

**Effectiveness of Short-Term Inpatient Psychotherapy based on Transactional
Analysis in Patients with Personality Disorders:
A Matched Control Study using Propensity Score**

Running title:

Effectiveness of Short-Term Inpatient Psychotherapy in PD patients

Eva K. Horn ^{a, b} Roel Verheul ^{c, d} Moniek Thunnissen ^{e, f} Jos Delimon ^c Mirjam Soons ^g Anke M.
M. A. Meerman ^h Uli M. Ziegler ⁱ Bert V. Rossum ^j Helene Andrea ^{k, l} Theo Stijnen ^m Paul M. G.
Emmelkamp ^{d, n} Jan J. V. Busschbach ^{a, b}

^a Viersprong Institute for Studies on Personality Disorders (VISPD), Bergen op Zoom; ^b Department of Psychiatry, section Medical Psychology and Psychotherapy, Erasmus Medical Centre, Rotterdam; ^c De Viersprong Netherlands Institute for Personality Disorders, Halsteren; ^d Department of Clinical Psychology, University of Amsterdam, Amsterdam; ^e GGZ WNB, Bergen op Zoom; ^f Private Practice, Bergen op Zoom; ^g PuntP, Amsterdam; ^h Centre of Psychotherapy Pro Persona, Lunteren; ⁱ Zaans Medical Centre, Zaandam; ^j Altrecht, Zeist; ^k GGZ Breburg, Tilburg; ^l Tranzo Scientific Centre for Care and Welfare, Faculty of Social and Behavioral Sciences, Tilburg University; ^m Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden; ⁿ The Center for Social and Humanities Research, King Abdulaziz University, Jeddah, Saudi Arabia

Corresponding author:

Eva K. Horn

Viersprong Institute for Studies on Personality Disorders (VISPD)

P.O. Box 7

4660 AA Halsteren

The Netherlands

Phone: +31-164-632 200

Fax: +31-164-632 220

Eva.Horn@deViersprong.nl

Abstract

Background

Controlled studies on the effectiveness of inpatient psychotherapy in personality disorders (PD) are rare. This study aims to compare 3-month short-term inpatient psychotherapy based on Transactional Analysis (STIP-TA) with other psychotherapies (OP) up to 36 months follow-up.

Methods

PD patients following STIP-TA were matched with OP patients using the propensity score. Primary outcome measure was general psychiatric symptomatology; secondary outcomes were psychosocial functioning and quality of life.

Results

In 67 pairs of patients, both STIP-TA and OP showed large symptomatic and functional improvements. However, STIP-TA showed more symptomatic improvement at all time points as compared to OP. At 36 months, 68% of STIP-TA patients were symptomatically recovered compared to 48% in OP.

Conclusions

STIP-TA outperformed OP in terms of improvements in general psychiatric symptomatology and quality of life. Superiority of STIP-TA was most pronounced at 12-month follow-up, but remained intact over the course of the three-year follow-up.

Declaration of Interest

None.

Keywords: personality disorders, cluster C personality disorder, personality disorder not otherwise specified, psychotherapy outcome research, treatment effectiveness, long-term outcome, inpatient treatment, transactional analysis

Introduction

There is now compelling evidence to support the hypothesis that personality disorders are treatable and changeable (e.g. Arnevik, Wilberg, Urnes, Johansen, Monsen, & Karterud, 2010; Binks, Fenton, McCarthy, Lee, Adams, & Duggan, 2006; Dixon-Gordon, Turner, & Chapman, 2011; Gabbard, 2000; Leichsenring & Leibing, 2003; Perry, Banon, & Ianni, 1999; Stoffers, Vollm, Rucker, Timmer, Huband, & Lieb, 2012), and psychotherapy is considered the treatment of first choice (e.g. Bartak, Soeteman, Verheul, & Busschbach, 2007; Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ, 2008; National Institute for Health and Clinical Excellence, 2009a, 2009b). Although most studies concern well-known treatments for borderline personality disorder (BPD), such as Schema-Focused Therapy (SFT), Transference-Focused Psychotherapy (TFP), Dialectical Behaviour Therapy (DBT), and Mentalization-Based Treatment (MBT) (e.g. Dixon-Gordon, Turner, & Chapman 2011; Gabbard, 2000), a steadily growing number of studies document psychotherapy's effectiveness in other PDs as well. In the past ten years, several reviews were done on the effectiveness of psychotherapeutic treatments in general PD patients. Positive results were found concerning psychodynamic psychotherapy (shorter than 41 weeks) and cognitive behavioural therapy (Leichsenring & Leibing, 2003; Matusiewicz, Hopwood, Banducci, & Lejuez, 2010; Town, Abbass, & Hardy, 2011; Verheul & Herbrink, 2007), and rather limited and conflicting evidence concerning long-term psychoanalytic psychotherapy (Smit, Huibers, Ioannidis, van Dyck, van Tilburg, & Arntz, 2012). Remarkably, the treatment literature on PD typically considers long-term treatments in an outpatient setting. An interesting and viable alternative might be Short-Term Inpatient Psychotherapy (STIP). For example, Bartak, Spreeuwenberg, Andrea, Holleman, Rijnerse, Rossum, Hamers, Meerman, Aerts, Busschbach, Verheul, Stijnen, and Emmelkamp (2010) and Soeteman, Verheul, Meerman, Ziegler, Rossum, Delimon, Rijnerse, Thunnissen, Busschbach, and Kim, (2011) have recently shown superior effectiveness and efficiency of STIP (with a maximum duration of six months) as compared to other treatment modalities in patients with a cluster C PD. In that particular study, STIP included a variety of durations (e.g. both 3-month and 6-month psychotherapy) and theoretical orientations (e.g. cognitive-behavioural and psychodynamic treatment). The current study focuses on the effectiveness of the shortest, 3-month variant, i.e. STIP based on Transactional Analysis (STIP-TA). Thunnissen, Duivenvoorden, Busschbach, Hakkaart-van Roijen, van Tilburg, Verheul, and Trijsburg (2008) reported on treatment outcome of STIP-TA in a naturalistic cohort study

and found an effect-size of 2.0 two years after start of the treatment and a recent review of studies on TA in psychotherapy found a significant positive effect of TA in more than 80% of the studies reviewed (Ohlsson, 2010). This study will compare the effectiveness of STIP-TA to other specialized psychotherapies (OPs) in PD patients in a non-randomized, but nevertheless rigorously controlled trial.

Methods

Study population and design

Patients were recruited from March 2003 to March 2006 from a consecutive series of admissions to six mental health care institutes in the Netherlands offering specialized psychotherapy for adult patients with PD. In this period of time, 1,379 patients completed the intake procedure. Of these, 837 had a DSM-IV-TR PD diagnosis (American Psychiatric Association, 2000), provided informed consent and were enrolled in the Study on Cost-Effectiveness of Personality Disorder Treatment (SCEPTRE), a large project in the Netherlands focusing on the effectiveness and cost-effectiveness of psychotherapeutic treatments in PD patients (Bartak et al., 2010). After informed consent, 70 patients refused to participate and 38 patients dropped out prematurely. Twenty-eight patients were lost to follow-up, seven patients could not participate due to illness or death. This left 694 patients of which 71 received STIP-TA at De Viersprong, Netherlands Institute for Personality Disorders, and 623 patients receiving other specialized psychotherapies (OP). It was not possible to find a good match for four STIP-TA patients, therefore, 67 patients were matched 1-by-1 by the logit of the propensity score, which left 134 patients for this trial (see figure 1).

--insert figure 1 about here--

Handling data

All outcome measures were assessed at baseline and several follow-up points. Various follow-up points had to be used due to logistic reasons: three treatment centers had their assessments at baseline, end of the treatment, and six and 12 months after the end of treatment, and at 36 and 60 months after baseline, whereas three other centers had their assessments at baseline, at 12, 24, 36,

and 60 months after baseline. Ninety-five percent of the STIP-TA patients had three or more follow-up measurements, as compared to 88% of the matched OP patients.

Statistical analyses

The matching of patients was done by matching each patient 1-to-1 on the logit of the propensity score (Rosenbaum & Rubin, 1985). Both, the propensity score and the matching are explained below.

The propensity score (PS) is defined as the conditional probability of assignment to one of two treatment groups given a set of observed pre-treatment variables (Bartak, Spreeuwenberg, Andrea, Busschbach, Croon, Verheul, Emmelkamp, & Stijnen, 2009; Rosenbaum & Rubin, 1983). To estimate the PS, we fitted one logistic regression model with group membership (STIP-TA or OP) as outcome. In order to create two similar samples at baseline, relevant confounders which were related to the outcome variable were used as independent variables (Brookhart, Schneeweiss, Rothman, Glynn, Avorn, & Sturmer, 2006). Fifty-four sociodemographic and clinical variables were selected as potential confounders based on clinical knowledge and a literature review. Of these, 27 variables had a significant influence on the outcome and were used in the estimation of the PS. Sociodemographic variables (e.g. age, sex, level of education) and clinical variables (e.g. motivation, baseline scores of the outcome measures, SIPP-118 scales, PD diagnoses) were included (a full list of these variables is provided in the appendix). Additionally, indirect and direct medical costs in a 12-month period before baseline were added as potential confounders in order to create a sample which can be used for future cost-effectiveness analyses. Since the logit of the PS is more likely to be normally distributed than the propensity score, matching was done on the logit of the PS. The logit is the logarithm of the PS divided by one minus the PS. In order to receive optimal balance between the treatment groups, a caliper distance of .2 of the standard deviation (SD) of the logit of the PS was used (Austin, 2011).

The matching was done by first ordering subjects of the STIP-TA group randomly. The control subject with 1) the smallest distance of the PS logit with the STIP-TA subject and 2) which PS logit fell within the caliper width was identified as match. This pair was then removed from the pool and a match was sought for the next STIP-TA subject. When no match could be found with a PS logit within the caliper, that STIP-TA subject was removed.

Multilevel models were used for the analysis of outcomes over time. We used multilevel modelling to deal with: (1) the dependency of the repeated measures within the same subject in time, (2) the dependency of patients within a pair, and (3) the longitudinal data with observations unequally spaced in time. The levels were (1) time within patients, (2) patients within pairs and (3) pairs. To estimate the treatment effects at 12, 24 and 36 months after baseline, we used random effects for pairs, patients within pairs and time within patients. We used a random intercept model with time as level I and pair number as level II and a second model with a random intercept and random slope with time as level I and patient number as level II. In addition to a linear time effect, we postulated knots (or splines) every six months which allowed the estimated course of the dependant variable to bend at these time points. Non-significant knots ($p < 0.05$) were deleted from the models until a parsimonious model was reached that did not differ significantly from the original saturated model. This resulted in a final best fitting model with the change scores (from baseline) observed during follow-up for each of the outcome measures as dependent variables and the following independent variables: group membership, time, interaction between group membership and time, the deviation of the overall baseline score (in order to take the baseline scores into account), a linear spline (knot point at 18 months), and interaction of the spline with the treatment group. Subsequently, we calculated within-group effect sizes (Cohen's d) to describe change from baseline to 12, 24, and 36 months in both groups (Cohen, 1988). We used the estimated pooled SDs combining the baseline SD with the follow-up SDs from the models.

The analyses of outcomes are based on the intention-to-treat (ITT) principle. All patients who received a 'minimal effective dosage' (defined as at least two sessions of outpatient psychotherapy or at least two treatment days of day hospital or inpatient psychotherapy) were followed and included in the analyses. The analyses are performed using IBM SPSS 20.0 (IBM Corporation, USA) for data preparation and SAS 9.2 (SAS Institute Inc, USA) for multi-level modeling.

Assessments

An extensive standard assessment battery of instruments was administered to the patients before treatment assignment.

Classification of PD

PDs were classified using the Dutch version of the Structured Interview for DSM-IV Personality (SIDP-IV) (De Jong, Derks, van Oel, & Rinne, 1996; Pfohl, Blum, & Zimmerman, 1997). This interview covers the 11 formal DSM-IV-TR Axis II diagnoses including PD Not Otherwise Specified (PDNOS). A PDNOS diagnosis could consist of either an appendix diagnosis (i.e. depressive, passive-aggressive PD, or self-defeating PD) or a mixed PD (meeting ten or more diagnostic criteria of various PDs). Interviewers were master-level psychologists, who were trained thoroughly, and who received monthly booster sessions to avoid deviation from the interviewer guidelines.

Main outcome measures

The primary outcome measure was general psychiatric symptomatology, expressed by the Global Severity Index (GSI) and measured by the Dutch version of the Brief Symptom Inventory (BSI). The BSI is a brief self-report questionnaire which covers nine symptom dimensions and is developed from the Symptom Checklist 90 - Revised (SCL-90-R) (Arrindell & Ettema, 2003; Derogatis, 1983). The GSI is calculated as the mean score of the 53 items of the BSI, ranging from zero to four (De Beurs & Zitman, 2006; Derogatis, 1992; Derogatis & Melisaratos, 1983).

As secondary outcomes, psychosocial functioning was measured using two subscales of the Outcome Questionnaire-45 (OQ-45), i.e. Interpersonal Relations and Social Role (Lambert, Morton, Hatfield, Harmon, Hamilton, Reid, Shimokawa, Christopherson, & Burlingame, 2004), and health-related quality of life was measured using the EQ-5D (EuroQolGroup, 1995). A study in the Netherlands measured and valuated the EQ-5D, resulting in the Dutch EQ-5D value set, which is used to calculate utilities for EQ-5D health states (Lamers, Stalmeier, McDonnell, Krabbe, & van Busschbach, 2005).

To evaluate clinical significant change at 36 months after baseline in terms of the GSI and the OQ-45, we followed the criteria by Jacobson and Truax (1991). The cut-off and reliable change index of the manual of the BSI (Derogatis, 2011) and of the article on the validation of the OQ-45 in a Dutch population (De Jong, Nugter, Polak, Wagenborg, Spinhoven, & Heiser, 2007), respectively, were used. First, we computed the percents of patients in each treatment group who achieved reliable change or clinical deterioration. Reliable change was defined as scores which decreased by the reliable change

index and clinical deterioration was defined as scores which increased by the reliable change index. Second, we computed the percentage of patients who moved from a dysfunctional range to a normative range. Finally, the percentage of patients who achieved both reliable change and who moved to a normative range were computed. These were considered as symptomatically recovered. Differences between treatment groups were analyzed using McNemar's test.

Additional baseline measures

Three additional self-report instruments to measure patient characteristics were included in the assessment battery at baseline and were used as potential confounders for the PS. First, the Dutch version of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ) measures type and degree of personality pathology (Livesly, 2002; van Kampen, 2002). We used patients' scores on this questionnaire on the four higher order factors: emotional dysregulation, dissocial behavior, inhibition, and compulsivity. Second, to measure patients' motivation for treatment, the two scales of the Motivation for Treatment Questionnaire (MTQ-8) were used: need for help and readiness to change (van Beek & Verheul, 2008). Third, the Severity Indices of Personality Problems-118 (SIPP-118) is a self-report questionnaire aimed to measure five core domains of personality pathology, i.e. self control, identity integration, responsibility, relational functioning, and social concordance (Verheul, Andrea, Berghout, Dolan, Busschbach, van der Kroft, Bateman, & Fonagy, 2008).

Interventions

STIP-TA

Transactional analysis is a theory of personality development, intrapsychic functioning, and interpersonal behavior developed by Eric Berne in the 1950s. It is based on psychoanalytical ideas and integrates elements of ego psychology, object relations theory and learning theory (Barnes, 2007). Aim of TA is to think, feel, and act differently by changing old patterns of behavior. TA is based on the observation that psychic functioning and social behavior are related to states of mind which are called ego states. These ego states are used to explain how individuals relate to themselves and to each other. The basic ego states consist of the extero-psychic (i.e. the 'Parent'; based on thoughts and

feelings copied from parents or parental figures), the neopsychic (i.e. the 'Adult', relates to thoughts and feelings in the here-and-now), and the archaeopsychic (i.e. the 'Child'; based on thoughts and feelings replayed from childhood) component (Berne, 1996). Although TA does not explicitly state to work on mentalizing and metacognition, these concepts are extensively addressed in TA (Hawkes, 2011). In the theory of TA, difficulties in metacognition and mentalizing are described as 'lacks in Adult functioning', whereby the Adult is the pattern of feeling, thinking and behavior characterized by adequate reality-testing and functioning in the here-and-now. A general aim of TA psychotherapy is to function more from an 'integrated Adult'-state, meaning that the person can undertake grown-up responsibilities, shows loyalty and reliability and is aware of his or her inner emotional states.

STIP-TA had incorporated the ideas of transactional analysis and was carefully developed as a 13 weeks inpatient psychotherapy program, divided in two halves of six weeks, with one week in between. The treatments groups were half-open: every six weeks a week off treatment was scheduled for all patients during which four patients who have completed treatment were discharged and four new patients started treatment. In the first period of six weeks, patients were 'junior' group members. They made their treatment contract and experienced the inpatient setting, being part of a group, and living together with seven other patients in one house. In order to create the individual treatment contract, a core conflict was determined and recorded which guided the treatment. In this contract, patients stated in which way they wanted to change their patterns of thinking, feeling, and behavior. The treatment contract was described in collaboration of the therapist with the patient oneself and was not valid until both therapist and patient agreed on it. The contract was an important part of the therapy as it guided the treatment. At the end of the treatment, therapists, patients, and the treatment group assessed whether and to what extent this core conflict was resolved.

STIP-TA was specifically designed for and tailored to the needs of patients with various personality disorders, particularly cluster C and personality disorder not otherwise specified. It usually involved patients with childhood traumatic experiences, such as severe illness, disability or death (sometimes by suicide) of a parent, emotional or physical neglect, or sexual or physical abuse. Their frame of reference is often severely shaken and they start realizing their self-restricting patterns of thinking, feeling and behaving. The language of TA helps them to understand how they function interpersonally, and start to see connections between the external and their internal world and the

roots of their patterns, originating in their childhood experiences. In this way, patients create a narrative with which they can understand their current problems in functioning as a result of their course through life. In the second part of the treatment specific therapeutic strategies, either in the group psychotherapy or in the non-verbal therapies, are used to lift blockades and to have corrective emotional experiences which help them to change their maladaptive patterns.

Specific for STIP-TA were the following nine characteristics, many of which correspond to the common features for effective psychotherapies as proposed by Bateman and Fonagy (Bateman & Fonagy, 2000): One shared vision which is carried out by the staff; one shared language; shared responsibility of patients and staff; mutual responsibility of patients; working with contracts on a focal spectrum of problems; structure in time, place and person; continuous evaluation of progress; active and well trained therapists (Delimon, 1999). STIP-TA included group psychotherapy, psychomotor and art therapy, sociotherapy, and milieu therapy. As the team of therapists (psychotherapist, nurses and non-verbal therapist) works very closely together and shares their experiences in working with the patient group two times a day, patients are followed carefully in their process. Non-verbal therapies like art- and movement therapy are included as for some of the patients, these are better entrances to explore their dysfunctional patterns compared to verbal therapies (Delimon, 1999; Thunnissen, 2007).

Other psychotherapies

The specialized psychotherapies consisted of treatments which varied widely in terms of setting (i.e. outpatient, day hospital, and inpatient), duration (i.e. 3-month to 36-month psychotherapy), and theoretical orientation (e.g. cognitive-behavioral and psychodynamic treatment). The final matched OP group consisted of six patients (9%) following short-term and 18 patients (27%) following long-term outpatient therapy (i.e. individual or group psychotherapy sessions, up to two sessions per week), seven patients (10%) following short-term and 11 patients (16%) following long-term day hospital therapy (i.e. at least one morning/afternoon per week, various forms of psychotherapeutic and psychosocial treatments, but sleeping at home), and nine patients (13%) following short-term and 16 patients (24%) following long-term inpatient therapy (i.e. staying at the institutions up to five days a week, various forms of psychotherapeutic and psychosocial treatments). Day hospital and inpatient programs typically consisted of group psychotherapy as a core element, mostly in combination with

one or more non-verbal or expressive group therapies (such as psychomotor and art therapy), individual psychotherapy, sociotherapy, milieu therapy, community meetings, and/or pharmacological treatment. Twenty-two percent of treatments had a cognitive-behavioural orientation, 30% had a psychodynamic orientation, 42% had an integrative orientation and the remaining 5% had a not specified orientation. The psychotherapists were all licensed psychiatrists or psychologists. All treatments under study were tailored to patients with personality pathology and can be considered highly representative of specialist mental health care in the Netherlands.

Treatment adherence

Since the Sceptre study initially focused on treatment dosage only, intended and realized treatments were described in terms of setting (outpatient, day hospital or inpatient psychotherapy) and duration of the treatments (short- or long-term). At the end of the treatment, 100% of the STIP-TA patients had completed the intended treatment dosage, compared to 85% of the OP ($\chi^2=10.81$, $df=1$, $p=.001$). Of the patients who changed treatment dosage, the deviation of the intended dosage was mainly due to the duration of treatment: 50% received a shorter treatment, while the remaining 50% received a longer treatment. One OP patient (2%) also changed the treatment setting.

Results

Sample

The sociodemographic and clinical characteristics of STIP-TA and OP patients are displayed in table 1. As expected, differences between the two groups were not significant (t-tests and χ^2 -tests). Patients in both groups were about 40 years of age, approximately two-thirds were female, and about half of the patients lived together with a partner. About 60% of both groups worked or was studying. Most patients were diagnosed with a cluster C PD (49% of STIP-TA, 39% of the OP group) and/or a PDNOS (42% STIP-TA, 49% OP). The largest part of the sample had a history of outpatient treatment.

--insert table 1 about here--

Figure 2 displays the distribution of the estimated probabilities of assignment to STIP-TA (propensity scores) before and after matching. After matching, the STIP-TA and the OP group were virtually equal with respect to the distribution of the propensity scores with STIP-TA having a propensity score of .259 (SD=.18), and matched controls of .260 (SD=.18), a non-significant difference ($t(66)=.26$; $p=.797$).

--insert figure 2 about here--

Treatment Outcome

At all time points (12, 24, and 36 months), both groups showed significant symptomatic improvements and large effect sizes (ES) (Cohen, 1992). Nevertheless, STIP-TA outperformed the OP group significantly at all time points, especially at twelve months after start of treatment ($b=.35$, $p<.001$; see table 2). Effect-sizes at 12 months follow-up were 2.02 for STIP-TA and 1.18 for OP. At 36 months, STIP-TA still outperformed OP ($b=.21$, $p=.0082$; see table 2) with effect-sizes of 1.93 for STIP-TA and 1.39 for OP. Figure 3 shows the course of general psychiatric symptomatology over 36 months.

--insert table 2 and figure 3 about here--

At 36 months after baseline, 44 pairs of patients (66%) had a 36-month follow-up measurement and could be used for analyses of clinically significant change. Based on the criteria outlined by Jackson and Truax (1991), 90% of the STIP-TA patients and 71% of the OP patients showed a reliable change ($p=.03$), and 0% of STIP-TA patients and 4% of OP patients demonstrated clinical deterioration.¹ Sixty-eight percent of STIP-TA patients and 48% of OP patients moved from a dysfunctional range to a normative range ($p=.08$). All patients who moved from a dysfunctional range to a normative range also demonstrated reliable change.

Two additional sensitivity analyses were done to further examine the data. First, a sensitivity analysis was done on the severity of PD pathology. Patients scoring 15 or less PD criteria (median number of traits) were considered less severe, patients scoring higher than 15 traits were considered more severe. Up to three years, no significant differences were found on the GSI between more or

¹ McNemar's test can be carried out only for a 2x2 table, where p must be greater than 1.

less severe patients in STIP-TA or OP on neither follow-up point. Differences in recovery at 36 months follow-up between high or less severe patients were also negligible. Second, a sensitivity analysis on the comparison of STIP-TA with different dosages in the OP group indicated that part of the superior effect of STIP-TA could be attributed to the lower effectiveness of inpatient treatments in the OP sample.

Concerning the secondary outcomes, psychosocial functioning (OQ-45) and health-related quality of life (EQ-5D), we found significant improvements from baseline to 36 months in both groups as well. Medium to large effect sizes were present.

Regarding psychosocial functioning, STIP-TA showed larger but non-significant ESs than OP on all scales at all time points, except for one (OQ-45 Interpersonal Relations, 24 months). Concerning the OQ-45 Interpersonal relations, at 36 months after baseline, 47 pairs of patients (70%) had a 36-month follow-up measurement. 51% of the STIP-TA patients and 34% of the OP patients showed a reliable change ($p=.17$) and 2% and 2%, respectively, demonstrated clinical deterioration ($p=1.00$). Thirty-eight percent of STIP-TA patients and 30% of OP patients moved from a dysfunctional range to a normative range ($p=.56$). Of these, 32% and 21% respectively moved from a dysfunctional range to a normative range and demonstrated reliable change ($p=.38$).

With regard to the OQ-45 Social Role, 46 pairs of patients (69%) had a 36-month follow-up measurement. We found that 50% of STIP-TA and 24% of OP showed a reliable change ($p=.01$). None of STIP-TA patients and 2% of OP patients demonstrated clinical deterioration. Fifty percent of STIP-TA and 37% of OP patients moved from a dysfunctional range to a normative range ($p=.26$) and of these, 30% of STIP-TA and 20% of OP moved from a dysfunctional range to a normative range and demonstrated reliable change ($p=.30$).

Regarding quality of life, significant larger effect sizes were found for STIP-TA compared to OP at 12 and 24 months, but not at 36 months (see table 2). The STIP-TA group approached the level of quality of life observed in the normal population (which was defined as a value of 0.88) at 12

months, with a small decline at 24 and 36 months, while the OP group stayed behind (Busschbach, Wolffenbuttel, Annemans, Meerding, & Koltowska-Haggstrom, 2011).

Significant differences between patients who completed their intended treatment dosage and patients who did not were found on the two outcome measures assessing psychosocial functioning at 12 months follow-up (OQ-45 Social Role: $b=3.46$, $p=.03$, OQ-45 Interpersonal Relations: $b=5.54$, $p=.02$). Patients who completed the intended treatment dosage showed significantly better psychosocial functioning than patients who did not.

Discussion

The current study showed that time-limited, short-term inpatient psychotherapy based on transactional analysis outperformed a variety of other specialized psychotherapies in terms of improvements in general psychiatric symptomatology and quality of life. Superiority of STIP-TA was most pronounced at 12-month follow-up, but remained intact over the course of the entire three-year follow-up period. These results support and extend findings of previous studies of STIP-TA (Thunnissen et al., 2008) and STIP in cluster C patients (Bartak et al., 2010), which strengthens the conclusions we can draw from this study.

In terms of general psychiatric symptomatology we found effect-sizes of 1.87 to 2.02 for STIP-TA patients and 1.18 to 1.39 for OP patients across three years follow-up. Our assessment of clinically meaningful change at 36 months after baseline showed a small but insignificant advantage in favour of STIP-TA with 68% of STIP-TA patients and 48% of OP patients being symptomatically recovered (demonstrating reliable change and movement from a dysfunctional range to a normative range). Other studies on reliable change or recovery rates of PD patients showed varying results. For example, Muran, Safran, Samstag, and Winston (2005) compared an outpatient, 20-session alliance-focused treatment (brief relational therapy) to short-term dynamic therapy and a cognitive-behavioral therapy in cluster C and PDNOS patients. All three treatments proved to be equally effective at six months follow-up in terms of symptoms and ten to 31% of patients showed reliable change. The study of Svartberg, Stiles, and Seltzer (2004) in Cluster C patients found higher recovery rates of 42% and 54% at two years follow-up after outpatient treatments. In cluster B, response or reliable change rates of 38-80% were found (Chiesa & Fonagy, 2003, McMain, Guimond, Streiner, Cardish, & Links, 2012;

Kröger, Harbeck, Armbrust, & Kliem, 2013). Thus, results of the current study showed higher effect sizes than other published effect studies of cluster C patients, and were within the range of rates found in studies of cluster B patients. However, varying results can be due to differences in effectiveness of treatments, to the follow-up period, as well as to patient characteristics such as the severity of personality pathology. Moreover, no constant definition is used for response or recovery which further hampers a comparison of these figures.

One could assume that the degree of severity of personality pathology influenced results in this study in that more severe patients showed less response than less severe patients. However, additional analyses did not find significant differences on the GSI between more or less severe patients in STIP-TA or OP on neither follow-up point up to three years. Differences in recovery at 36 months follow-up between high or less severe patients were also negligible.

Concerning the secondary outcome measures, we found that STIP-TA patients approached the level of quality of life observed in the normal population, while the OP patients lagged behind. At 36 months, significant differences in quality of life between the two groups found earlier, vanished. Less pronounced results were found concerning psychosocial functioning. About one third of STIP-TA patients and one fifth of OP patients could be considered recovered after 36 months. One could speculate whether short-term treatments are sufficient to achieve structural change which would be visible in improved social functioning. However, the limited impact of treatment on social functioning found in the current study is consistent with other studies showing that the treatment of PD has limited impact on social functioning (e.g. Arnevik et al., 2010; Chiesa & Fonagy, 2003; Skodol, Pagano, Bender, Shea, Gunderson, Yen, Stout, Morey, Sanislow, Grilo, Zanarini, & McGlashan, 2005; McMain et al., 2012). As these studies mainly studied long-term treatments, and about two-thirds of the OP treatments in the current study were considered long-term, the length of treatment may not be the main reason why we found lower effects in this area. The results rather support the hypothesis that maladaptive interpersonal patterns are core features of PDs and improve more slowly than psychiatric symptomatology. More research is necessary to study which variables need to be targeted in psychotherapy to receive better results and to be able to tailor treatments to the needs of individual PD patients (Dimaggio, Nicolò, Semerari, & Carcione, 2013). Chiesa and Fonagy (2003) suggested the following reasons for the slow pace of improvement: (1) social adjustment refers to a more complex dimension that is more difficult to change, (2) social adjustment could be “intrinsically linked to the

concept of PD, with the notion of durability and persistence”, (3) parts of social adjustment could be heavily dependent on external factors like unemployment due to economic circumstances, and (4) instruments used could be unsuitable in PD patients.

Methodological considerations

In light of possible limitations of this study, several issues have to be acknowledged: First, this study is not a randomized controlled trial (RCT), but a matched-control study which merely mimics a RCT. However, this can be considered a limitation as well as a strength. RCTs are generally criticized for their limited external validity: treatments take place under strictly controlled or experimental circumstances, only patients are studied who agree to randomization, and exclusion criteria are typically stringent (e.g. Hodgson, Bushe, & Hunter, 2007). In contrast, the treatments in the current study can be considered highly representative of specialist mental health care in the Netherlands (Bartak et al., 2010). Furthermore, this limitation is somewhat mitigated by our rigorous matching procedure that ensures the similarity of the patient groups and partly rules out the possibility of selection bias. Nevertheless, since we matched patients on the PS which was computed using predetermined and observable variables, it is possible that other patient differences that were either not predetermined or not observable affected assignment to treatment and confounded the observed differences in treatment effectiveness (Austin, 2008). The likelihood of confounding, however, is mitigated by the fact that we tested a large number of possible confounding variables.

Second, the generalizability of this study to a general PD population is limited due to a threefold aspects: First, 91% of patients were diagnosed with a cluster C PD and/or a PDNOS diagnosis, second, only PD patients who were seeking treatment and who were admitted for treatment were studied, and third, almost all patients already had a history of outpatient treatment.

Third, we only used self-report instruments as outcome measures. We do not have information whether the treatments were also able to change the PD diagnosis of patients or whether therapists considered the change sufficient.

Fourth, we do not have information about the treatment fidelity and the adherence of the treating psychotherapists for neither groups. However, we assumed that treatments delivered in specialized psychotherapeutic institutes by experienced psychotherapists are of sufficient quality. We know that in STIP-TA there is considerable effort to maintain treatment integrity, whether or not this successful, we

do not know. Furthermore if we had monitored the adherence of treating psychotherapists, other psychotherapies could not be considered treatment as usual anymore, as therapists might change their attitude and therapy.

Fifth, the OP group consisted of different treatments in terms of modality, as well as in terms of theoretical orientation which reduced the internal validity as we do not know which treatment is less effective than STIP-TA. However, as this heterogeneous mix of treatments is a good representation of the treatments as usual offered to PD patients (in the Netherlands), this fact heightens the external validity of this study. As inpatient or residential psychotherapeutic treatments in PD patients are not part of treatment as usual in other countries, and part of the superior effect of STIP-TA could be attributed to the lower effectiveness of inpatient treatments in the OP sample, the generalizability of the results to other countries is limited.

Implications and future directions

The results of our study are consistent with several studies showing that psychological treatments tailored to PD are generally very effective. Effect sizes of STIP-TA are even larger as compared to those typically observed in previous studies (e.g. Arnevik et al., 2010; Leichsenring & Leibling, 2003). This study therefore strongly suggests that STIP-TA is a promising treatment for patients with cluster C PD and PDNOS. However, since STIP-TA patients are hospitalized during 13 weeks, it is an expensive treatment, which raises the question whether the costs are worth the effects. These high initial costs may be an obstacle in the reimbursement of this treatment in some countries. Nevertheless, Soeteman et al. (2011) showed that in cluster C PDs, STIP seemed to be most cost-effective compared to other treatment modalities with different theoretical orientations and varying from short-term outpatient to long-term inpatient treatments. A cost-effectiveness study comparing STIP-TA to OP adds more evidence to the knowledge of effective and cost-effective treatments in PD patients and will be performed in the future.

The interpretation of the results is limited by the variation of treatment modalities in the OP condition and this study therefore does not clarify the observed superiority of STIP-TA. For example, it is not at all clear from this study whether the observed superiority of STIP-TA is best accounted for by either setting (inpatient), duration (short-term), psychotherapeutic orientation (transactional analysis), a combination thereof, or even another factor such as the consistent application of the theoretic

framework which has been hypothesized to be one of the general ingredients of effective treatments (Bateman & Fonagy, 2000). For instance, we found that almost half of OP patients received psychotherapies with an integrative character, thus therapies which used different psychotherapeutic orientations which is in sharp contrast to the STIP-TA treatment in which much effort is done to comply with the concepts of its framework (TA). On the other hand, studies just point to generic effective principles of change in psychotherapeutic treatments of PD. Castonguay and Beutler (2006) described general treatment principles common to all psychotherapeutic treatments and distilled three different variables as potential factors related to outcome: participant characteristics (e.g. resistance or coping style), therapeutic relationship variables, and technical factors (e.g. treatment intensity or the level of focus of therapy). Notably, the quality of the therapeutic relationship seemed to be as important as the specific treatment method in terms of treatment outcome. Elements such as the therapeutic alliance, the ability of the therapist to repair ruptures in the alliance, and the cohesion in group therapy seemed to be important factors of therapy relationship and therefore important factors in the effectiveness of treatment (Castonguay & Beutler, 2006; Norcross, 2002; Tufekcioglu, Muran, Safran, & Winston, 2013). As STIP-TA is known for its low drop-out rates and ruptures in the alliance were a main issue of the therapy, these may be factors which could explain part of its high effectiveness.

Recent studies point to the need of a refinement of existing treatments in PD. Until now, treatments studied have mostly focused on BPD, on distinct areas of pathology and on comparisons of different psychotherapies. There is a paucity of research to guide treatment and enhance outcome in PD patients (Critchfield & Benjamin, 2006). As most research showed that differences in effectiveness of active treatments in PD are negligible ("equivalence effect") (Budge, Moore, Del Re, Wampold, Baardseth, & Nienhuis, 2013; Dimaggio, 2013), a recent issue of the *Journal of Personality Disorders* stressed the need for the development of comprehensive and integrated treatments in PD patients (Dimaggio & Livesley, 2012). As personality disorders are "multidimensional phenomena" with a large degree of axes I and II comorbidity, treatments that address several problem areas are required (Nelson, Beutler, & Castonguay, 2012). Treatments ought to be based on generic change methods with the addition of specific psychotherapeutic interventions of different psychotherapeutic schools to target specific problems of the individual patient. Attention should be given to common factors such as the therapeutic alliance, the use of a coherent approach, and a careful case formulation (Livesly, 2012; Critchfield, 2012; Clarkin, 2012). Dimaggio (2012) further stresses that treatments should be

adapted to the different phases of the treatment process. Instead of further comparing different treatments, Clarkin (2012) pleads to concentrate research on the active ingredients of treatments as it is unknown which therapeutic ingredients are effective and which are not. This highlights the need of dismantling studies on the principles of change, and research into the efficacious ingredients of effective treatments in PD patients, including STIP-TA.

STIP-TA is known as a high-pressure and potentially destabilizing treatment and as such will not be the first choice of treatment in severely disturbed patients as not all patients seem to be able to withstand the “pressure cooker” of this kind of treatment (Chiesa, Fonagy, & Gordon, 2009) which could result in more dropouts, suicidality, or psychotic decompensation. For example, Gullestad, Johansen, Hoglend, Karterud, and Wilberg, (2013) have found that day hospital treatments and its group therapy format were too demanding in patients with low reflective functioning (i.e. a low level of mentalizing) which resulted in worse treatment effects compared to individual therapy. Factors such as the multiple interactions in group treatments and the use of different therapists were assumed to play a role in these results. From an attachment point of view, it can be hypothesized that the amount and intensity of contact in group treatments can lead to a (hyper)activation of the attachment system, which can result in a high level of stress in patients and therefore a less effective treatment (Hutsebaut, Bales, Busschbach, & Verheul, 2012). On the other hand, cluster C PD patients might be able to handle the pressure and intensity of these treatments when provided in a safe and holding environment. Patients with a cluster C PD often show rigid patterns of behavior, motivated by anxiety, which makes them afraid for change. An intensive, inpatient treatment is pre-eminently suited for these patients. The holding environment in the hospital and the supportive elements in the living together in a group with the treatment contract as an anchor, gives safety. This makes it possible for patients to explore painful experiences in their current life and in their past, and to experiment with new behaviors. A short hospitalization further has less impact on daily life and reduces the risks of iatrogenic effects which can be associated with long-term inpatient treatments. However, it is of practical relevance to study whether more seriously disordered patients could also profit from this intensive kind of treatment and whether it is possible to expand the target population to (relatively mild) cluster A and B PD patients. In a previous study on the effectiveness of aftercare following STIP-TA, almost 30% of the studied patient population was diagnosed with either a cluster A or B PD

(Thunnissen et al., 2008). A secondary analysis on these patient groups showed that patients with a cluster A or B PD did show a different pattern of improvement over time (Thunnissen, 2007). At 24 months follow-up, however, patients showed similar symptom levels and large effect sizes. This might indicate that STIP-TA can also be effective for these patient groups. It is therefore clinically relevant to further investigate the safety and applicability of STIP-TA in cluster A and B PD.

This study showed that STIP-TA is a very promising and effective treatment option in mainly cluster C PD and PDNOS patients. To make this treatment available to more patients, additional research on the effectiveness and cost-effectiveness of this treatment is recommended.

References

- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. Washington, DC: American Psychiatric Association.
- Arnevik, E., Wilberg, T., Urnes, O., Johansen, M., Monsen, J. T., & Karterud, S. (2010). Psychotherapy for personality disorders: 18 months' follow-up of the ullevål personality project. *Journal of Personality Disorders, 24*(2), 188-203.
- Arrindell, W. A., & Ettema, J. H. M. (2003). *SCL-90. Manual for a Multidimensional Indicator of Psychopathology*. Lisse, The Netherlands: Swets & Zeitlinger.
- Austin, P. C. (2008). A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Statistics in Medicine, 27*(12), 2037-2049.
- Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics, 10*(2), 150-161.
- Barnes, G. (2007). Not Without the Couch: Eric Berne on Basic Differences between Transactional Analysis and Psychoanalysis. *Transactional Analysis Journal, 37*(1), 41-50.
- Bartak, A., Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2007). Strengthening the status of psychotherapy for personality disorders: an integrated perspective on effects and costs. *Canadian Journal of Psychiatry, 52*(12), 803-810.
- Bartak, A., Spreeuwenberg, M. D., Andrea, H., Busschbach, J. J., Croon, M. A., Verheul, R., Emmelkamp, P. M., & Stijnen, T. (2009). The use of propensity score methods in psychotherapy research. A practical application. *Psychotherapy and Psychosomatics, 78*(1), 26-34.
- Bartak, A., Spreeuwenberg, M. D., Andrea, H., Holleman, L., Rijnierse, P., Rossum, B. V., Hamers, E. F., Meerman, A. M., Aerts, J., Busschbach, J. J., Verheul, R., Stijnen, T., & Emmelkamp, P. M. (2010). Effectiveness of Different Modalities of Psychotherapeutic Treatment for Patients with Cluster C Personality Disorders: Results of a Large Prospective Multicentre Study. *Psychotherapy and Psychosomatics, 79*(1), 20-30.
- Bateman, A. W., & Fonagy, P. (2000). Effectiveness of psychotherapeutic treatment of personality disorder. *British Journal of Psychiatry, 177*, 138-143.
- Berne, E. (1996). Principles of transactional analysis. *Indian Journal of Psychiatry, 38*(3), 154-159.

- Binks, C. A., Fenton, M., McCarthy, L., Lee, T., Adams, C. E., & Duggan, C. (2006). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews* (1), CD005652.
- Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Sturmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, *163*(12), 1149-1156.
- Budge, S. L., Moore, J. T., Del Re, A. C., Wampold, B. E., Baardseth, T. P., & Nienhuis, J. B. (2013). The effectiveness of evidence-based treatments for personality disorders when comparing treatment-as-usual and bona fide treatments. *Clinical Psychology Review*, *33*(8), 1057-1066.
- Busschbach, J. J., Wolffenbuttel, B. H., Annemans, L., Meering, W. J., & Koltowska-Haggstrom, M. (2011). Deriving reference values and utilities for the QoL-AGHDA in adult GHD. *The European Journal of Health Economics*, *12*(3), 243-252.
- Castonguay, L. G., & Beutler, L. E. (2006). Principles of therapeutic change: A task force on participants, relationships, and techniques factors. *Journal of Clinical Psychology*, *62*(6), 631-638.
- Chiesa, M., & Fonagy, P. (2003). Psychosocial treatment for severe personality disorder. 36-month follow-up. *British Journal of Psychiatry*, *183*, 356-362.
- Chiesa, M., Fonagy, P., & Gordon, J. (2009). Community-Based Psychodynamic Treatment Program for Severe Personality Disorders: Clinical Description and Naturalistic Evaluation. *Journal of Psychiatric Practice*, *15*(1), 12-24.
- Clarkin, J. F. (2012). An Integrated Approach to Psychotherapy Techniques for Patients with Personality Disorder. *Journal of Personality Disorders*, *26*(1), 43-62.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (rev. ed.). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1).
- Critchfield, K. L. (2012). Tailoring Common Treatment Principles to Fit Individual Personalities. *Journal of Personality Disorders*, *26*(1), 108-125.
- Critchfield, K. L., & Benjamin, L. S. (2006). Principles for psychosocial treatment of personality disorder: summary of the APA Division 12 Task Force/NASPR review. *Journal of Clinical Psychology*, *62*(6), 661-674.

- De Beurs, E., & Zitman, F. G. (2006). The Brief Symptom Inventory (BSI): the reliability and validity of a brief alternative of the SCL-90. *Maandblad Geestelijke Volksgezondheid*, 61(2), 120-141.
- De Jong, C. A. J., Derks, F. C. A., van Oel, C. J., & Rinne, T. (1996). *Gestructureerd Interview voor de DSM-IV persoonlijkheidsstoornissen (SIDP-IV)* [Structured Interview for DSM-IV personality (SIDP-IV)]. Sint Oedenrode, The Netherlands: Stichting Verslavingszorg Oost Brabant.
- De Jong, K., Nugter, M. A., Polak, M. G., Wagenborg, J. E. A., Spinhoven, P., & Heiser, W. J. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clinical Psychology & Psychotherapy*, 14(4), 288-301.
- Delimon, J. (1999). Kortdurende klinische psychotherapie [Short inpatient psychotherapy]. In T. J. C. Berk, M. P. Bolten, E. Gans & H. G. Y. Koksma (Eds.), *Handboek Groepspsychotherapie* (pp. F5.1-f5.31). Houten, The Netherlands: Bohn Stafleu van Loghum.
- Derogatis, L. (1983). *SCL-90-R: Administration, Scoring and procedures manual-II*. Towson, MD: Clinical Psychometric Research.
- Derogatis, L. (1992). *BSI: Administration, scoring and procedures manual-II*. Baltimore, MD: Clinical Psychometric Research.
- Derogatis, L. (2011). *BSI en BSI 18. Herziene Nederlandse Editie [BSI and BSI 18. Revised Dutch edition]*. Amsterdam, The Netherlands: Pearson.
- Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory report. *Psychological Medicine*, 13(3), 595-605.
- Dimaggio, G. (2013). Hitting the Bull's Eye in Personality Disorders Psychotherapy. *Journal of Contemporary Psychotherapy*, 1-6.
- Dimaggio, G., & Livesley, J. (2012). Introduction to the Special Feature on the Integrated Treatment of Personality Disorder. *Journal of Personality Disorders*, 26(1), 1-6.
- Dimaggio, G., Nicolò, G., Semerari, A., & Carcione, A. (2013). Investigating the personality disorder psychotherapy process: The roles of symptoms, quality of affects, emotional dysregulation, interpersonal processes, and mentalizing. *Psychotherapy Research*, 23(6), 624-632.
- Dixon-Gordon, K. L., Turner, B. J., & Chapman, A. L. (2011). Psychotherapy for personality disorders. *International Review of Psychiatry*, 23(3), 282-302.
- EuroQolGroup (1995). *EQ-5D user guide*. Rotterdam, The Netherlands: Sanders Institute EUR.

- Gabbard, G. O. (2000). Psychotherapy of personality disorders. *The Journal of Psychotherapy Practice and Research*, 9(1), 1-6.
- Gullestad, F. S., Johansen, M. S., Hoglend, P., Karterud, S., & Wilberg, T. (2013). Mentalization as a moderator of treatment effects: Findings from a randomized clinical trial for personality disorders. *Psychotherapy Research*, 23(6), 674-689.
- Hawkes, L. (2011). With You and Me in Mind: Mentalization and Transactional Analysis. *Transactional Analysis Journal*, 41(3), 230-240.
- Hodgson, R., Bushe, C., & Hunter, R. (2007). Measurement of long-term outcomes in observational and randomised controlled trials. *British Journal of Psychiatry Supplement*, 50, s78-84.
- Hutsebaut, J., Bales, D. L., Busschbach, J. J., & Verheul, R. (2012). The implementation of mentalization-based treatment for adolescents: a case study from an organizational, team and therapist perspective. *International Journal of Mental Health Systems*, 6(1), 10.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12-19.
- Kröger, C., Harbeck, S., Armbrust, M., & Kliem, S. (2013). Effectiveness, response, and dropout of dialectical behavior therapy for borderline personality disorder in an inpatient setting. *Behaviour Research and Therapy*, 51(8), 411-416.
- Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. C., Shimokawa, K., Christopherson, C., & Burlingame, G. B. (2004). *Administration and Scoring Manual for the OQ-45*. Orem, UT: American Professional Credentialing Services.
- Lamers, L. M., Stalmeier, P. F., McDonnell, J., Krabbe, P. F., & van Busschbach, J. J. (2005). Kwaliteit van leven meten in economische evaluaties: het Nederlands EQ-5D-tarief. *Nederlands Tijdschrift voor Geneeskunde*, 149(28), 1574-1578.
- Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ (2008). *Multidisciplinaire Richtlijn Persoonlijkheidsstoornissen* [Multidisciplinary guideline for personality disorders]. Utrecht, The Netherlands: Trimbos-instituut.
- Leichsenring, F., & Leibling, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: a meta-analysis. *American Journal of Psychiatry*, 160(7), 1223-1232.

- Livesly, W. J. (2002). *Manual for the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ)*. Port Huron, MI: Sigma Press.
- Livesley, W. J. (2012). Integrated Treatment: A Conceptual Framework for an Evidence-Based Approach to the Treatment of Personality Disorder. *Journal of Personality Disorders*, 26(1), 17-42.
- Matusiewicz, A. K., Hopwood, C. J., Banducci, A. N., & Lejuez, C. W. (2010). The Effectiveness of Cognitive Behavioral Therapy for Personality Disorders. *Psychiatric Clinics of North America*, 33(3), 657-685.
- McMain, S. F., Guimond, T., Streiner, D. L., Cardish, R. J., & Links, P. S. (2012). Dialectical Behavior Therapy Compared With General Psychiatric Management for Borderline Personality Disorder: Clinical Outcomes and Functioning Over a 2-Year Follow-Up. *American Journal of Psychiatry*, 169(6), 650-661.
- Muran, J. C., Safran, J. D., Samstag, L. W., & Winston, A. (2005). Evaluating an alliance-focused treatment for personality disorders. *Psychotherapy: Theory, Research, Practice, Training*, 42(4), 532-545.
- National Institute for Health and Clinical Excellence (2009a). *Antisocial Personality Disorder: Treatment, Management and Prevention*. Retrieved April 11, 2013, from <http://www.nice.org.uk/CG77>.
- National Institute for Health and Clinical Excellence (2009b). *Borderline Personality Disorder: Treatment and Management*. Retrieved April 11, 2013, from <http://www.nice.org.uk/CG78>.
- Nelson, D. L., Beutler, L. E., & Castonguay, L. G. (2012). Psychotherapy Integration in the Treatment of Personality Disorders: A Commentary. *Journal of Personality Disorders*, 26(1), 7-16.
- Norcross, J. C. (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. New York, NY, US: Oxford University Press.
- Ohlsson, T. (2010). Scientific evidence base for transactional analysis in the year 2010. *International Journal of Transactional Analysis Research*, 1(1), 4-29.
- Perry, J. C., Banon, E., & Ianni, F. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry*, 156(9), 1312-1321.
- Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured Interview for DSM-IV Personality (SIDP-IV)*. Washington, DC: American Psychiatric Press.

- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control-group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33-38.
- Skodol, A. E., Pagano, M. E., Bender, D. S., Shea, M. T., Gunderson, J. G., Yen, S., Stout, R. L., Morey, L. C., Sanislow, C. A., Grilo, C. M., Zanarini, M. C., & McGlashan, T. H. (2005). Stability of functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder over two years. *Psychological Medicine*, *35*(3), 443-451.
- Smit, Y., Huibers, M. J. H., Ioannidis, J. P. A., van Dyck, R., van Tilburg, W., & Arntz, A. (2012). The effectiveness of long-term psychoanalytic psychotherapy-A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *32*(2), 81-92.
- Soeteman, D. I., Verheul, R., Meerman, A. M., Ziegler, U., Rossum, B. V., Delimon, J., Rijniere, P., Thunnissen, M., Busschbach, J. J., & Kim, J. J. (2011). Cost-effectiveness of psychotherapy for cluster C personality disorders: a decision-analytic model in the Netherlands. *Journal of Clinical Psychiatry*, *72*(1), 51-59.
- Stoffers, J. M., Vollm, B. A., Rucker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*, *8*, CD005652.
- Svartberg, M., Stiles, T. C., & Seltzer, M. H. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *American Journal of Psychiatry*, *161*(5), 810-817.
- Thunnissen, M. (2007). *Long-term prognosis and aftercare in short-term inpatient psychotherapy of personality disorders: A randomised clinical trial of two methods of aftercare* (Doctoral dissertation). Erasmus MedicalCenter, Rotterdam, The Netherlands.
- Thunnissen, M., Duivenvoorden, H., Busschbach, J., Hakkaart-van Roijen, L., van Tilburg, W., Verheul, R., & Trijsburg, W. (2008). A randomized clinical trial on the effectiveness of a reintegration training program versus booster sessions after short-term inpatient psychotherapy. *Journal of Personality Disorders*, *22*(5), 483-495.

- Town, J. M., Abbass, A., & Hardy, G. (2011). Short-Term Psychodynamic Psychotherapy For Personality Disorders: A Critical Review Of Randomized Controlled Trials. *Journal of Personality Disorders, 25*(6), 723-740.
- Tufekcioglu, S., Muran, J. C., Safran, J. D., & Winston, A. (2013). Personality disorder and early therapeutic alliance in two time-limited therapies. *Psychotherapy Research, 23*(6), 646-657.
- Van Beek, N., & Verheul, R. (2008). Motivation for treatment in patients with personality disorders. *Journal of Personality Disorders, 22*(1), 89-100.
- Van Kampen, D. (2002). The DAPP-BQ in the Netherlands: Factor structure and relationship with basic personality dimensions. *Journal of Personality Disorders, 16*(3), 235-254.
- Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J., Bateman, A. W., & Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability, and validity. *Psychological Assessment, 20*(1), 23-34.
- Verheul, R., & Herbrink, M. (2007). The efficacy of various modalities of psychotherapy for personality disorders: a systematic review of the evidence and clinical recommendations. *International Review of Psychiatry, 19*(1), 25-38.

Tables and figures.

Table 1. Sociodemographic and clinical characteristics of STIP-TA and OP patients.

Table 2. Mean outcome and effect size in STIP-TA and OP patients.

Figure 1. Patient flowchart.

Figure 2. Distribution of estimated probabilities of assignment to STIP-TA treatment (propensity scores) before and after matching.

Figure 3. Symptom severity over the course of three years.

Table 1. Sociodemographic and clinical characteristics of STIP-TA and OP patients.

Sociodemographic and clinical characteristics	STIP-TA n=67	OP n=67
	<i>mean (s.d.)</i>	
Age, years	39.4 (9.8)	39.3 (10.2)
	<i>n (%)</i>	
Male gender	22 (33)	18 (27)
Education		
High (EQF ¹ ≥6)	26 (39)	24 (36)
Medium (EQF ¹ 3-5)	27 (40)	31 (46)
Low (EQF ¹ ≤2)	14 (21)	12 (18)
General way of living		
Alone (with or without child)	31 (46)	25 (37)
With partner (with or without child)	33 (49)	33 (49)
With parent(s)	2 (3)	6 (9)
With other people	1 (2)	3 (5)
Care for child(ren)	19 (28)	27 (40)
Civil status		
Married/steady relationship	23 (34)	26 (39)
Divorced/widowed	8 (12)	7 (10)
Never married	36 (54)	34 (51)
Mode of employment		
Paid work/study	39 (58)	38 (57)
Unemployed/other	28 (42)	29 (43)
Presence PD*		
Cluster A	1 (2)	2 (3)
Cluster B	8 (12)	7 (11)
Cluster C	33 (49)	26 (39)
PDNOS	28 (42)	33 (49)
	<i>mean (s.d.)</i>	
Number of PDs	1.8 (0.9)	1.6 (0.9)
Number of PD criteria	17.1 (5.6)	15.9 (6.2)
Previous treatment		
Outpatient treatment	62 (93)	55 (82)
Inpatient treatment	11 (16)	16 (24)
Medication treatment	41 (61)	34 (51)

¹European Qualifications Framework

²Since it is possible to have more than one diagnosis, the sum of the prevalence is higher than 100%.

Table 2. Mean outcome and effect size in STIP-TA and OP patients.

mean (s.d.)	Baseline	12 months	24 months	36 months	Cohen's d (within-group)		
					12 months	24 months	36 months
GSI							
STIP-TA (n=67)	1.59 (0.58)	0.55 (0.45)	0.62 (0.46)	0.57 (0.47)	2.02	1.87	1.93
OP (n=67)	1.59 (0.58)	0.90 (0.60)	0.80 (0.59)	0.78 (0.59)	1.18	1.36	1.39
b		0.35*	0.18*	0.21*			
OQ-45							
Social Role							
STIP-TA (n=67)	16.53 (4.31)	11.35 (6.63)	11.02 (6.62)	10.83 (6.67)	0.93	0.99	1.02
OP (n=67)	15.63 (4.50)	11.74 (6.90)	10.86 (6.52)	10.57 (6.39)	0.67	0.86	0.92
b		1.29	0.75	0.64			
Interpersonal Relations							
STIP-TA (n=67)	21.22 (6.17)	15.49 (4.26)	15.67 (4.37)	14.87 (4.53)	1.09	1.05	1.18
OP (n=67)	20.39 (5.10)	16.37 (4.89)	15.10 (4.68)	14.77 (4.52)	0.81	1.09	1.05
b		1.71	0.26	0.73			
EQ-5D							
STIP-TA (n=67)	0.54 (0.27)	0.82 (0.20)	0.81 (0.21)	0.80 (0.22)	1.19	1.1	1.04
OP (n=67)	0.53 (0.26)	0.69 (0.27)	0.73 (0.26)	0.73 (0.25)	0.61	0.76	0.79
b		0.12*	0.07*	0.05			

GSI=Global Severity Index. OQ-45=Outcome Questionnaire-45. EQ-5D=EuroQol-5D.

b= regression-coefficient of the difference of the outcome measures, * regression-coefficient b significant at p<0.05.

Figure 1. Patient flowchart.

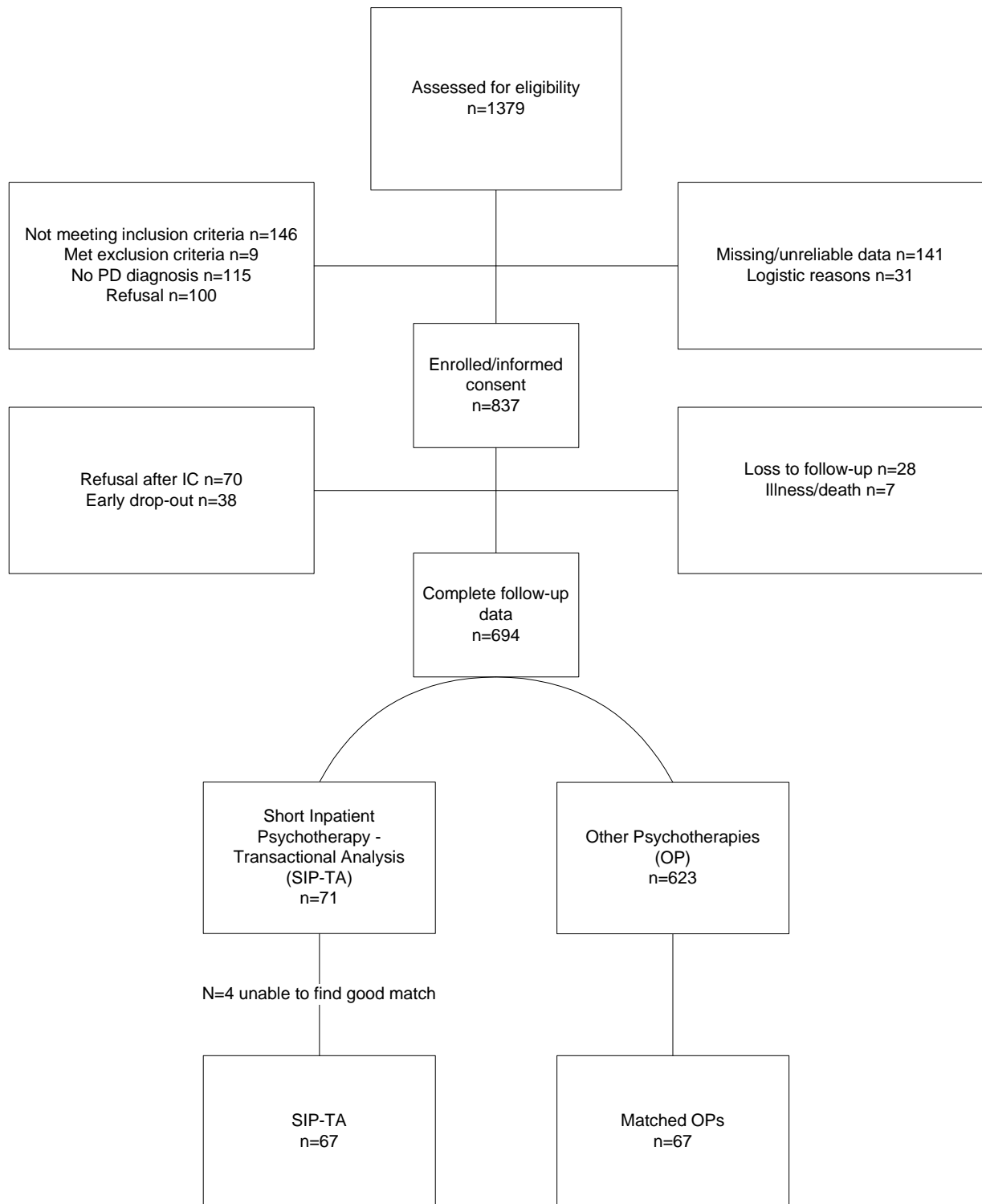


Figure 2. Distribution of estimated probabilities of assignment to STIP-TA treatment (propensity scores) before and

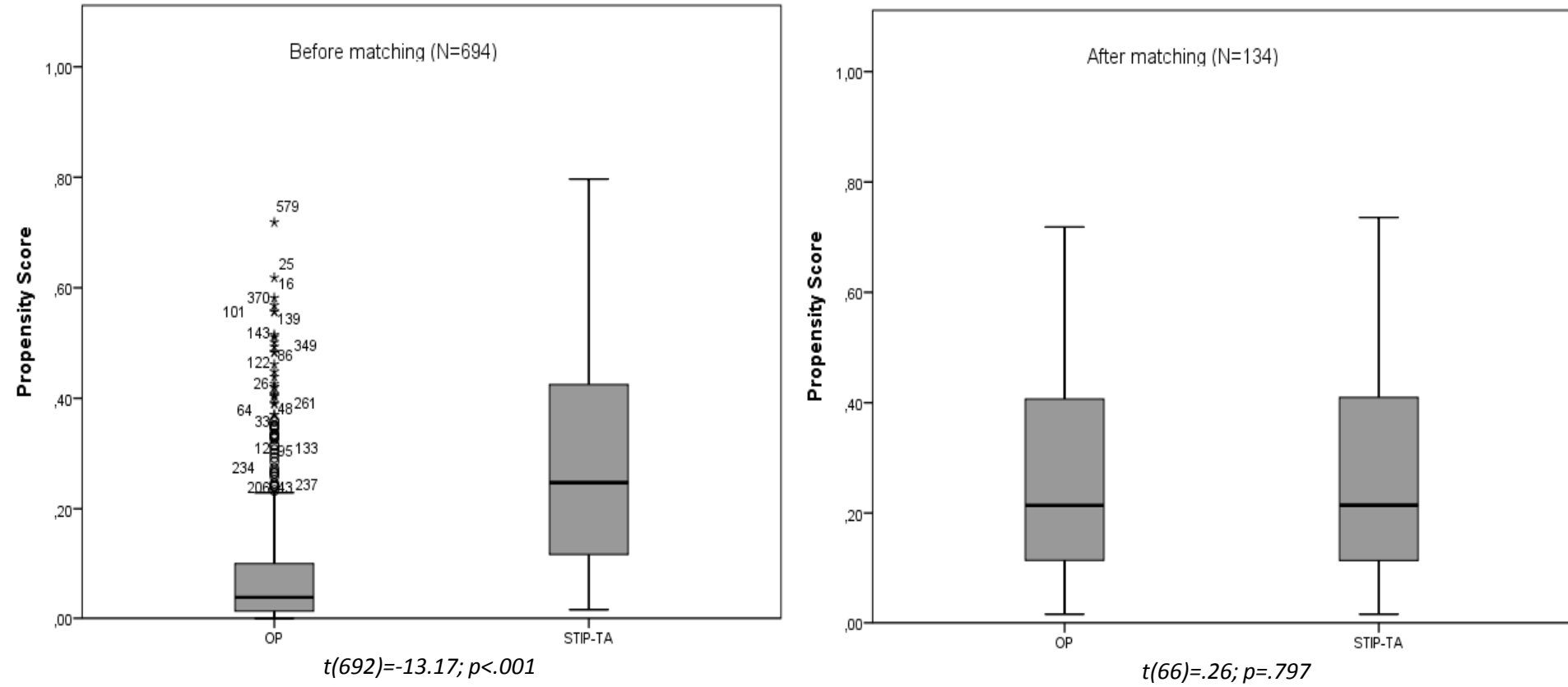


Figure 3. Symptom severity over the course of three years.

