–100.
- 8 Alden L: Short-term structured treatment for avoidant personality disorder. *J Consult Clin Psychol* 1989;57:756–764.
- 9 Emmelkamp PMG, Benner A, Kuipers A, Feiertag GA, Koster HC, Van Apeldoorn FJ: Comparison of brief dynamic and cognitive-behavioural therapies in avoidant personality disorder. *Br J Psychiatry* 2006;189:60–64.
- 10 Hellerstein DJ, Rosenthal RN, Pinsker H, Samstag LW, Muran JC, Winston A: A randomized prospective study comparing supportive and dynamic therapies: outcome and alliance. *J Psychother Pract Res* 1998;7:261–271.
- 11 Stravynski A, Belisle M, Marcouiller M, Lavallee YJ, Elie R: The treatment of avoidant personality disorder by social skills training in the clinic or in real-life settings. *Can J Psychiatry* 1994;39:377–383.
- 12 Svartberg M, Stiles TC, Seltzer MH: Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *Am J Psychiatry* 2004;161:810–817.
- 13 Winston A, Laikin M, Pollack J, Samstag LW, McCullough L, Muran JC: Short-term psychotherapy of personality disorders. *Am J Psychiatry* 1994;151:190–194.
- 14 Karterud S, Pedersen G, Bjordal E, Brabrand J, Friis S, Haaseth O, Haavaldsen G, Irion T, Leirvag H, Torum E, Urnes O: Day treatment of patients with personality disorders: experiences from a Norwegian treatment research network. *J Pers Disord* 2003;17:243–262.
- 15 Wilberg T, Urnes O, Friis S, Irion T, Pedersen G, Karterud S: One-year follow-up of day treatment for poorly functioning patients with personality disorders. *Psychiatr Serv* 1999;50:1326–1330.
- 16 Gude T, Vaglum P: One-year follow-up of patients with cluster C personality disorders: a prospective study comparing patients with 'pure' and comorbid conditions within cluster C, and 'pure' C with 'pure' cluster A or B conditions. *J Pers Disord* 2001;15:216–228.
- 17 Teusch L, Bohme H, Finke J, Gastpar M: Effects of client-centered psychotherapy for personality disorders alone and in combination with psychopharmacological treatment: an empirical follow-up study. *Psychother Psychosom* 2001;70:328–336.
- 18 Arnevik E, Wilberg T, Urnes O, Johansen M, Monsen JT, Karterud S: Psychotherapy for personality disorders: short-term day hospital psychotherapy versus outpatient individual therapy – a randomized controlled study. *Eur Psychiatry* 2009;24:71–78.
- 19 Black N: Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215–1218.
- 20 Zeeck A, Weber S, Sandholz A, Wetzler-Burmeister E, Wirsching M, Hartmann A: Inpatient versus day clinic treatment for bulimia nervosa: a randomized trial. *Psychother Psychosom* 2009;78:152–160.
- 21 Facchinetti F, Ottolini F, Fazzio M, Rigatelli M, Volpe A: Psychosocial factors associated with preterm uterine contractions. *Psychother Psychosom* 2007;76:391–394.
- 22 Forstmeier S, Rueddel H: Improving volitional competence is crucial for the efficacy of psychosomatic therapy: a controlled clinical trial. *Psychother Psychosom* 2007;76:89–96.
- 23 Golkaramnay V, Bauer S, Haug S, Wolf M, Kordy H: The exploration of the effectiveness of group therapy through an Internet chat as aftercare: a controlled naturalistic study. *Psychother Psychosom* 2007;76:219–225.
- 24 Grossman P, Tiefenthaler-Gilmer U, Raysz A, Kesper U: Mindfulness training as an intervention for fibromyalgia: evidence of postintervention and 3-year follow-up benefits in well-being. *Psychother Psychosom* 2007;76:226–233.
- 25 Van Manen J: How do intake clinicians use patient characteristics to select treatment for patients with personality disorders? *Psychother Res* 2008;18:711–718.
- 26 Vervaeke GAC, Emmelkamp PMG: Treatment selection: What do we know? *Eur J Psychol Assess* 1998;14:50–59.
- 27 DeJong CAJ, Derks FCH, Van Oel CJ, Rinne T: Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen (SIDP-IV). Sint Oedenrode, Stichting Verslavingszorg Oost Brabant, 1996.
- 28 Pfohl B, Blum N, Zimmerman M: Structured Interview for DSM-IV Personality (SIDP-IV). Washington, American Psychiatric Press, 1997.
- 29 van Kampen D: The DAPP-BQ in the Netherlands: factor structure and relationship with basic personality dimensions. *J Pers Disord* 2002;16:235–254.
- 30 Livesley WJ, Jackson DN: Manual for the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ). Port Huron, Sigma Press, 2002.
- 31 Verheul R, Andrea H, Berghout CC, Dolan C, Busschbach JJV, Van der Kroft PJA, Bateman AW, Fonagy P: Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability and validity. *Psychol Assess* 2008;20:23–34.
- 32 van Beek N, Verheul R: Motivation for treatment in patients with personality disorders. *J Personal Disord* 2008;22:89–100.
- 33 Derogatis LR, Melisaratos N: The Brief Symptom Inventory: an introductory report. *Psychol Med* 1983;13:595–605.
- 34 de Beurs E, Zitman FG: De Brief Symptom Inventory (BSI): de betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90 (The Brief Symptom Inventory (BSI): the reliability and validity of a brief alternative of the SCL-90). *Maandblad Geestelijke Volksgezondheid* 2006;61:120–141.
- 35 Derogatis LR: SCL-90-R Administration, Scoring and Procedures: Manual-II for the Revised Version. Townson, Clinical Psychometric Research, 1986.
- 36 Arrindell WA, Ettema JHM: SCL-90-R: herziene handleiding bij een multidimensionale psychopathologie-indicator (SCL-90-R: Revised Manual for a Multidimensional Indicator of Psychopathology). Lisse, Swets & Zeitlinger, 2003.
- 37 Lambert MJ, Burlingame GM, Umphress V, Hansen NB, Vermeersch DA, Clouse GC, Yanchar SC: The reliability and validity of the outcome questionnaire. *Clin Psychol Psychother* 1996;3:249–258.
- 38 Brooks R, Rabin R, de Charro F: The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective. Evidence from the EuroQol BIOMED Research Programme. Dordrecht, Kluwer Academic, 2003.
- 39 Cohen J: Statistical Power Analysis for the Behavioral Sciences. Hillsdale, Lawrence Erlbaum Associates, 1988.
- 40 Rosenbaum PR, Rubin DB: The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
- 41 Imbens GW: The role of the propensity score in estimating dose-response functions. *Biometrika* 2000;87:706–710.
- 42 Spreeuwenberg MD, Bartak A, Croon MA, Hagens JA, Busschbach JJV, Andrea H, Twisk J, Stijnen T: The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health. *Medical Care*, in press.
- 43 Bartak A, Spreeuwenberg MD, Andrea H, Busschbach JJV, Croon MA, Verheul R, Emmelkamp PMG, Stijnen T: The use of propensity score methods in psychotherapy research: a practical application. *Psychother Psychosom* 2009;78:26–34.
- 44 Hodgson R, Bushe C, Hunter R: Measurement of long-term outcomes in observational and randomised controlled trials. *Br J Psychiatry* 2007;191:78–84.
- 45 Gunderson JG, Daversa MT, Grilo CM, McGlashan TH, Zanarini MC, Shea MT, Skodol AE, Yen S, Sanislow CA, Bender DS, Dyck IR, Morey LC, Stout RL: Predictors of 2-year outcome for patients with borderline personality disorder. *Am J Psychiatry* 2006;163:822–826.

- 46 Links PS, Mitton MJE, Steiner M: Stability of borderline personality-disorder. *Can J Psychiatry* 1993;38:255–259.
- 47 McGlashan TH: The prediction of outcome in borderline personality disorders; in McGlashan TH (ed): *The Borderline: Current Empirical Research*. Washington, American Psychiatric Press, 1985, pp 61–98.
- 48 Ogrodniczuk JS, Joyce AS, Lynd LD, Piper WE, Steinberg PI, Richardson K: Predictors of premature termination of day treatment for personality disorder. *Psychother Psychosom* 2008;77:365–371.
- 49 Plakun EM: Prediction of outcome in borderline personality disorder. *J Pers Disord* 1991;5:93–101.
- 50 Ryle A, Golyukina K: Effectiveness of time-limited cognitive analytic therapy of borderline personality disorder: factors associated with outcome. *Br J Med Psychol* 2000;73(part 2):197–210.
- 51 Chiesa M, Fonagy P: Prediction of medium-term outcome in cluster b personality disorder following residential and outpatient psychosocial treatment. *Psychother Psychosom* 2007;76:347–353.
- 52 Leichsenring F, Hoyer J, Beutel M, Herpertz S, Hiller W, Irle E, Joraschky P, König HH, de Liz TM, Nolting B, Pohlmann K, Salzer S, Schauenburg H, Stangier U, Strauss B, Subic-Wrana C, Vormfelde S, Weniger G, Willutzki U, Wiltink J, Leibing E: The social phobia psychotherapy research network. The first multicenter randomized controlled trial of psychotherapy for social phobia: rationale, methods and patient characteristics. *Psychother Psychosom* 2009;78:35–41.
- 53 McCrone P, Marks IM, Greist JH, Baer L, Kobak KA, Wenzel KW, Hirsch MJ: Cost-effectiveness of computer-aided behaviour therapy for obsessive-compulsive disorder. *Psychother Psychosom* 2007;76:249–250.
- 54 Chiesa M, Fonagy P, Gordon J: Community-based psychodynamic treatment program for severe personality disorders: clinical description and naturalistic evaluation. *J Psychiatr Pract* 2009;15:12–24.
- 55 Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF: Evaluating three treatments for borderline personality disorder: A multiwave study. *Am J Psychiatry* 2007;164:922–928.
- 56 Giesen-Bloo J, Van Dyck R, Spinhoven P, Van Tilburg W, Dirksen C, Van Asselt T, Kremers I, Nadort M, Arntz A: Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs. transference-focused psychotherapy. *Arch Gen Psychiatry* 2006;63:649–658.
- 57 Bateman A, Fonagy P: Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *Am J Psychiatry* 2001;158:36–42.
- 58 Lorentzen S, Hoglend PA: Moderators of the effects of treatment length in long-term psychodynamic group psychotherapy. *Psychother Psychosom* 2008;77:321–322.