

Effectiveness of Outpatient, Day Hospital, and Inpatient Psychotherapeutic Treatment for Patients with Cluster B Personality Disorders

Anna Bartak^{a,c} Helene Andrea^{c,e} Marieke D. Spreeuwenberg^{f,g} Uli M. Ziegler^h
Jack Dekker^b Bert V. Rossumⁱ Elisabeth F.M. Hamers^d Wubbo Scholte^j
Janneke Aerts^k Jan J.V. Busschbach^{c,e} Roel Verheul^{a,d} Theo Stijnen^l
Paul M.G. Emmelkamp^a

^aDepartment of Clinical Psychology, University of Amsterdam, and ^bArkin, Amsterdam; ^cInstitute for Studies on Personality Disorders (VISPD), and ^dPsychotherapeutic Centre De Viersprong, Halsteren; ^eDepartment of Medical Psychology and Psychotherapy, Erasmus Medical Centre, Rotterdam; ^fHogeschool Zuyd, Heerlen; ^gDepartment of Methodology and Statistics, Tilburg University, Tilburg; ^hZaans Medical Centre, Zaandam; ⁱAltrecht, Zeist; ^jDe 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 · Antisocial personality disorder · Borderline personality disorder · Histrionic personality disorder · Narcissistic personality disorder

Abstract

Background: For patients with cluster B personality disorders there is no consensus regarding the optimal treatment setting. The aim of this study was to compare the effectiveness of different psychotherapeutic settings for patients with cluster B personality disorders, i.e. outpatient, day hospital, and inpatient treatment. **Methods:** The study was conducted between March 2003 and June 2008 in 6 mental health care centres in the Netherlands, with a sample of 207 patients with a DSM-IV-TR axis II cluster B diagnosis. Patients were assigned to 3 different settings of psychotherapeutic treatment and effectiveness was assessed at 18 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 3 settings improved significantly in terms of psychiatric symptoms, social and interpersonal functioning, and quality of life 18 months after baseline. The inpatient group showed the largest improvements. The comparison of outpatient and inpatient treatment regarding psychiatric symptoms showed a marginally significant result ($p = 0.057$) in favour of inpatient treatment. **Conclusions:** Patients with cluster B personality disorders improved in all investigated treatment settings, with a trend towards larger improvements of psychiatric symptoms in the inpatient setting compared to the outpatient setting. Specialised inpatient treatment should be considered as a valuable treatment option for cluster B personality disorders, both in research and in clinical practice.

Copyright © 2010 S. Karger AG, Basel

Introduction

Cluster B personality disorders (PD) affect a considerable percentage of the general population (3.1–4.5%) [1, 2] and are highly prevalent in psychiatric patients (13.0%) [3]. Borderline, histrionic, narcissistic, and antisocial PD are not only associated with individual suffering [4–7], but also with early institutional care and criminality [8]. They place a heavy burden on both individuals and society as a whole [9–13]. To date, a number of specific treatment protocols tailored to this patient group have been developed. This is especially the case for borderline PD [14, 15]. Examples are dialectical behaviour therapy [16], mentalization-based treatment [17], schema-focused therapy [18], and transference-focused therapy [19]. As a consequence, methodologically sound research on psychotherapeutic treatments for cluster B PD has increased during the last 10 years. However, no study has directly compared outpatient, day hospital, and inpatient treatment within one patient sample. The available published studies have focused either on the development of specific treatment manuals and their evaluation [20–29] or on the comparison of different theoretical schools within outpatient settings [30–32]. Another focus of psychotherapy outcome research was the effectiveness of specific treatment ingredients, such as transference interpretations [33, 34] or crisis support outside office hours [35]. So until now, hardly any study has addressed possible differences in treatment effectiveness between different settings and across theoretical schools. Exceptions are the studies by Chiesa et al. [36–38] comparing inpatient treatment and step-down/outpatient treatment and the study by Arnevik et al. [39] comparing day hospital treatment and outpatient treatment.

Comparing different treatment settings is relevant as it can be anticipated that the type of setting will affect costs for both patients and society, as well as patients' time. One of the reasons this comparison (e.g. outpatient versus day hospital versus inpatient treatment) has not been made previously is probably the difficulty of randomly assigning patients to different treatment settings in clinical samples due to practical or ethical constraints [40]. Furthermore, even if researchers were successful in setting up and starting a randomised treatment setting study, external validity would be doubtful because a high number of patients would probably refuse to participate. In non-randomised studies, however, researchers often cannot draw valid conclusions due to the strong influence of selection bias. Accordingly, there is a need to design and conduct non-randomised effectiveness studies

while minimising the influence of selection bias. This kind of study is increasingly being found in the literature [41–44]. In the present study, this has been undertaken for the first time in a population of cluster B PD patients using the propensity score method to minimise the influence of selection bias.

The purpose of our study was to compare the effectiveness of different treatment settings (outpatient, day hospital, inpatient) in psychotherapy for cluster B PD 18 months after baseline. The primary outcome measure in this study was psychiatric symptomatology. Additionally, effects on interpersonal and social functioning as well as on quality of life were evaluated.

Method

Participants

Participants ($n = 207$) 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 and Roosendaal; Arkin, Amsterdam). These institutions offer outpatient, day hospital, and/or inpatient psychotherapy for patients with PD. From March 2003 to March 2006, 1,379 patients completed the intake and screening procedure and were selected for treatment (fig. 1). The intake and screening data were collected for all applicants, irrespective of study participation later on. The data obtained from this initial assessment served as baseline data for our study. As it was part of the standard screening procedure, informed consent for the baseline data collection was not mandatory under Dutch law.

Of these 1,379 patients, 146 (10.6%) were excluded from the study for not meeting one of the following inclusion criteria: age between 18 and 70 years ($n = 13$), 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 eligible participants, of whom 31 (2.5%) could not participate due to logistic reasons (e.g. no appointment could be made to provide informed consent). A total of 133 patients (10.9%) had to be excluded due to missing or unreliable baseline data during the intake and screening procedure. For the majority of these 133 patients ($n = 106$), no standardised axis II diagnosis was obtained. This was mostly due to a practical problem, i.e. a shortage of interviewers at the start of the study ($n = 101$). Because of this logistic reason, it can be assumed that these data were 'missing completely at random' and therefore they do not threaten internal validity (as they were unrelated to specific patient characteristics). Five patients had an unreliable axis II diagnosis as determined by the interviewer, for example because the patient was too anxious or too depressed to obtain a reliable diagnosis with a semi-structured interview. A few patients did not return their assessment booklet during the intake procedure ($n = 27$). Thirty-eight pa-

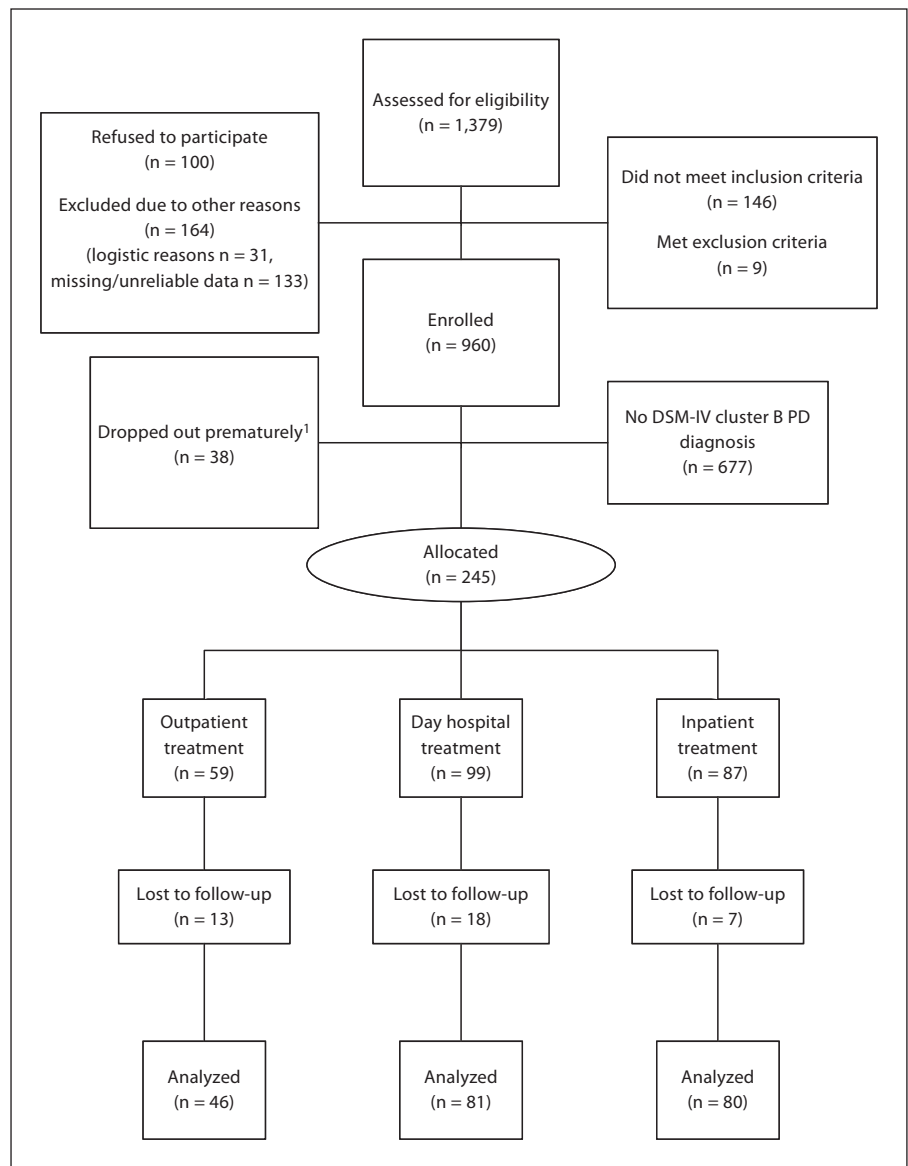


Fig. 1. Patient flow. ¹These patients did not receive a ‘minimal effective dose’ of treatment, defined as two sessions for outpatients and two treatment days for day hospital patients and inpatients, and were therefore not included in the initial study sample.

tients (3.1%) received less than 2 treatment sessions of outpatient therapy or less than 2 days of inpatient or day hospital therapy. For example, some of the patients left treatment after 1 session, and some did not even show up for the first session. They were therefore excluded beforehand from the study sample in which we only included patients with a ‘minimal effective dose’ of treatment. One hundred patients (8.2%) refused to participate. The remaining 922 patients were informed about the study and its procedure, provided written informed consent, and entered the study. Of those, 245 patients (26.6%) had 1 or more cluster B PD.

In the absence of explicit guidelines for treatment assignment in PD [45], the selection procedure was based on the expert opinion of clinicians who used their clinical experience combined with patient data from standardised instruments. To clarify the criteria used for the assignment process, our research group had

recently conducted a study with intake clinicians from the participating treatment centres. We found evidence of substantial consensus among clinicians concerning the criteria used for treatment decision making [46].

Patients were assigned to 1 of 3 setting groups: 59 to outpatient, 99 to day hospital, and 87 to inpatient psychotherapy. Follow-up data were not available for 38 patients (15.5%, patients who did not respond to any follow-up assessment or patients where follow-up measurements were as yet not 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 the comparison within the 3 treatment groups). The final sample consisted of 207 patients who were included in the intention-to-treat 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 settings, the following 3 treatment groups were compared:

(1) Patients in the outpatient treatment group ($n = 46$, 22.2% of the study sample). These patients came for individual (89.2%) or group (10.9%) psychotherapy sessions for up to 2 sessions (mean 0.7 sessions/week, SD 0.4, median 0.5) per week (mean treatment duration 14.5 months, SD 6.6, median 12.0).

(2) Patients in the day hospital treatment group ($n = 81$, 39.1% of the study sample). These patients came to the institutions at least 1 morning/afternoon per week (mean 3.5 days/week, SD 1.4, median 3.0) and received different forms of psychotherapeutic and psychosocial treatment, but slept at home (mean treatment duration 10.4 months, SD 4.8, median 12.0).

(3) Patients in the inpatient treatment group ($n = 80$, 38.6% of the study sample). These patients stayed at the institutions for 5 days a week and received different forms of psychotherapeutic and psychosocial treatment (mean treatment duration 9.1 months, SD 3.0, median 10.0).

Day hospital and inpatient programmes typically consisted of group psychotherapy as a core element, mostly in combination with 1 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 therapists were 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. The study protocol was approved by the Medical Ethics Committee of the Erasmus University Medical Centre in Rotterdam.

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 (SIDP-IV) [47, 48]. 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.). They received monthly booster sessions to avoid deviation from the interviewer guidelines. Interrater reliability was evaluated in 25 videotaped 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 type and degree of personality pathology [49, 50]. We used patients' scores on this questionnaire on the 4 higher-order factors: emotional dysregulation, dissociative behaviour, inhibition and compulsivity. To mea-

sure 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 [51]. 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 [52].

Outcome Measures. The primary outcome measure was general psychiatric symptomatology. This was measured using the Dutch version of the Brief Symptom Inventory [53, 54], a validated self-report scale derived from the revised Symptom Checklist 90 [55, 56]. 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), interpersonal relations and social role functioning [57]. Health-related quality of life was measured using the EuroQol EQ-5D [58]. All 4 outcome measures, i.e. 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 follow-up at approximately 12, 24, and 36 months after baseline; the other 3 treatment centres conducted follow-up at the end of treatment, subsequently after about 6 and 12 months, 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 18 months after baseline. We used multilevel modelling to deal with the dependency of repeated measures on the same subject in time and longitudinal data with observations unequally spaced in time (see Outcome Measures). To estimate the uncorrected treatment effect at 18 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 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) [59] to describe change from baseline to 18 months per treatment group.

However, treatment groups cannot be compared based on the uncorrected results in this non-randomised clinical trial as these findings might be confounded by initial patient differences. To adjust for these differences and to 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 pretreatment variables [60]. The multiple propensity score is an extension of the classic propensity score to more than 2 treatment groups [61]. 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 [62]. 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; a complete list of potential/identified confounders for all outcome variables is 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 Dissocial behaviour	Lacking regard for others
DAPP-BQ Inhibition	Deriving little enjoyment from intimate relationships
DAPP-BQ Compulsivity	Compulsivity and absence of oppositional behaviour
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	The capacity to tolerate, use and control one's own emotions and impulses
SIPP Social concordance	The ability to value someone's identity, withhold aggressive impulses towards others and to work together with others
SIPP Identity integration	Coherence of identity; the ability to see oneself and one's own life as stable, integrated and purposive
SIPP Relational capacities	The 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	The 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
SIDP-IV Cluster C PD	Number of cluster C PD
SIDP-IV Dimensional score cluster C PD	Dimensional score of cluster C PD characteristics
Borderline PD	Diagnosis of borderline PD
Narcissistic PD	Diagnosis of narcissistic PD

A 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 [63]. From looking at the overlap between the 3 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. [62] and Spreeuwenberg et al. [63].

To compare change in outcome variables across treatment groups, a more sophisticated multilevel model, now including

multiple propensity scores, was used. Dependent variables were all available change scores observed during follow-up for each of the outcome measures. The following independent variables were entered in the initial model: dummy variables indicating group membership, time, the multiple propensity scores (with their mutual interactions), the 2-way interactions between group membership and time, the 2-way interactions between propensity scores and time, the 2-way interactions between propensity scores and group membership, and the 3-way interactions between propensity scores, time, and group membership. Then variables were eliminated from the model by backward selection to obtain a final

best-fit model. This model estimated differences in change scores at 18 months after baseline in pairwise comparisons of the 3 treatment groups.

Follow-up response was high, enhancing the robustness of the multilevel analyses: 12.1% of the respondents completed 1 follow-up measurement, 6.8% completed 2 follow-up measurements, 39.6% completed 3 follow-up measurements, 39.1% completed 4 follow-up measurements, and 2.4% completed 5 follow-up measurements. The analyses were performed using SPSS 17.0 for data preparation and Proc Mixed of SAS 9.1.3 for multilevel modelling (SAS Institute Inc., Cary, N.C., USA).

Results

Sample Characteristics

The majority of patients (71%) were female. The mean age was 31.3 years (SD 8.5). The level of education was low for 33.3%, medium for 19.3%, and high for 47.3%. Furthermore, 78.7% were unmarried, 12.6% were married, and 8.7% were divorced or widowed. Most patients (77.3%) had a diagnosis of borderline PD, 22.7% had a diagnosis of narcissistic PD, 12.6% had a diagnosis of histrionic PD, and 8.7% a diagnosis of antisocial PD. There was considerable overlap with PD from other clusters: 40.6% had 'pure' cluster B PD (i.e., no comorbid cluster A or C PD), 44.9% had a combination of cluster B PD and cluster C PD, 3.4% had a combination of cluster B PD and cluster A PD, and 11.1% had a combination of cluster B PD and both clusters A and C PD.

Treatment Compliance

Before the start of treatment, every patient received an allocation to a certain 'treatment dosage'. According to our registration at the end of treatment, about one third of all patients (n = 62) underwent exactly the intended treatment (31.1% of the outpatient group, 23.8% of the day hospital group, and 36.7% of the inpatient group). Of the 145 patients whose received treatment deviated from their intended treatment, 60 (41.4% of the deviating patients) stayed in treatment shorter than planned (28.1% of the deviating patients in the outpatient group, 45.2% in the day hospital group, and 45.1% in the inpatient group). Of these 60 patients, 40 decided in agreement with their therapist that treatment was no longer beneficial or necessary, whereas only 20 patients dropped out of treatment prematurely or were forced to leave earlier by the staff. From these 20 dropouts, 12 patients were from the day hospital group (14.8% of the total day hospital group) and 8 patients were from the inpatient group (10.0% of the total inpatient group). All dropouts were treatment drop-

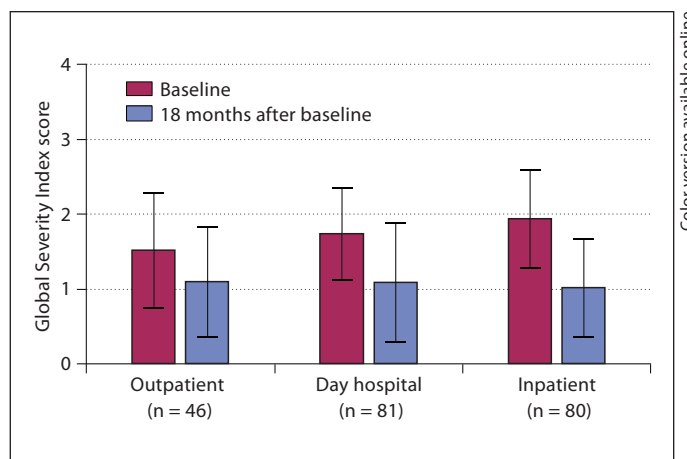


Fig. 2. GSI uncorrected mean scores at baseline and 18-month follow-up.

outs (and no study dropouts) who completed follow-up measures and were included in the intention-to-treat analyses.

Treatment Outcome

Eighteen months after baseline, patients in all 3 settings improved remarkably in terms of psychiatric symptoms (GSI), the primary outcome measure (fig. 2; table 2). Within-group effect sizes of the uncorrected scores were 0.55 (medium effect) for outpatient psychotherapy, 0.97 (large effect) for day hospital psychotherapy, and 1.37 (very large effect) for inpatient psychotherapy.

Improvements were also observed in terms of psychosocial functioning and quality of life, as can be seen in table 2. Effect sizes for these outcome measures were somewhat smaller compared to psychiatric symptoms, with effect sizes varying between 0.64 and 0.87 for social role functioning (OQ-45), between 0.30 and 0.89 for interpersonal relations (OQ-45), and between 0.37 and 0.80 for quality of life (EQ-5D).

Improvement between baseline and 18 months proved to be significant for patients within all setting groups on all 4 outcome measures ($p < 0.05$).

Group Comparisons

After correction for observed pretreatment differences by means of the multiple propensity score, the differences in improvement of psychiatric symptoms between outpatient and day hospital treatment and between day hospital and inpatient treatment were rather small, with $\beta = 0.11$ ($p = 0.44$) and $\beta = 0.18$ ($p = 0.14$), respectively

Table 2. Uncorrected outcomes (mean \pm SD) and effect sizes in 3 treatment groups for all outcome variables

Variable	Treatment group	Baseline	18 months	Within-group effect size, Cohen's d
GSI	outpatient (n = 46)	1.52 \pm 0.78	1.10 \pm 0.75	0.55
	day hospital (n = 81)	1.74 \pm 0.62	1.09 \pm 0.72	0.97
	inpatient (n = 80)	1.94 \pm 0.66	1.02 \pm 0.69	1.37
OQ-45 Social role	outpatient (n = 46)	15.63 \pm 4.35	12.75 \pm 4.70	0.64
	day hospital (n = 81)	16.40 \pm 4.46	12.89 \pm 4.68	0.77
	inpatient (n = 80)	17.33 \pm 5.33	12.83 \pm 5.05	0.87
OQ-45 Interpersonal relations	outpatient (n = 46)	21.30 \pm 7.45	19.01 \pm 7.89	0.30
	day hospital (n = 81)	21.34 \pm 5.99	17.65 \pm 6.41	0.60
	inpatient (n = 80)	23.59 \pm 5.94	17.92 \pm 6.83	0.89
EQ-5D	outpatient (n = 46)	0.57 \pm 0.28	0.67 \pm 0.26	0.37
	day hospital (n = 81)	0.47 \pm 0.27	0.66 \pm 0.26	0.72
	inpatient (n = 80)	0.50 \pm 0.27	0.70 \pm 0.23	0.80

Table 3. Comparison of change scores from baseline to 18 months after propensity score correction, all outcome variables

Variable	Treatment groups compared	β^a	95% CI	p value
GSI	outpatient (n = 46) vs. day hospital (n = 81)	0.11	-0.17 to 0.40	0.44
	outpatient (n = 46) vs. inpatient (n = 80)	0.30	-0.01 to 0.60	0.057
	day hospital (n = 81) vs. inpatient (n = 80)	0.18	-0.06 to 0.42	0.14
OQ-45 Social role	outpatient (n = 46) vs. day hospital (n = 81)	-0.79	-3.13 to 1.54	0.50
	outpatient (n = 46) vs. inpatient (n = 80)	-0.82	-3.38 to 1.73	0.53
	day hospital (n = 81) vs. inpatient (n = 80)	-0.03	-1.90 to 1.84	0.98
OQ-45 Interpersonal relations	outpatient (n = 46) vs. day hospital (n = 81)	0.89	-1.70 to 3.49	0.50
	outpatient (n = 46) vs. inpatient (n = 80)	2.09	-0.65 to 4.83	0.14
	day hospital (n = 81) vs. inpatient (n = 80)	1.19	-0.96 to 3.35	0.28
EQ-5D	outpatient (n = 46) vs. day hospital (n = 81)	0.02	-0.08 to 0.11	0.71
	outpatient (n = 46) vs. inpatient (n = 80)	0.07	-0.03 to 0.17	0.16
	day hospital (n = 81) vs. inpatient (n = 80)	0.05	-0.03 to 0.13	0.18

^a Positive coefficients indicate that the last treatment group shown is superior, negative coefficients indicate that the first treatment group is superior.

(table 3). However, the difference in improvement between outpatient and inpatient treatment proved to be marginally significant ($\beta = 0.30$; $p = 0.057$) in favour of inpatient treatment.

Group differences in the improvement of psychosocial functioning were smaller than the differences in the improvement of psychiatric symptoms (table 3). The same holds true for the differences in change scores of quality of life, with results that were far from significant.

From the inspection of the uncorrected results (i.e., without propensity score correction), it appeared that propensity score correction reduced the differences between treatment groups considerably. In other words, pretreatment differences played an important role in the differential effectiveness observed.

Discussion

Main Findings

This is the first study comparing the effectiveness of psychotherapy in 3 treatment settings in a large population of patients with cluster B PD. Patients in the outpatient, day hospital, and inpatient psychotherapy groups showed low dropout rates. They improved in all 3 treatment groups on psychiatric symptoms, psychosocial functioning, and quality of life at 18 months after baseline. Patients in the inpatient psychotherapy group showed the strongest improvement, particularly in psychiatric symptoms. This result proved to be marginally statistically significant ($p = 0.057$ for the comparison of inpatient and outpatient treatment) even when pretreatment differences of patients were controlled for with the multiple propensity score.

Strengths and Limitations

A clear strength of the present study is its external validity and clinical utility since it was conducted in regular clinical practice, not under experimental conditions [64]. Remarkably, dropout rates were low compared to other studies conducted in PD populations [65]. A second strength is the powerful statistical control of potential confounders, using the multiple propensity score methodology. Finally, a major asset of this study is its large sample size. All this enabled the comparison of outpatient, day hospital, and inpatient psychotherapeutic treatment while keeping sufficient statistical power.

Despite these strengths, the study had several limitations. First, even though we controlled for all observed pretreatment differences, it cannot be ruled out that results were influenced by unobserved confounders. To minimise this risk as much as possible, a broad range of possible confounders was carefully selected and measured, based on both clinical and empirical knowledge [62], 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 [66–71]. Only variables significantly related to treatment outcome were included in the propensity score. In the present sample (and contrary to previous research), this appeared not to be the case for previous hospitalisation and substance misuse. 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 [72], were not measured. Second, for

ethical reasons, a control group receiving no treatment at all was not included. This omission is mitigated by the fact that several previous studies have convincingly shown that specialised psychotherapeutic treatment provides better outcomes than various control conditions (for example waiting list controls) [20, 24]. Third, even though the study sample covers the whole range of cluster B PD (borderline PD, narcissistic PD, histrionic PD, and antisocial PD), the majority of patients had a diagnosis of borderline PD. This makes the conclusions most applicable to this borderline group and only to a lesser extent to the other 3 diagnostic groups. Fourth, the cut-off point between outpatient and day hospital treatment was arbitrary. We defined outpatient treatment as a low-intensive treatment with a maximum of 2 sessions of therapy per week. As soon as more therapy elements were added, we called it ‘day hospital’ treatment. This implies that our group of day hospital treatments was a heterogeneous one, varying from 1 morning/afternoon per week to 5 days per week. Therefore, conclusions about the effectiveness of specific day hospital treatments have to be drawn carefully, bearing this cut-off score in mind.

Research Implications

What are the implications of the present results for future research?

First, inpatient treatment has been largely marginalised in the literature. Only a handful of studies have investigated its usefulness, most of them with promising results [36, 73–76]. Moreover, in the last 2 decades, inpatient treatment for PD has become infrequent in clinical practice [38]. In recent practice guidelines (NICE) [77, 78] – based on scientific knowledge – inpatient treatment is only recommended as a short-term crisis intervention. In a report of the NIMH on treatment of PD, it is also explicitly stated that there are ‘no plans to extend this level of residential provision’ [79]. In the light of the present results, however, specialised inpatient treatment deserves to be considered as a valuable treatment option for patients with cluster B PD and future research should not a priori exclude this treatment setting from effectiveness studies. In a recent study on cluster C PD, short inpatient treatment also emerged as the treatment with the best results, compared to other treatment modalities [80]. On the other hand, Chiesa et al. [38] showed in a naturalistic study that specialised low-intensive treatment can equal (or even surpass) the results of traditional inpatient treatment. And indeed, in our study, we did not find a significant difference between day hospital and inpatient treatment. Therefore, future research should focus more on

the question of ‘ideal dosage’ of treatment for patients with PD. As our study covers the whole range of cluster B PD and therefore a heterogeneous group of patients in many aspects, results might be different across specific subgroups of patients [81, 82]. For instance, very vulnerable patients may not be able to tolerate the therapeutic ‘pressure’ that is often characteristic of full-time inpatient treatment or the subculture of patients’ living together [65]. This patient group might benefit much more from a long-term structured day hospital programme, for example mentalization-based treatment [17]. It is recommended to further investigate these potential matching factors. This will enable clinicians to make specific treatment recommendations for different subgroups of cluster B PD patients and to develop more sophisticated clinical practice guidelines.

Second, when randomised designs are impossible or unfeasible, important research questions have to be answered by means of non-randomised clinical trials. In these cases, researchers have to control for as many patient variables as possible in a statistically sound way. This is essential in order not to draw premature and possibly misleading conclusions.

Public Health Significance

The marginally significant corrected results make it difficult to definitely determine the best treatment setting for patients with cluster B PD. After propensity score correction, the favourable effect of inpatient psychotherapy on psychiatric symptoms is of marginal statistical significance, but still clinically worthwhile. This effect cannot be explained by the influence of dropout rates, as treatment compliance was even lower in the inpatient group (10.0% dropout rate) compared to the outpatient group (0.0% dropout rate). However, the marginal statistical significance does not allow us to draw definite conclusions. Possibly, this may be caused by a lack of power, which could be solved by an even larger sample size. The

effect might also become larger when the observation period is increased. A question for future research is therefore whether a certain treatment setting is superior over a longer-term perspective. It could be speculated that the outcome areas psychosocial functioning and quality of life have a slower pace of change compared to psychiatric symptoms and might also show a differential effectiveness of treatments at a later stage. Yet this remains conjecture and future research is needed urgently to shed more light on this question. Another future area of exploration should be cost-effectiveness. Expensive treatments can earn back the investment made and may even lead to cost savings in the long run, because patients probably consume less additional forms of care after leaving an intensive (and effective) treatment. However, in the absence of long-term cost-effectiveness data, an answer to this question cannot be provided here.

In conclusion, this study indicates that patients with cluster B PD improve in outpatient, day hospital, and inpatient psychotherapeutic treatment. In addition, we observed a small trend towards larger improvements of psychiatric symptoms in the inpatient setting compared to the outpatient setting. Future studies are needed to further investigate the effect of different treatment dosages and possible differences in treatment effectiveness for different subgroups of patients. Meanwhile, inpatient therapy should still be considered a valuable option for patients with cluster B PD and should receive as much attention as other treatment options, both in research and clinical practice.

Acknowledgements

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

References

- 1 Torgersen S, Kringlen E, Cramer V: The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry* 2001; 58:590–596.
- 2 Samuels J, Eaton WW, Bienvenu Ili OJ, Brown CH, Costa PT Jr, Nestadt G: Prevalence and correlates of personality disorders in a community sample. *Br J Psychiatry* 2002;180:536.
- 3 Zimmerman M, Rothschild L, Chelminski I: The prevalence of DSM-IV personality disorders in psychiatric outpatients. *Am J Psychiatry* 2005;162:1911–1918.
- 4 Lieb K, Zanarini MC, Schmahl C, Linehan MM, Bohus M: Borderline personality disorder. *Lancet* 2004;364:453–461.
- 5 Nestadt G, Romanoski AJ, Chahal R, Merchant A, Folsstein MF, Gruenberg EM, McHugh PR: An epidemiological study of histrionic personality disorder. *Psychol Med* 1990;20:413–422.
- 6 Ogrodniczuk JS, Piper WE, Joyce AS, Steinberg PI, Duggal S: Interpersonal problems associated with narcissism among psychiatric outpatients. *J Psychiatr Res* 2009;43:837–842.

- 7 Duggan C: A treatment guideline for people with antisocial personality disorder: overcoming attitudinal barriers and evidential limitations. *Crim Behav Ment Health* 2009; 19:219–223.
- 8 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.
- 9 Bender DS, Dolan RT, Skodol AE, Sanislow CA, Dyck IR, McGlashan TH, Shea MT, Zanarini MC, Oldham JM, Gunderson JG: Treatment utilization by patients with personality disorders. *Am J Psychiatry* 2001; 158:295–302.
- 10 Skodol AE, Gunderson JG, Shea MT, McGlashan TH, Morey LC, Sanislow CA, Bender DS, Grilo CM, Zanarini MC, Yen S, Pagano ME, Stout RL: The collaborative longitudinal personality disorders study (CLPS): overview and implications. *J Pers Disord* 2005;19:487–504.
- 11 Cramer V, Torgersen S, Kringlen E: Personality disorders and quality of life. A population study. *Compr Psychiatry* 2006;47:178–184.
- 12 van Asselt AD, Dirksen CD, Arntz A, Sevens JL: The cost of borderline personality disorder: societal cost of illness in BPD-patients. *Eur Psychiatry* 2007;22:354–361.
- 13 Soeteman DI, Hakkaart-van Roijen L, Verheul R, Busschbach JVV: The economic burden of personality disorders in mental health care. *J Clin Psychiatry* 2008;69:259–265.
- 14 de Groot ER, Verheul R, Trijsburg RW: An integrative perspective on psychotherapeutic treatments for borderline personality disorder. *J Pers Disord* 2008;22:332–352.
- 15 Zanarini MC: Psychotherapy of borderline personality disorder. *Acta Psychiatr Scand* 2009;120:373–377.
- 16 Linehan MM: *Cognitive-Behavioral Treatment of Borderline Personality Disorder*. New York, Guilford Press, 1993.
- 17 Bateman AW, Fonagy P: *Psychotherapy for Borderline Personality Disorder: Mentalization-Based Treatment*. Oxford, Oxford University Press, 2004.
- 18 Young JE, Klosko JS, Weishaar ME: *Schema Therapy: A Practitioner's Guide*. New York, Guilford Press, 2003.
- 19 Yeomans FE, Clarkin JF, Kernberg OF: *A Primer of Transference-Focused Psychotherapy for the Borderline Patient*. Lanham, Jason Aronson Inc, 2002.
- 20 Bateman A, Fonagy P: Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry* 1999;156: 1563–1569.
- 21 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.
- 22 Bateman A, Fonagy P: Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry* 2009;166:1355–1364.
- 23 Turner RM: Naturalistic evaluation of dialectical behavior therapy-oriented treatment for borderline personality disorder. *Cogn Behav Pract* 2000;7:413–419.
- 24 Koons CR, Robins CJ, Tweed JL, Lynch TR, Gonzalez AM, Morse JQ, Bishop GK, Butterfield MI: Efficacy of dialectical behavior therapy in women veterans with borderline personality disorder. *Behav Ther* 2001;32: 371–390.
- 25 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.
- 26 Davidson K, Norrie J, Tyrer P, Gumley A, Tata P, Murray H, Palmer S: The effectiveness of cognitive behavior therapy for borderline personality disorder: results from the borderline personality disorder study of cognitive therapy (BOSCOT) trial. *J Pers Disord* 2006;20:450–465.
- 27 Farrell JM, Shaw IA, Webber MA: A schema-focused approach to group psychotherapy for outpatients with borderline personality disorder: a randomized controlled trial. *J Behav Ther Exp Psychiatry* 2009;40:317–328.
- 28 Soler J, Pascual JC, Tiana T, Cebria A, Barachina J, Campins MJ, Gich I, Alvarez E, Perez V: Dialectical behaviour therapy skills training compared to standard group therapy in borderline personality disorder: a 3-month randomised controlled clinical trial. *Behav Res Ther* 2009;47:353–358.
- 29 McMMain SF, Links PS, Gnam WH, Guimond T, Cardish RJ, Korman L, Streiner DL: A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *Am J Psychiatry* 2009;166:1365–1374.
- 30 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.
- 31 Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF: Evaluating three treatments for borderline personality disorder: a multi-wave study. *Am J Psychiatry* 2007;164:922–928.
- 32 Cottraux J, Note ID, Boutitie F, Milliere M, Genouihac V, Yao SN, Note B, Mollard E, Bonasse F, Gaillard S, Djamousian D, Guillard Cde M, Culem A, Gueyffier F: Cognitive therapy versus Rogerian supportive therapy in borderline personality disorder. Two-year follow-up of a controlled pilot study. *Psychother Psychosom* 2009;78:307–316.
- 33 Hoglend P, Amlo S, Marble A, Bogwald KP, Sorbye O, Sjaastad MC, Heyerdahl O: Analysis of the patient-therapist relationship in dynamic psychotherapy: an experimental study of transference interpretations. *Am J Psychiatry* 2006;163:1739–1746.
- 34 Hoglend P, Bogwald KP, Amlo S, Marble A, Ulberg R, Sjaastad MC, Sorbye O, Heyerdahl O, Johansson P: Transference interpretations in dynamic psychotherapy: do they really yield sustained effects? *Am J Psychiatry* 2008;165:763–771.
- 35 Nadort M, Arntz A, Smit JH, Giesen-Bloo J, Eikelenboom M, Spinhoven P, van Asselt T, Wensing M, van Dyck R: Implementation of outpatient schema therapy for borderline personality disorder with versus without crisis support by the therapist outside office hours: a randomized trial. *Behav Res Ther* 2009;47:961–973.
- 36 Chiesa M, Fonagy P, Holmes J, Drahorad C: Residential versus community treatment of personality disorders: a comparative study of three treatment programs. *Am J Psychiatry* 2004;161:1463–1470.
- 37 Chiesa M, Fonagy P, Holmes J: Six-year follow-up of three treatment programs to personality disorder. *J Pers Disord* 2006;20: 493–509.
- 38 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.
- 39 Arnevik E, Wilberg T, Urnes O, Johansen M, Mosen 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.
- 40 Black N: Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215–1218.
- 41 Facchinetti F, Ottolini F, Fazzio M, Rigatelli M, Volpe A: Psychosocial factors associated with preterm uterine contractions. *Psychother Psychosom* 2007;76:391–394.
- 42 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.
- 43 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.
- 44 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.
- 45 Vervaeke GAC, Emmelkamp PMG: Treatment selection: what do we know? *Eur J Psychol Assess* 1998;14:50–59.

- 46 van Manen J, Kamphuis JH, Visbach G, Ziegler U, Gerritsen A, Van Rossum B, Rijnerse P, Timman R, Verheul R: How do intake clinicians use patient characteristics to select treatment for patients with personality disorders? *Psychother Res* 2008;18:711-718.
- 47 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.
- 48 Pfohl B, Blum N, Zimmerman M: *Structured Interview for DSM-IV Personality (SIDP-IV)*. Washington, American Psychiatric Press, 1997.
- 49 Livesley WJ, Jackson DN: *Manual for the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ)*. Port Huron, Sigma Press, 2002.
- 50 van Kampen D: The DAPP-BQ in The Netherlands: factor structure and relationship with basic personality dimensions. *J Pers Disord* 2002;16:235-254.
- 51 Verheul R, Andrea H, Berghout CC, Dolan C, Busschbach JJV, Van der Kroft PJA, Bate-man AW, Fonagy P: Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability and validity. *Psychol Assess* 2008;20:23-34.
- 52 van Beek N, Verheul R: Motivation for treatment in patients with personality disorders. *J Pers Disord* 2008;22:89-100.
- 53 Derogatis LR, Melisaratos N: The Brief Symptom Inventory: an introductory report. *Psychol Med* 1983;13:595-605.
- 54 de Beurs E, Zitman FG: De Brief Symptom Inventory (BSI): de betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. *Maandbl Geestelijke Volksgezond* 2006;61:120-141.
- 55 Derogatis LR: *SCL-90-R: Administration, Scoring and Procedure. Manual-II for the Revised Version*. Townson, Clinical Psychometric Research, 1986.
- 56 Arrindell WA, Ettema JHM: *SCL-90-R: herziene handleiding bij een multidimensionele psychopathologie-indicator*. Lisse, Swets & Zeitlinger, 2003.
- 57 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.
- 58 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 Publishers, 2003.
- 59 Cohen J: *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, Erlbaum, 1988.
- 60 Rosenbaum PR, Rubin DB: The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41-55.
- 61 Imbens GW: The role of the propensity score in estimating dose response functions. *Biometrika* 2000;87:706-710.
- 62 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.
- 63 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. *Med Care* 2010;48:166-174.
- 64 Hodgson R, Bushe C, Hunter R: Measurement of long-term outcomes in observational and randomised controlled trials. *Br J Psychiatry* 2007;191:78-84.
- 65 Chiesa M, Drahorad C, Longo S: Early termination of treatment in personality disorder treated in a psychotherapy hospital. Quantitative and qualitative study. *Br J Psychiatry* 2000;177:107-111.
- 66 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.
- 67 Plakun EM: Prediction of outcome in borderline personality disorder. *J Pers Disord* 1991;5:93-101.
- 68 Links PS, Mitton MJE, Steiner M: Stability of borderline personality disorder. *Can J Psychiatry* 1993;38:255-259.
- 69 Ryle A, Golyunkina K: Effectiveness of time-limited cognitive analytic therapy of borderline personality disorder: factors associated with outcome. *Br J Med Psychol* 2000;73:197-210.
- 70 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.
- 71 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.
- 72 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.
- 73 Bohus M, Haaf B, Simms T, Limberger MF, Schmahl C, Unckel C, Lieb K, Linehan MM: Effectiveness of inpatient dialectical behavioral therapy for borderline personality disorder: a controlled trial. *Behav Res Ther* 2004;42:487-499.
- 74 Kroger C, Schweiger U, Sipos V, Arnold R, Kahl KG, Schunert T, Rudolf S, Reinecker H: Effectiveness of dialectical behaviour therapy for borderline personality disorder in an inpatient setting. *Behav Res Ther* 2006;44:1211-1217.
- 75 Fassbinder E, Rudolf S, Bussiek A, Kroger C, Arnold R, Greggersen W, Huppe M, Sipos V, Schweiger U: Effektivität der dialektischen Verhaltenstherapie bei Patienten mit Borderline-Persönlichkeitsstörung im Langzeitverlauf - Eine 30-Monats-Katamnese nach stationärer Behandlung. *Psychother Psychosom Med Psychol* 2007;57:161-169.
- 76 Kleindienst N, Limberger MF, Schmahl C, Steil R, Ebner-Priemer UW, Bohus M: Do improvements after inpatient dialectical behavioral therapy persist in the long term? A naturalistic follow-up in patients with borderline personality disorder. *J Nerv Ment Dis* 2008;196:847-851.
- 77 National Institute for Health and Clinical Excellence: *Antisocial Personality Disorder: Treatment, Management and Prevention*. London, NICE, 2009. www.nice.org.uk.
- 78 National Institute for Health and Clinical Excellence: *Borderline Personality Disorder: Treatment and Management*. London, NICE, 2009. www.nice.org.uk.
- 79 National Institute for Mental Health in England: *Personality Disorder: No Longer a Diagnosis of Exclusion - Policy Implementation Guidance for the Development of Services for People with Personality Disorder*. London, DoH, 2003.
- 80 Bartak A, Spreeuwenberg MD, Andrea H, Holleman L, Rijnerse P, Rossum BV, Hamers EFM, Meerman AMMA, Aerts J, Busschbach JJV, Verheul R, Stijnen T, Emmelkamp PMG: Effectiveness of different modalities of psychotherapeutic treatment for patients with cluster C personality disorders: results of a large prospective multicentre study. *Psychother Psychosom* 2010;79:20-30.
- 81 Lorentzen S, Hoglend PA: Moderators of the effects of treatment length in long-term psychodynamic group psychotherapy. *Psychother Psychosom* 2008;77:321-322.
- 82 Digre EI, Reece J, Johnson AL, Thomas RA: Treatment response in subtypes of borderline personality disorder. *Pers Ment Health* 2009;3:56-67.