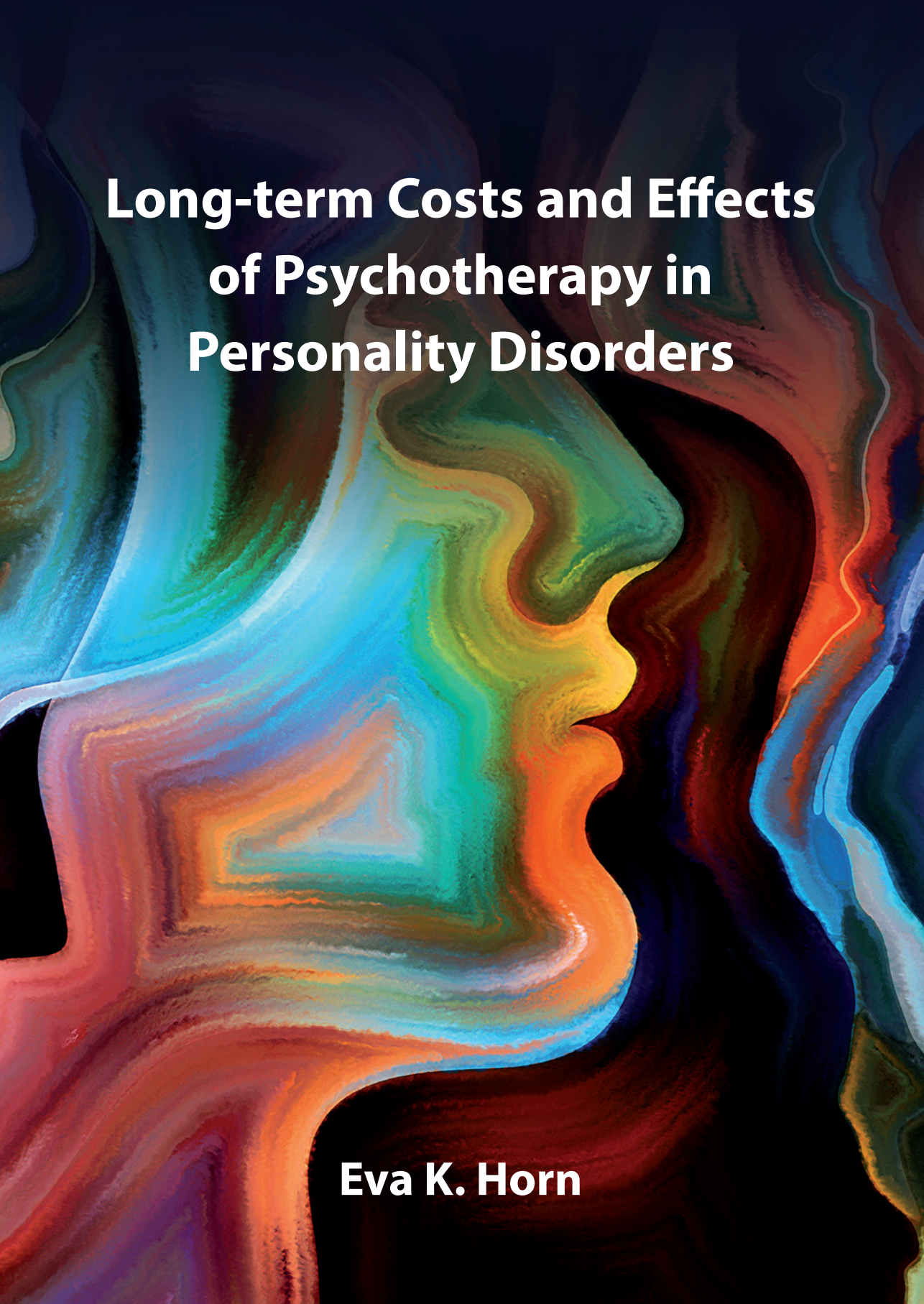
Long-term Costs and Effects of Psychotherapy in Personality Disorders



Eva K. Horn

Long-term Costs and Effects of Psychotherapy in Personality Disorders

Eva K.Horn

Cover: Ridderprint BV, Ridderkerk, the Netherlands
Layout: Ridderprint BV - www.ridderprint.nl
Printed by: Ridderprint BV - www.ridderprint.nl
ISBN: 978-94-6299-324-2

Long-Term Costs and Effects of Psychotherapy in Personality Disorders

Kosten en Effecten van Psychotherapie bij Persoonlijkheidsstoornissen

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
dinsdag 19 april 2016 om 13:30 uur

door

Eva Karin Horn
geboren te Saarbrücken, Duitsland

Erasmus University Rotterdam

Promotiecommissie

Promotor:	Prof.dr. J.J. van Busschbach
Overige leden:	Prof.dr. C.L. Mulder
	Prof.dr. W.B.F. Brouwer
	Prof.dr. P. Luyten
Copromotor:	Dr. R. Verheul

CONTENTS

CHAPTER 1

Introduction

The Study of Cost Effectiveness of Personality Disorder Treatment (SCEPTRE) showed that patients with personality disorder (PD) benefit from psychotherapy (Bartak, 2011; Soeteman, 2010). It should be noted, however, that especially in patients with a Cluster C PD, treatment effects depend on treatment modality (outpatient, day hospital or inpatient treatment) and length of treatment. Furthermore, some particular combinations of treatment modality and length of treatment were found to have a favorable cost-effectiveness. So far, analyses in the SCEPTRE study are limited to the use of three year follow-up data and patients with PD Not Otherwise Specified (PDNOS) were excluded. The (cost-)effectiveness studies in this thesis use the SCEPTRE data up to five years follow-up and do include patients with PDNOS. What is more, the (cost-)effectiveness of one specific promising treatment was investigated in-depth, namely Short-Term Inpatient Treatment based on Transactional Analysis (STIP-TA).

This introduction starts with an outline of the context of and the rationale for the SCEPTRE investigation. What follows next is a description of the concept of evidence based medicine, i.e. the scientific paradigm on which this research effort is based, and an overview of the scientific evidence of the effectiveness of psychotherapy in PD. Next, the design of SCEPTRE and the observed outcomes are reported. This introductory chapter concludes with the objectives and the content of the current thesis.

NEED FOR (COST-)EFFECTIVENESS RESEARCH

Over the years, mental health care has become more accessible and, at the same time, seeking help for mental health problems is becoming more accepted (Schomerus et al., 2012). The liberalization of the mental health care market in the Netherland has led to the availability of more health care providers and more and different treatments (Bijenhof, Folkertsma, Kommer, Slobbe, & Polder, 2012). This trend is inevitably accompanied with an increase in costs. In the Netherlands, from 2005 to 2010, mental health care costs for patients with PD went up by approximately 70% for men and by even 187% for women (Bijenhof et al., 2012). The recent budget constrains made it even more necessary to justify interventions in terms of their effectiveness and cost-effectiveness and in addition shifted attention to evidence based treatments. This is especially true for expensive interventions with limited published evidence for its effectiveness, such as inpatient treatment of PD.

The publication of the DSM-III in 1980 (American Psychiatric Association, 1987), in which the criteria of personality disorders were explicitly established for the first time and which was followed by the DSM-IV in 1994 (American Psychiatric Association, 1994), led

to more attention for personality disorders and their treatment. The general diagnostic criteria of the DSM-IV-TR (American Psychiatric Association, 2000) and which are incorporated in the DSM-5 (American Psychiatric Association, 2013) are listed below.

According to DSM-IV-TR, the diagnosis of a personality disorder must satisfy the following general criteria, in addition to the specific criteria listed under the specific personality disorder under consideration.

A. An enduring pattern of inner experience and behavior deviating markedly from the expectations of the individual's culture. This pattern is manifested in two (or more) of the following areas:
   (1) cognition (perception and interpretation of self, others and events)
   (2) affect (the range, intensity, lability and appropriateness of emotional response)
   (3) interpersonal functioning
   (4) impulse control
B. The enduring pattern is inflexible and pervasive across a broad range of personal and social situations.
C. The enduring pattern leads to clinically significant distress or impairment in social, occupational or other important areas of functioning.
D. The pattern is stable and of long duration and its onset can be traced back at least to adolescence or early adulthood.
E. The enduring pattern is not better accounted for as a manifestation or consequence of another mental disorder.
F. The enduring pattern is not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition (e.g., head trauma).

Textbox. General diagnostic criteria for a DSM-IV-TR Axis II personality disorder (APA, 2000)

DE VIERSPRONG; INITIATOR OF THE *SCEPTRE* STUDY

The recent history of *De Viersprong* nicely illustrates these developments. A specialized psychotherapeutic center for the treatment of PD patients in the Netherlands, *De Viersprong* started as a "*therapeutic community*" in the 1950s with long-term inpatient treatments, which could last up to two years. The past decade has seen a reduction in its inpatient treatments, both in the number of beds and in the length of treatments. At the same time the institute has been investing in research on the effectiveness of treatments and the relationship between costs and effects, with the aim to offer evidence based treatments to patients with PD. Such investigations should ultimately provide evidence whether the high costs of inpatient treatments can be justified with possible savings

elsewhere, and/or better health outcomes. This train of thought led to the initiation of the SCEPTRE study in 2003, i.e. a large quasi-experimental investigation of the effectiveness and cost-effectiveness of different dosages of psychotherapies for PD (Bartak, 2011; Soeteman, 2010).

EVIDENCE BASED TREATMENTS: THE BASIS FOR SCEPTRE

The SCEPTRE investigation was based on the concept of evidence based medicine, defined in 1996 by Sackett, Rosenberg, Gray, Haynes, and Richardson as *"the conscientious, explicit and judicious use of current best evidence in making decisions about the care of the individual patient. It means integrating individual clinical expertise with the best available external clinical evidence from systematic research*." An increasing interest in evidence based medicine stimulated studies on psychotherapeutic treatments in PD patients; such studies had so far been relatively rare, with most research concentrating on various manualized psychotherapies for patients with borderline PD. The SCEPTRE initiative tried to fill that gap with a large controlled study in all clusters of PD, including the PDNOS.

SCIENTIFIC EVIDENCE ON EFFECTIVENESS OF PSYCHOTHERAPY FOR PD PATIENTS

The evidence of psychotherapeutic treatments in PD patients is mixed and largely depends on the PD diagnosis. In a systematic overview of studies involving cluster A, B and C PDs, Perry, Banon and Ianni (1999) estimated effect sizes of 1.1 to 1.3 pre-treatment to post-treatment. More specifically, for cognitive behavior therapies a mean effect size of 0.87 was found and for psychodynamic therapies a mean effect size of 1.79 (Leichsenring & Leibing, 2003). The very few studies on cluster A patients consist of case reports or have otherwise included small study populations. Some studies show that cluster A patients hardly respond to psychotherapy whereas other studies show that they do (Dixon-Gordon, Turner, & Chapman, 2011; Gude & Vaglum, 2001; Thunnissen, 2007). A possible explanation for this discrepancy is that studies often employ only a limited follow-up period and that these patients seem to improve slowly (Thunnissen, 2007). Alternatively, selection bias might have played a role. Cluster B, and more specifically BPD, has been studied more extensively. Several systematic reviews addressed the psychological and pharmacological treatment of borderline (Lieb, Vollm, Rucker, Timmer, & Stoffers, 2010; Stoffers et al., 2012) and antisocial PD (Gibbon et al., 2010; Glenn, Johnson, & Raine, 2013; Khalifa et al., 2010). The authors of a recent Cochrane review on psychological treatments in BPD concluded that these are effective and deserve an

important place in the management of patients with BPD (Stoffers et al., 2012). A recent review on antisocial PD showed some effectiveness of CBT and MBT (Glenn et al., 2013). Other cluster B PDs, i.e. narcissistic and histrionic PD, have been less researched. People with a narcissistic PD seem to profit from psychotherapeutic treatments but drop-out is high (Levy, Reynoso, Wasserman, & Clarkin, 2007). Studies on cluster C PD patients showed that psychodynamic, cognitive (behavioural) therapy and brief relational therapy all seemed to be effective. Medium to large effects were found up to three years after baseline (Dixon-Gordon et al., 2011). At present no explicit treatment studies on patients with PDNOS have been published. In the day hospital treatment study of Karterud et al. (2003) in a mixed population, the best treatment results were found in PDNOS patients (Karterud et al., 2003).

STUDY OF COST-EFFECTIVENESS OF PERSONALITY DISORDER TREATMENT (SCEPTRE)

The studies referred to above provide evidence that psychotherapeutic treatments in PD patients are effective. Remarkably, typically there are hardly any or no differences in effectiveness between specific treatments and theoretical orientations (Bartak, Soeteman, Verheul, & Busschbach, 2007; Budge et al., 2013). This is a common finding in studies of psychological treatments and is often referred to as the "*Dodo Bird verdict*". The Dodo Bird verdict implies that all psychotherapies lead to approximately equivalent effects seeing that common factors such as 'the belief in the treatment' and 'the therapeutic alliance' underlie the effectiveness of psychotherapy (Wampold et al., 1997).

Importantly, many if not all comparative studies have focussed on equal dosages of psychotherapy. As various studies have observed dose-effect associations in PD, it is plausible to hypothesize that treatment effectiveness is driven in whole or in part by the setting (outpatient, day hospital or inpatient treatment) and the length of treatment. Furthermore, as setting and length of treatment largely determine the costs of treatment, the SCEPTRE study could link variance in costs to variance in effectiveness.

In the SCEPTRE trial, 1,379 patients completed the intake procedure (2003-2006) and were selected for various treatment dosages in six mental health care centres in the Netherlands: *De Viersprong*, Netherlands Institute for Personality Disorders, Halsteren; *GGZ WNB*, Bergen op Zoom and Roosendaal; *Centre of Psychotherapy Pro Persona*, Lunteren; *Altrecht*, Utrecht; *Zaans Medical Centre*, Zaandam; and the *Centre of Psychotherapy Arkin*, Amsterdam (van Manen et al., 2011). The treatments under study varied greatly in setting, length of treatment, and theoretical orientation. The dosages, which were

the focus of the study, were defined as followed: outpatient, day hospital and inpatient treatments. More specific, in the cluster C study, the length of treatment (short- or long-term) could be added to the setting to make a further distinction. That was not possible in clusters A and B, because of the small numbers of patients in the resulting 6 subgroups. Assignment to treatment was based on the results of standardized assessments and the expert opinions of clinicians from the participating health care centres, as it had appeared to be impossible to randomize patients to treatments with this great variation. The SCEPTRE study thus follows a quasi-experimental study design. To correct for the initial differences between patients in different treatments, we made use of the propensity score and the multiple propensity score (Spreeuwenberg et al., 2010). The studies presented in this thesis included 921 of those 1,379 patients, in different compositions. Previous studies so far had used patient data up to three year follow-up in the analyses. Anna Bartak presented short- (12 months follow-up) and mid-term (18 months follow-up) effectiveness results in her thesis (2010). She showed that dosage had some influence on effectiveness (Bartak et al., 2010, 2011a, 2011b). In cluster A, day hospital and inpatient treatments were found associated with the best results. However, as the treatment groups were not readily comparable, the general conclusion as that cluster A patient can profit from psychotherapy (Bartak et al., 2011a). In cluster B, inpatient treatments were found most effective (Bartak et al., 2011b), and in cluster C large effects in favor of short-term inpatient treatments were found (Bartak et al., 2010). The thesis of Djøra Soeteman (2010) showed that dosage had some influence on the cost-effectiveness of psychotherapy. More precisely, when the costs were added to the analyses and the costs and effects were modeled over five years, outpatient and day hospital treatments were the most cost-effective treatments for cluster B (Soeteman et al., 2010), while short-term day hospital and short-term inpatient treatments were the most cost-effective treatments in cluster C (Soeteman et al., 2011). Cost-effectiveness was not modeled for cluster A patients due to the low number of patients.

The studies in the present thesis draw on the studies presented by Anna Bartak and Djøra Soeteman. We used the now available data up to five years, presented the three and five years' long-term effects for clusters A, B, C PD, and PDNOS, and combined effectiveness with cost-effectiveness in a short-term inpatient treatment study. Furthermore, one could argue that patients low on psychological strength cannot benefit from treatment with the high level of destabilization that is typical for short-term inpatient psychotherapy. We therefore took into account the interaction between the patient's psychological strength and the treatment's level of destabilization with respect to outcome in all patients and over all treatments.

OBJECTIVE AND RESEARCH QUESTIONS OF THIS THESIS

The objective was to address aspects of the SCEPTRE study that had been left unexplored so far and which have not yet been addressed in earlier publications. We formulated the following three research questions:

1. Are the improvements previously observed in patients with a cluster A, B, or C PD or PDNOS stable over five years of follow-up? Are there differences in the effectiveness of outpatient, day hospital, and inpatient treatments on the long-term outcome in patients with a cluster A, B, or C PD or PDNOS?
2. What are the effectiveness and cost-effectiveness of Short-Term Inpatient Psychotherapy based on Transactional Analysis (STIP-TA) in PD patients?
3. Is there an interaction between the patient's psychological strength and the treatment's level of destabilization with respect to outcome?

CONTENT OF THIS THESIS

This thesis presents data collected from a quasi-experimental study. **Chapter 2** reports on the long-term effectiveness in patients with a cluster A, B and/or C PD. **Chapter 3** explores the long-term effectiveness in patients with a PDNOS. **Chapter 4** compares the effectiveness of a specific treatment in PD, namely "*STIP-TA*" to that of other psychotherapies in patients with mainly a cluster C PD or PDNOS. **Chapter 5** compares the cost-effectiveness of STIP-TA to that of other psychotherapies over three years' follow-up. **Chapter 6** reports on the matching hypothesis that patients high on psychological strengths profit more from predominantly destabilizing treatments while patients low on psychological strengths profit more from predominantly stabilizing treatments. **Chapter 7** is the general discussion where the findings of this thesis are summarized and the implications for clinical practice and future research are discussed.

REFERENCES

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Washington, DC: American Psychiatric Association.

Bartak, A. (2011). *On the Effectiveness of Psychotherapy in Personality Disorders* (Doctoral dissertation). Retrieved from dare.uva.nl/document/184990.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Thunnissen, M., Ziegler, U. M., Dekker, J., … Emmelkamp, P. M. G. (2011a). Patients with cluster A Personality Disorders in Psychotherapy: An Effectiveness Study. *Psychotherapy and Psychosomatics, 80*, 28-38.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Ziegler, U. M., Dekker, J., Rossum, B. V., … Emmelkamp, P. M. G. (2011b). Effectiveness of Outpatient, Day Hospital, and Inpatient Psychotherapeutic Treatment for Patients with Cluster B Personality Disorder. *Psychotherapy and Psychosomatics, 80*, 88-99.

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., Holleman, L., Rijnierse, P., Rossum, B. V., Emmelkamp, P. M. G. (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.

Bijenhof, A., Folkertsma, M., Kommer, G., Slobbe, L., & Polder, J. (2012). *Notitie Kostenontwikkeling GGZ* [Note Development of costs in Mental Health]. Bilthoven, NL: RIVM.

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.

Dixon-Gordon, K. L., Turner, B. J., & Chapman, A. L. (2011). Psychotherapy for personality disorders. *International Review of Psychiatry, 23*(3), 282-302.

Gibbon, S., Duggan, C., Stoffers, J., Huband, N., Vollm, B. A., Ferriter, M., … Lieb, K. (2010). Psychological interventions for antisocial personality disorder. *Cochrane Database of Systematic Reviews*(6), CD007668.

Glenn, A. L., Johnson, A. K., & Raine, A. (2013). Antisocial Personality Disorder: A Current Review. *Current Psychiatry Reports, 15*(12).

Gude, T., & Vaglum, P. (2001). One-year follow-up of patients with cluster C personality disorders: a prospective study comparing patients with "pure" and comorbid conditions within cluster C, and "pure" C with "pure" cluster A or B conditions. *Journal of Personality Disorders, 15*(3), 216-228.

Karterud, S., Pedersen, G., Bjordal, E., Brabrand, J., Friis, S., Haaseth, O. … Urnes, O. (2003). Day treatment of patients with personality disorders: experiences from a Norwegian treatment research network. *Journal of Personality Disorders, 17*(3), 243-262.

Khalifa, N., Duggan, C., Stoffers, J., Huband, N., Vollm, B. A., Ferriter, M., … Lieb, K. (2010). Pharmacological interventions for antisocial personality disorder. *Cochrane Database of Systematic Reviews*(8), CD007667.

Leichsenring, F., & Leibing, 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.

Levy, K. N., Reynoso, J., Wasserman, R. H., & Clarkin, J. F. (2007). Narcissistic personality disorder. In W. O'Donohue, K. A. Fowler & S. O. Lilienfeld (Eds.), *Personality Disorders: Toward the DSM-V* (pp. 233-277). Thousand Oaks, CA, US: Sage Publications, Inc.

Lieb, K., Vollm, B., Rucker, G., Timmer, A., & Stoffers, J. M. (2010). Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. *British Journal of Psychiatry, 196*(1), 4-12.

Perry, J. C., Banon, E., & Ianni, F. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry, 156*(9), 1312-1321.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal, 312*(7023), 71-72.

Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., Angermeyer, M.C. (2012). Evolution of public attitudes about mental illness: a systematic review and meta-analysis. *Acta Psychiatrica Scandinavica, 125*(6), 440-452.

Soeteman, D. I. (2010). *Economic Evaluation of Psychotherapy for Personality Disorders: burden of disease, cost-effectiveness, and the value of further research and active implementation* (Doctoral dissertation). Retrieved from http://repub.eur.nl/pub/19865.

Soeteman, D. I., Verheul, R., Delimon, J., Meerman, A. M. M. A., van den Eijnden, E., Rossum, B. V., … Kim, J. J. (2010). Cost-effectiveness of psychotherapy for cluster B personality disorders. *British Journal of Psychiatry, 196*(5), 396-403.

Soeteman, D. I., Verheul, R., Meerman, A. M. M. A., Ziegler, U. M., Rossum, B. V., Delimon, 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.

Spreeuwenberg, M. D., Bartak, A., Croon, M. A., Hagenaars, J. A., Busschbach, J. J. V., Andrea, H., … Stijnen, T. (2010). The Multiple Propensity Score as Control for Bias in the Comparison of More Than Two Treatment Arms An Introduction From a Case Study in Mental Health. *Medical Care, 48*(2), 166-174.

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.

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). Retrieved from http://repub.eur.nl/pub/8322.

van Manen, J. G., Andrea, H., van den Eijnden, E., Meerman, A. M., Thunnissen, M. M., Hamers, E. F., … Verheul, R. (2011). Relationship between patient characteristics and treatment allocation for patients with personality disorders. *Journal of Personality Disorders, 25*(5), 656-667.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially, "all must have prizes." *Psychological Bulletin, 122*(3), 203-215.

CHAPTER 2

Long-Term Effectiveness of Psychotherapy in Personality Disorders

Eva K. Horn[a, b], Anna Bartak[c, d], Anke M. M. A. Meerman[e], Bert V. Rossum[f], Uli M. Ziegler[g], Mirjam Soons[h], Moniek Thunnissen[i, j], Paul M. G. Emmelkamp[d, k], Theo Stijnen[l], Roel Verheul[d, m], 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] Bos en Lommer Private Practice, Amsterdam; [d] Department of Clinical Psychology, University of Amsterdam, Amsterdam; [e] Centre of Psychotherapy Pro Persona, Lunteren; [f] Altrecht, Zeist; [g] Zaans Medical Centre, Zaandam; [h] Arkin, Amsterdam; [i] GGZ WNB, Bergen op Zoom; [j] Private Practice, Bergen op Zoom; [k] The Center for Social and Humanities Research, King Abdulaziz University, Jeddah, Saudi Arabia; [l] Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden; [m] De Viersprong, Netherlands Institute for Personality Disorders, Halsteren, the Netherlands

SUMMARY

Background

The effectiveness of specialized psychotherapy in personality disorders (PDs) is well documented. Previous research on short- and mid-term follow-up showed some dosage-effect relationships, with superiority of short-term inpatient psychotherapy in patients with cluster C PD being the most significant finding. This manuscript reports about the 5-year follow-up of patients with PD in various treatment modalities.

Methods

Five hundred nineteen patients with a DSM-IV-TR cluster A, B, and/or C PD, assigned to outpatient, day hospital or inpatient treatments were followed up to five years. In patients with cluster C PD, short-term treatments (< 6 months) were further distinguished from long-term treatments. Primary outcome was symptom severity (GSI). Secondary outcomes were psychosocial functioning (OQ-45) and health related quality of life (EQ-5D). Multiple propensity scores were used to correct for initial baseline differences.

Results

Uncorrected results showed that all patient groups except cluster A outpatients reported significantly less symptom severity (effect sizes .65 to 1.82), and that initial positive outcomes were maintained over the 5-year follow-up. Corrected differences between the modalities were small and mostly non-significant. In cluster C patients, short-term inpatient treatments were superior over most of the other modalities.

Conclusions

After five year follow-up, patients still maintained a better level of symptoms and functioning than before treatment. Differences between treatment modalities were small. The previously observed superiority of short-term inpatient psychotherapy in patients with cluster C PDs at 12 month follow-up was still present at 5- year follow-up.

INTRODUCTION

The effectiveness and cost-effectiveness of specialized psychotherapy in personality disorders (PDs) is well documented both on the short- and on the long-term (Binks et al., 2006; Dixon-Gordon, Turner, & Chapman, 2011; Hadjipavlou & Ogrodniczuk, 2010; Leichsenring & Leibing, 2003; Perry, Banon, & Ianni, 1999). This evidence is also reflected in current guidelines of e.g. the UK (National Institute for Health and Clinical Excellence, 2009a, 2009b), Australia (National Health and Medical Research Council, 2012), and the Netherlands (Trimbos-institute, 2008). Differences in effects among psychotherapeutic orientations turned out to be small (Bartak, Soeteman, Verheul, & Busschbach, 2007; Budge et al., 2013). Earlier research on the short- and mid-term effectiveness showed small differences between treatment modalities in terms of length and the setting of the treatment: outpatient, day hospital an inpatient treatment (Bartak et al, 2010, 2011a, 2011b). This article reports about the long-term outcome of these studies at 5 years after baseline.

Almost half of all patients (46%) in mental health care have a DSM-IV personality disorder (PD) (Zimmerman, Rothschild, & Chelminski, 2005). This makes PDs one of the most frequent disorders treated by psychiatrists, psychologists and other mental health care workers in outpatient care with diagnoses of cluster C PD and PDNOS being most prevalent (Zimmerman et al., 2005). The burden of disease is high: PDs are characterized by enduring maladaptive patterns of behaviour, which often lead to impairments in different areas, such as occupational or social functioning, additional psychopathology, a diminished global functioning, and a low quality of life (Kvarstein & Karterud, 2012; Samuels, 2011; Soeteman, Verheul, & Busschbach, 2008). Since one aspect of PDs is their long-lasting, persisting pathology, long-term follow-up of patients is significant to determine whether treatment is capable to produce long-lasting change.

Until now, treatments studied have mostly focused on patients with a specific personality pathology (e.g. borderline personality disorders [BPD]), or on comparisons of different psychotherapies (Bateman & Fonagy, 2009; Giesen-Bloo et al., 2006). A recent review showed that so-called evidence-based treatments were more effective than treatment as usual (TAU) but comparative trials on active treatments have not yet provided conclusive evidence for the superiority of one theoretical orientation over another ("*equivalence effect*" or "*dodo bird effect")* (Budge et al., 2013). Studies further suggest that treatment characteristics other than theoretical orientation might be more important determinants of (cost-) effectiveness, such as a well-structured and coherent theoretic framework (Bateman & Fonagy, 2000), the strength of the therapeutic alliance (Falkenstrom, Granstrom, & Holmqvist, 2013), or modality (setting and/or duration) of treatment (Bartak et al., 2007). The long-term effectiveness of different modalities (setting and/or duration) of psychotherapy is the focus of the present study.

In cluster A, studies on psychotherapeutic treatments are rare and existing studies have rather small study populations and often focus on schizotypal PD. Results of these studies are mixed with a few studies showing that cluster A patients have poor results but other studies showing that patients with a cluster A PD can profit from treatment as much as patients with a cluster B and/or C PD (Dixon-Gordon et al., 2011; Gude & Vaglum, 2001; Thunnissen, 2007). A possible explanation for these mixed results is the finding that cluster A patients seem to improve more slowly and thus need more time to show positive results (Thunnissen, 2007).

Cluster B, and more specifically BPD, have been studied more extensively. Systematic reviews have been conducted concerning the psychological and pharmacological treatment of borderline (Lieb, Vollm, Rucker, Timmer, & Stoffers, 2010; Stoffers et al., 2012) and antisocial PD (Gibbon et al., 2010; Khalifa et al., 2010). The authors of a recent Cochrane review on psychological treatments in BPD concluded that psychotherapeutic treatments are effective, although more studies are necessary (Stoffers et al., 2012). Less research is done on the narcissistic, antisocial, and histrionic PD. A review on antisocial PD showed some effectiveness of cognitive behavioural treatment (CBT) and mentalization based treatment (MBT) (Glenn, Johnson, & Raine, 2013).

Studies on cluster C patients showed that psychodynamic, cognitive (behavioural) therapy and brief relational therapy all seemed to be effective. Medium to large effects were evident up to three years after baseline (Dixon-Gordon et al., 2011).

Between 2010 and 2011 several articles were published on the effectiveness of psychotherapy in personality disorders (PDs) on the basis of the SCEPTRE study (trial register ISRCTN: 73817429), a longitudinal multicentre study, which includes more than 900 patient treated for different personality disorders in The Netherlands. As earlier studies showed that the choice of theoretical framework had little effect on treatment effectiveness, SCEPTRE focused on the dosage or modality of treatment as the moderator of effect. The published articles of Bartak et al. presented the results up to 18 month of follow-up (Bartak et al., 2010, 2011a, 2011b). Significant improvements of symptoms, interpersonal functioning, and quality of life, and some significant differences between treatment modalities were found. In cluster A, day hospital and inpatient treatments, in cluster B inpatient treatments, and in cluster C short-term inpatient treatments seemed to produce best results. Five year follow-up using data of patients with PD not otherwise specified (PDNOS) was published recently and showed mostly comparable effectiveness of these treatment dosages at long-term follow-up (Horn et al., 2014).

In the present study we present the long-term results up to five years after baseline using data of patients with cluster A, B, and C PD obtained from SCEPTRE. The study protocol was approved by the Medical Ethics Committee of the Erasmus University Medical Centre in Rotterdam, the Netherlands. This study aims to extend the evidence on the effectiveness of psychotherapy for patients with cluster A, B, and C PDs. It is the first study to investigate the long-term effectiveness of various modalities of psychotherapy in these patient groups. The effectiveness of outpatient, day hospital, and inpatient psychotherapy is investigated over five years after baseline.

The following research questions will be addressed:

1) Is the improvement in symptom severity, interpersonal functioning, and quality of life found in earlier short-term studies stable over time?
2) Are there differences in the effectiveness of outpatient, day hospital and/or inpatient treatments on the long-term?

METHODS

Study population and design

During a 3-year period, 1,379 patients completed the intake procedure in six mental health care centres in the Netherlands and were selected for treatment (De Viersprong, Netherlands Institute for Personality Disorders, Halsteren; GGZ WNB, Bergen op Zoom and Roosendaal; Centre of Psychotherapy Pro Persona, Lunteren; Altrecht, Utrecht; Zaans Medical Centre, Zaandam; Centre of Psychotherapy Arkin, Amsterdam). Assignment to treatment was based on the results of standardized assessments and the expert opinion of clinicians from the participating health care centres. Of the 1,379 patients who were selected for treatment, 959 were enrolled in the SCEPTRE study. One hundred and forty six patients did not meet the inclusion criteria: age between 18 and 70 years (n = 13), significant personality pathology (n = 34), referred for treatment aimed at personality pathology (n = 99). Nine patients met the exclusion criteria: insufficient command of the Dutch language (n = 6), organic cerebral impairment (n = 1), mental retardation (n = 1), and schizophrenia (n = 1). One hundred and thirty-three patients were excluded due to unreliable or missing baseline data, 100 patients refused to participate and 32 patients were excluded due to logistic reasons (e.g. it was not possible to make an appointment to inform patients about the study). Five hundred and eighty eight patients had a Cluster A, B, and/or C PD and had received at least two sessions of outpatient psychotherapy or at least two days of day hospital or inpatient psychotherapy (see figure 2.1). Of these, 519 (88%) had completed at least one follow-up assessment and were included in the present study.

Figure 2.1. Patient Flow.

Treatments

The six mental health care centres offered a variety of psychotherapeutic treatments tailored to a PD patient population. Their treatments differed according to several features. As this study focused on different treatment modalities, three modalities were compared. Outpatient (i.e., individual or group psychotherapy sessions, up to two sessions per week), day hospital (i.e., at least one morning/afternoon per week, various forms of psychotherapeutic and psychosocial treatment, but patients sleep at home), and inpatient (i.e., patients stay at the institutions up to five days a week, various forms of psychotherapeutic and psychosocial treatments) psychotherapy with the following characteristics (mean ±SD):

1. outpatient psychotherapy, .85 ±. 44 sessions per week, 13.58 ± 7.26 months, treatments with integrative (31%), cognitive-behavioural (28%), psychodynamic (25%) or an other (16%) orientation;
2. day hospital psychotherapy, 3.21 ± 1.44 days per week, 9.38 ± 4.04 months, treatments with integrative (82%), psychodynamic (13%) or cognitive-behavioural (6%) orientation;
3. inpatient psychotherapy 4.97 ±. 16 days per week, 7.75 ± 3.68 months, treatments with integrative (40%), psychodynamic (31%) or cognitive-behavioural (29%) orientation.

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, individual psychotherapy, milieu therapy, coaching for social problems, community meetings, and/or pharmacological treatment (Bartak et al., 2010). One hundred-and-one psychotherapists who were all licensed psychiatrists or psychologists participated in this study. On average, they had 14.9 ± 10.1 years 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. Analyses were carried out on the basis of intended treatments.

Assessments

An extensive standard assessment battery of instruments was administered to the patients before treatment assignment. PDs were measured 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). The primary outcome measure was symptom severity. Symptom severity was measured using the Dutch version of the Brief Symptom Inventory (De Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983), a validated self-report scale derived from the revised Symptom Checklist 90 (Arrindell &

Ettema, 2003; Derogatis, 1983). 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 zero to four. Secondary outcome measures were psychosocial functioning and health-related quality of life. Psychosocial functioning was measured using two subscales of the Outcome Questionnaire-45 (OQ-45), i.e. Interpersonal Relations and Social Role (Lambert et al., 2004). Health-related quality of life was measured using the EQ-5D (EuroQolGroup, 1995). A recent study in the Netherlands measured and valuated the EQ-5D, resulting in the Dutch EQ-5D value set, which was used to calculate utilities for EQ-5D health states (Lamers, Stalmeier, McDonnell, Krabbe, & van Busschbach, 2005).

The assessment battery included three additional instruments, which were used as potential confounders for the estimation of the propensity score (see below). First, the Dutch version of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ) was used to measure type and degree of personality pathology (Livesley, 2002; van Kampen, 2002). 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 core components of personality pathology were measured using the Severity Indices of Personality Problems (SIPP-118), a 118-item questionnaire aimed to measure five core domains of personality pathology, i.e. Self-Control, Identity Integration, Responsibility, Relational Functioning, and Social Concordance (Verheul et al., 2008).

All outcome measures were assessed at baseline and several follow-up points. Three treatment centres had assessments at baseline, end of treatment, at six and 12 months after the end of treatment, and at 36 and 60 months after baseline. Three other centres had assessments at baseline and at 12, 24, 36, and 60 months after baseline. Different time points were used due to logistic reasons (Bartak et al., 2010). Follow-up response was high, thereby enhancing the robustness of the multi-level analyses. We included all patients with at least one follow-up measure, as multi-level models make optimal use of incomplete repeated measures data and are robust for drop-out when the drop-out is missing at random (Little & Rubin, 1987). Twenty-eight percent had one to three follow-ups, 37% had four follow-ups, 29% had five follow-ups, and 7% had six follow-ups. No significant differences ($p<.05$) were found between patients with and without follow-up concerning the outcome measures at baseline.

Statistical analyses

We first examined the uncorrected results on all outcome measures over five years after baseline. We used multilevel modelling to deal with: (1) the dependency of the repeated measures within the same subject in time, and (2) the longitudinal data with observations

unequally spaced in time (see 'Assessments'). To estimate the uncorrected treatment effects over five years after baseline, we used a random intercept and random slope model with time as level I and patient number as level II. In addition to a linear time effect, we modelled the time effect through linear splines with knots every six months, which allowed the estimated course of the dependant variable to bend at these time points. Non-significant knots ($p \geq 0.05$) 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 following independent variables: dummy variables indicating group membership, time, a knot-point at three years allowing a change in the slope of the time effect at 3 years, interaction between group membership and time, and interaction between group membership and the change in slope at three years. Subsequently, we calculated within-group effect sizes (ES, Cohen's d) using the estimated pooled standard deviations from the models to describe change from baseline up to five years in each group (Cohen, 1988).

To evaluate clinically significant change at five years after baseline in terms of the GSI, we followed the criteria by Jacobson and Truax (Jacobson & Truax, 1991). We only used patients with a five year follow-up measure and we did not correct the results for initial baseline differences. The cut-off and reliable change index of the manual of the BSI were used (Derogatis, 2011). Fisher's exact test was used to test for differences between the treatment groups.

Since this is a non-randomised study, the comparison of the groups had to be corrected for the influence of potential confounders, i.e. initial patient differences between modalities and related to outcome. To adjust for these differences and avoid bias in effect estimation, we included a multiple propensity score in our analyses. The classic propensity score is defined as the conditional probability of assignment to one of two modalities given a set of observed pre-treatment variables (Bartak et al., 2009; Rosenbaum & Rubin, 1983). The multiple propensity score is an extension of the classic propensity score to more than two modalities and has proven to be feasible in mental health research (Spreeuwenberg et al., 2010). To identify relevant confounders, a comprehensive list of social, economic, and diagnostic variables based on the literature and clinical knowledge was carefully considered by both clinicians and researchers (Bartak et al., 2009). All variables significantly related to a specific outcome were used to estimate the multiple propensity scores in a multinomial logistic regression analysis, with group membership as the dependent variable (Brookhart et al., 2006). For illustration, the propensity score for cluster C for the GSI was estimated using the following variables: age, level of education, baseline scores of the GSI, EQ-5D, DAPP-BQ (Emotional Dysregulation, Dissocial Behaviour, Inhibition), OQ-45 (Symptom Distress, Interpersonal Relations, Social Role), MTQ-8 (Need for Help, Readiness to change), SIPP-118 (Self-Control, Identity Integration, Relational Functioning, Social Concordance, Responsibility), SIDP-IV Dimensional score (cluster A PD, cluster

C PD and total score), amount of cluster C traits, and diagnoses of avoidant and dependant PD (see appendix table for a complete list of potential and included confounders for all patient groups and all outcome measures). By using the propensity score method, the overlap in propensity score distributions (and thus the overlap in relevant variables) between modalities can be visualised and judged. In earlier investigations about the effectiveness of psychotherapy on short- and mid-term follow-up with patients with cluster A, B and C PDs, this visual judgement led to the exclusion of a group (in cluster C), a less refined differentiation between modalities (in cluster A), and to a less firm interpretation of results (cluster A), as the propensity score distributions were too far apart (Bartak et al., 2010, 2011a, 2011b). Abovementioned remarks about the cluster A and C studies also applied to the current study, despite the longer follow-up period. The analyses on the total group were carried out on the uncorrected results only, as the groups were not readily comparable, the propensity scores were too far apart. Subsequently a multilevel model, including multiple propensity scores to correct for initial patient differences, was used to compare change in outcome variables across modalities. Dependent variables were the change scores (from baseline) observed during follow-up for each of the outcome measures. Independent variables were dummy variables indicating group membership, time, a knot point at three years, interaction between group membership and time, interaction between group membership and a knot point at three years, and the multiple propensity scores with their mutual interactions. This model was used to estimate differences in change scores at 60 months after baseline in pair wise comparisons of the six modalities.

The analyses were performed using IBM SPSS Statistics 20.0 for data preparation and SAS 9.2 for multi-level modelling.

Additional analyses

The five modalities of cluster C PD.

In order to be able to compare results of this study with the earlier publication on cluster C (Bartak et al., 2010), additional analyses were done to inspect long-term treatment effects in more detail. The three main modalities were split up into five subgroups, i.e. long-term (more than six months) outpatient treatment, short- (up to six months) and long-term day-hospital treatment, and short- and long-term inpatient treatment. The eighteen patients following short-term outpatient treatment were not included in the analyses. See Bartak et al. (Bartak et al., 2010) for a description of the modalities.

Treatment Compliance.

Patients were allocated to their treatment modality in terms of setting and duration preceding their treatment. A secondary analysis was carried out on patients who followed

their initial, intended treatment modality and patients who changed their treatment modality.

Comorbidity.

The patient groups in the current study showed considerable comorbidity on axis 2 of the DSM-IV (see table 2.1). Therefore, additional analyses were carried out to compare

Table 2.1. Patient characteristics, treatment variables and PD diagnoses of cluster A, B, and C PD and the total PD group.[1]

Patient characteristics	Cluster A	Cluster B	Cluster C	Total
	N=59	N=210	N=416	N=519
Mean ±SD				
Age	29.3 ± 8.1	31.5 ± 8.5	33.8 ± 9.6	33.2 ± 9.4
N (%)				
Gender				
Male	18 (31)	61 (29)	126 (30)	160 (31)
Education*				
High (EQF≥6)	14 (24)	65 (31)	147 (35)	177 (34)
Medium (EQF 3 to 5)	18 (31)	73 (35)	167 (40)	202 (39)
Low (EQF≤2)	27 (46)	72 (34)	102 (25)	140 (27)
General way of living (with or without children)				
Alone	19 (32)	92 (44)	169 (41)	222 (43)
With partner	22 (37)	75 (36)	162 (39)	190 (37)
With parent(s)	12 (20)	23 (11)	52 (13)	67 (13)
With other people	6 (10)	20 (10)	33 (8)	40 (8)
Children				
Care for child(ren)	10 (17)	35 (17)	94 (23)	108 (21)
Civil status				
Married/steady relationship	8 (14)	26 (12)	93 (22)	102 (20)
Divorced/widowed	4 (7)	18 (9)	39 (9)	47 (9)
Never married	47 (80)	166 (79)	284 (68)	370 (71)
Mode of employment				
Paid work/study	40 (68)	130 (62)	270 (65)	333 (64)
Previous treatment				
Outpatient	44 (75)	173 (82)	344 (83)	428 (83)
Inpatient	13 (22)	44 (21)	73 (18)	98 (19)

Table 2.1. Patient characteristics, treatment variables and PD diagnoses of cluster A, B, and C PD and the total PD group.[1] (continued)

Treatment variables	Cluster A	Cluster B	Cluster C	Total
	N=59	N=210	N=416	N=519
Mean ±SD				
Frequency				
Outpatient (sessions/week)	.96 ±.64	.72 ±.41	.89 ±.47	.84 ±.46
Day Hospital (days/week)	3.32 ±1.60	3.47 ±1.41	3.14 ±1.48	3.30 ±1.50
Inpatient (days/week)	5.00 ±.00	4.97 ±.14	4.97 ±.14	5.00 ±.20
Duration (months)				
Outpatient	13.21 ±6.03	14.54 ±6.56	13.36 ±6.71	13.74±6.84
Day Hospital	10.32 ±4.49	10.30 ±4.76	8.89 ±3.99	9.34±4.12
Inpatient	8.53 ±2.34	9.07 ±2.97	7.83 ±3.42	8.15±3.40
N (%)				
Theoretical orientation				
Outpatient				
Cognitive-behavioural	5 (24)	16 (34)	27 (27)	34 (28)
Psychodynamic	4 (19)	5 (11)	27 (27)	31 (25)
Integrative	10 (48)	14 (30)	29 (29)	36 (30)
Other	2 (10)	12 (26)	16 (16)	21 (17)
Day Hospital				
Cognitive-behavioural	3 (16)	6 (7)	10 (6)	12 (6)
Psychodynamic	3 (16)	16 (20)	18 (11)	25 (13)
Integrative	13 (36)	60 (73)	137 (83)	158 (81)
Inpatient				
Cognitive-behavioural	4 (21)	30 (37)	38 (25)	58 (29)
Psychodynamic	1 (5)	12 (15)	54 (35)	63 (31)
Integrative	14 (74)	39 (48)	63 (41)	81 (40)
PD diagnoses	**Outpatient** N=122	**Day hospital** N=195	**Inpatient** N=202	**Overall N=519**
N (%)				
Cluster A only	4 (3)	2 (1)	5 (3)	11 (2)
Cluster B only	15 (12)	31 (16)	39 (19)	85 (16)
Cluster C only	63 (52)	105 (54)	112 (55)	280 (54)
Cluster A and B	4 (3)	0 (0)	3 (2)	7 (1)
Cluster A and C	8 (7)	6 (3)	4 (2)	18 (4)
Cluster B and C	23 (19)	40 (21)	32 (16)	95 (18)
Cluster A, B, and C	5 (4)	11 (6)	7 (4)	23 (4)

*EQF= European Qualifications Framework.

[1] As the patient groups were overlapping, the total N is lower than the sum of cluster A, B, and C.

patients with one or more PDs of one cluster (pure) with patients with PDs from more than one cluster (mixed) on the primary outcome.

RESULTS

Total sample

Description total group

Mean age of the sample was 33.2 years and the majority (69%) of all patients was female (see table 2.1). More than 80% of patients had a history of outpatient treatment; almost 20% had a history of inpatient treatment. The largest part of patients (80%) had a cluster C PD, 41% had a cluster B PD and 11% had a diagnosis of cluster A (patients could have more than one PD diagnosis). Fifty-three percent of all treatments were described as integrative, 23% of treatments were psychodynamic, 20% cognitive behavioural, and 4% were described as an other theoretical orientation.

Outcomes total group

Sixty months after baseline, uncorrected results showed within-group effect sizes from .89 (large effect, outpatient treatment) to 1.30 and 1.38 (very large effect, day hospital and inpatient treatment, see table 2.2). The improvement was significant for all dosages

Table 2.2. Effect sizes at 60 months for three treatment modalities split up in Cluster A, B, and C (uncorrected).

Cluster A (n=59)

Variable	Treatment group	N	Mean ±SD		Within-group effect size (Cohen's d)
			Baseline	60 months	
GSI[1]	Outpatient	21	1.39 ±.68	1.09 ±.83	.41
	Day hospital	19	1.99 ±.62	.90 ±.62	1.81*
	Inpatient	19	1.92 ±.83	.72 ±.48	1.82*
OQ-45 Social Role	Outpatient	21	15.91 ±4.71	10.58 ±4.81	1.15
	Day hospital	19	18.84 ±5.36	10.70 ±5.72	1.51*
	Inpatient	19	17.33 ±5.67	10.83 ±4.26	1.33*
OQ-45 Interpersonal Relations	Outpatient	21	21.22 ±5.41	17.66 ±7.97	.54
	Day hospital	19	23.67 ±6.50	15.65 ±6.81	1.24*
	Inpatient	19	25.49 ±4.61	14.90 ±6.39	1.95*
EQ-5D[1]	Outpatient	21	.67 ±.18	.71 ±.36	.14
	Day hospital	19	.45 ±.28	.66 ±.37	.66*
	Inpatient	19	.54 ±.30	.81 ±.25	1.00*

Table 2.2. Effect sizes at 60 months for three treatment modalities split up in Cluster A, B, and C (uncorrected). (continued)

Cluster B (n=210)

Variable	Treatment group	N	Mean ±SD		Within-group effect size (Cohen's d)
			Baseline	60 months	
GSI	Outpatient	47	1.53 ±.77	1.05 ±.73	.65*
	Day hospital	82	1.75 ±.61	.87 ±.68	1.37*
	Inpatient	81	1.93 ±.66	.88 ±.77	1.47*
OQ-45 Social Role	Outpatient	47	15.51 ±4.51	11.86 ±4.95	.78*
	Day hospital	82	16.42 ±4.48	10.46 ±5.72	1.17*
	Inpatient	81	17.32 ±5.42	11.53 ±5.34	1.08*
OQ-45 Interpersonal Relations	Outpatient	47	21.36 ±7.14	16.65 ±8.69	.60*
	Day hospital	82	21.26 ±6.00	14.02 ±8.27	1.01*
	Inpatient	81	23.49 ±5.97	16.35 ±7.79	1.04*
EQ-5D	Outpatient	47	.58 ±.27	.67 ±.32	.31
	Day hospital	82	.48 ±.27	.71 ±.29	.83*
	Inpatient	81	.50 ±.27	.77 ±.20	1.14*

Cluster C (n=416)

Variable	Treatment group	N	Mean ±SD		Within-group effect size (Cohen's d)
			Baseline	60 months	
GSI	Outpatient	99	1.39 ±.69	.85 ±.64	.82*
	Day hospital	162	1.57 ±.64	.80 ±.68	1.17*
	Inpatient	155	1.77 ±.65	.77 ±.70	1.49*
OQ-45 Social Role	Outpatient	99	15.32 ±4.73	11.79 ±5.27	.71*
	Day hospital	162	15.85 ±4.93	10.74 ±4.90	1.04*
	Inpatient	155	17.28 ±4.39	11.71 ±5.06	1.18*
OQ-45 Interpersonal Relations	Outpatient	99	21.00 ±6.30	15.81 ±7.85	.73*
	Day hospital	162	21.61 ±5.97	15.33 ±7.79	.91*
	Inpatient	155	24.02 ±5.38	15.72 ±7.22	1.31*
EQ-5D	Outpatient	99	.60 ±.24	.74 ±.26	.56*
	Day hospital	162	.55 ±.27	.74 ±.26	.72*
	Inpatient	155	.50 ±.27	.78 ±.24	1.10*

GSI=Global Severity Index. OQ-45=Outcome Questionnaire-45. EQ-5D=EuroQol-5D.
*significant differences ($p<.05$) between baseline and follow-up.
[1]significant differences on baseline ($p<0.05$)

Table 2.3. Clinical significant change rates at 60 months for Cluster A, B, and C split up in the three treatment modalities.

GSI		N	CSC[1]
Cluster A	Outpatient	14	43%
	Day hospital	13	46%
	Inpatient	10	40%
Cluster B*	Outpatient	33	24%
	Day hospital	42	45%
	Inpatient	50	54%
Cluster C*	Outpatient	68	31%
	Day hospital	104	49%
	Inpatient	102	60%
Total group*	Outpatient	86	43%
	Day hospital	115	49%
	Inpatient	129	59%

** significant difference between treatment groups (p<.05)*

[1] CSC=clinical significant change; patients moved from a dysfunctional range to a normative range and also demonstrated reliable change

(p<.0001). Thirty-four percent of patients following outpatient psychotherapy, 49% of patients following day hospital psychotherapy, and 59% of patients following inpatient psychotherapy showed clinical significant change (see table 2.3).

Cluster A

Description cluster A

Of the 59 patients with a cluster A PD, 85% had a diagnosis of paranoid PD, 10% of schizoid PD, and 10% of schizotypal PD. The largest part (98%) had one cluster A PD, 2% had two cluster A PDs. We found significant differences (p<.05) between the three modalities at baseline concerning age, civil status, living situation, education, care for children, and baseline scores of the MTQ-8 (Need for Help), GSI, OQ-45 (Total Score), EQ-5D, and the SIPP-118 (Identity Integration). Overall, the outpatient group deviated from the day hospital and inpatient group in having less severe problems, whilst differences between the latter two groups were negligible.

Outcomes cluster A

Sixty months after baseline, uncorrected results showed within-group effect sizes from .41 (medium effect, outpatient treatment) to 1.81 and 1.82 (very large effect, day hospital and inpatient treatment; see table 2.2). The improvement was significant for day hospital and inpatient treatments, but marginally significant for the outpatient treatments (d=.41, p=.053). See figure 2.2 for a visual display of the modelled GSI scores. Fourty to 46% of

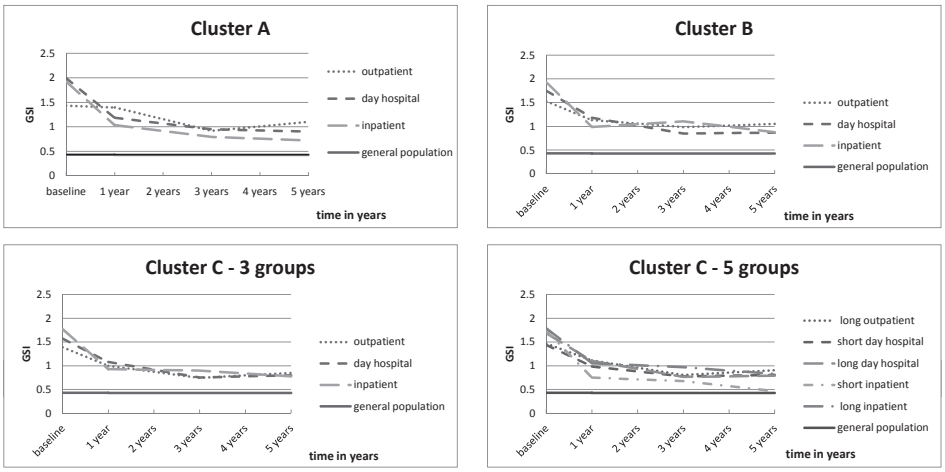


Figure 2.2. Course of the GSI up to five years.

all cluster A patients showed clinical significant change (see table 2.3). In the secondary outcome variables, comparable results were found. Notably, concerning quality of life as outcome variable, patients in the outpatient group already started with a good quality of life and showed almost no improvement after five years (EQ-5D, d=.14, p=.60).

When the analyses were corrected for the influence of baseline differences by the propensity score, day hospital and inpatient treatments were significantly more effective in decreasing symptom severity than outpatient treatments (see table 2.4). On the secondary outcomes, no significant differences were present. However, since the propensity scores were too far apart which means that patients in the three modalities were too different to be readily comparable, these results should be interpreted cautiously.

Cluster B

Description cluster B

In cluster B, 77% had a diagnosis of borderline PD, 23% of narcissistic PD, 12% of histrionic PD, and 9% of antisocial PD. The largest part had one cluster B disorder (81%), 16% had two cluster B PDs, and 2% had three cluster B PDs.

Outcomes cluster B

In cluster B, the improvement after five years was significant for all three groups on all outcome measures, except for outpatients showing only marginal significance on quality of life (EQ-5D, d=.31, p=.07). Effect sizes for the primary outcome (GSI) ranged from .65 (medium effect, outpatient treatment) to 1.47 (large effect, inpatient treatment; see

table 2.2). See figure 2.2 for the modelled GSI scores. Twenty-four percent of outpatients to 54% of inpatients showed clinical significant change (table 2.3). Corrected results showed that patients in inpatient treatments outperformed patients in outpatient treatments concerning quality of life (see table 2.4). No significant differences between modalities were found concerning symptom severity and interpersonal functioning.

Table 2.4. Difference scores at 60 months for three treatment groups split up in Cluster A, B, and C (corrected).

Cluster A (n=59)

Variable	Treatment group	N	Outpatient	Day hospital
GSI	Outpatient	21		
	Day hospital	19	.57*	
	Inpatient	19	.86**	.27
OQ-45 Social Role	Outpatient	21		
	Day hospital	19	1.47	
	Inpatient	19	1.89	.42
OQ-45 Interpersonal Relations	Outpatient	21		
	Day hospital	19	1.56	
	Inpatient	19	3.44	1.88
EQ-5D	Outpatient	21		
	Day hospital	19	.08	
	Inpatient	19	.17	.09

Cluster B (n=210)

Variable	Treatment group	N	Outpatient	Day hospital
GSI	Outpatient	47		
	Day hospital	82	.23	
	Inpatient	81	.29	.06
OQ-45 Social Role	Outpatient	47		
	Day hospital	82	.93	
	Inpatient	81	-.42	-1.36
OQ-45 Interpersonal Relations	Outpatient	47		
	Day hospital	82	2.08	
	Inpatient	81	.94	-1.14
EQ-5D	Outpatient	47		
	Day hospital	82	.06	
	Inpatient	81	.15*	.08

Table 2.4. Difference scores at 60 months for three treatment groups split up in Cluster A, B, and C (corrected). (Continued)

Cluster C (n=416)

Variable	Treatment group	N	Outpatient	Day hospital
GSI	Outpatient	99		
	Day hospital	162	.07	
	Inpatient	155	.16	.08
OQ-45 Social Role	Outpatient	99		
	Day hospital	162	.65	
	Inpatient	155	.13	-.53
OQ-45 Interpersonal Relations	Outpatient	99		
	Day hospital	162	.31	
	Inpatient	155	1.14	.83
EQ-5D	Outpatient	99		
	Day hospital	162	.01	
	Inpatient	155	.07	.06

GSI=Global Severity Index. OQ-45=Outcome Questionnaire-45. EQ-5D=EuroQol-5D.
Positive coefficients indicate that the treatment group shown in the left column is superior, negative coefficients indicate that the treatment group in the above row is superior.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Cluster C

Description cluster C

In cluster C, 63% had a diagnosis of avoidant PD, 49% had a diagnosis of obsessive-compulsive PD, and 22% of dependent PD. More than two third (71%) had one cluster C disorder, 24% had two cluster C PDs, and 5% had three cluster C PDs.

Outcomes cluster C

In the cluster C population, effect sizes on the GSI can be considered large (.82 for outpatient and 1.17 for day hospital treatment) or very large (1.49 for inpatient treatment; see table 2.2). Effect sizes of the secondary outcome measures ranged from .56 (EQ-5D, outpatient treatment) to 1.31 (OQ-45 Interpersonal Relations, inpatient treatment). The improvements found were significant in all three groups for all outcome measures. Thirty-one percent of outpatients to 60% of inpatients showed clinical significant change (table 2.3). No significant differences between the three modalities on the primary outcome measure were visible when corrected for initial differences (see table 2.4).

Additional analyses

The five modalities of cluster C PD

When the three modalities were split up into five modalities by introducing length of treatment as a differentiating factor, uncorrected results showed that most improvement was observed in the short-term inpatient group with an effect size of 2.28 for the GSI. Long-term day hospital and long-term inpatient treatments reached effect sizes of 1.41 (long day hospital) and 1.24 (long inpatient) on the GSI, whereas long-term outpatient and short-term day hospital treatments reached effect sizes of .82 and .95, respectively. The clinical significant change rates ranged from 32% for patients in long-term outpatient psychotherapy to 70% for patients in short-term inpatient therapies. We found strong and significant effects in improvement of symptom severity for patients in short-term inpatient treatments compared to long-term outpatient (b=.39, p=.01), short-term day hospital (b=.33, p=.02), and long-term inpatient treatment (b=-.35, p=.01), and marginally significant differences compared to long-term day hospital treatment (b=.31, p=.07).

Treatment Compliance

Nineteen percent of patients changed their intended treatment modalities during treatment (cluster A: 25%, cluster B: 17%, cluster C: 19%). Analyses on differences between patients who stayed in their planned modality and patients who did not were carried out every twelve months after baseline. No significant differences were found on the GSI up to five year follow-up between patients who stayed in their planned modality and patients who did not (at 60 month follow-up; cluster A: b=.14, p=.85, cluster B: b=-.64, p=.25, cluster C: b=-.27, p=.41).

Comorbidity

Results between patients with DSM-IV axis 2 comorbidity and patients without showed that only within patients with a cluster C PD who followed day hospital treatment, a significant difference emerged on the primary outcome: patients with comorbidity showed more progress than the group without comorbidity (b=-.28, p=.03). In cluster C patients, additional analyses were done on five treatment modalities. These analyses showed a significant difference in the long-term day hospital cluster C group in favour of the group with comorbidity (b=-.47, p=.01).

DISCUSSION

The current study showed that the improvement in symptoms, interpersonal functioning, and quality of life after psychotherapy in PD patients was stable or further increased compared to 12 or 18 months after the start of treatment (Bartak et al., 2010, 2011a, 2011b). Psychotherapy seemed to produce good and long-lasting results up to five years in a severely disordered PD group. Little differences in effectiveness were found between three treatment modalities (outpatient, day hospital, and inpatient) at five years after baseline. Exceptions were found in cluster A, where we found a superiority of day hospital and inpatient treatments compared to outpatient treatments in terms of psychiatric symptomatology. These results, however, have to be interpreted cautiously due to the differences between the patient groups. In cluster B at five year follow-up, inpatients reported a better quality of life compared to outpatients. In cluster C we only found significant differences when treatment modalities were split up in five modalities. This resulted in a superiority of short-term inpatient treatments compared to long-term outpatient, short-term day hospital and long-term inpatient treatments in terms of symptom severity.

In all three clusters, we found most progress during the first year after baseline. During the follow-up, gained progress was maintained or further improved. This is in line which research that showed that most improvement occurs during treatment (Simon, 2009), and that gained progress is maintained or further improved (e.g. Zanarini, Frankenburg, Reich, & Fitzmaurice, 2010; Gunderson et al., 2011).

Strengths & limitations

The most obvious strengths of this study are the high number of patients, the long-term follow-up over five years, a high follow-up response, the wide range of treatments under study, and the naturalistic design with the use of a sophisticated statistical technique to mimic a randomized allocation to treatments (multiple propensity score). Obviously, the need for such a sophisticated statistical technique also marks the most important limitation, i.e. patients were not randomized to treatment. This leaves the possibility that the propensity score has not been able to fully compensate the skewness of the background variables. In cluster A, baseline differences were substantial and the overlap of the propensity score distributions was limited. Consequently, it was not possible to directly compare patients in the three treatment modalities and we do not know whether treatment effects can solely be attributed to the treatment or whether these should be rather attributed to patient characteristics (Bartak et al., 2011a). Another limitation of the propensity score is that it can only correct for background variables included in

the propensity score. Given the high number of variables analyzed and included in the propensity score, we did the best we can to cover a wide range of variables.

A second limitation of this study is that the theoretical orientation of the therapy is possibly associated with setting and duration. However, as plead in the introduction, we think that the theoretical orientation of the therapy had little effect on outcome, and therefore could have only limited effect on the results found. Other aspects of treatment, such as a coherent theoretical structure or the strength of the therapeutic alliance seem more effective ingredients of treatment than the theoretical orientation (Bateman & Fonagy, 2000).

A further limitation of this study is the absence of a control group without treatment. We do not know which part of the improvement is due to therapy and which part is due to e.g. natural recovery. For ethical reasons, it would not be possible to compare an active treatment with a placebo or waiting list. The current study can be understood as an example of "*comparative effectiveness*" (Institute of Medicine, 2009). The changeability of PDs is still an ongoing discussion (Morey & Hopwood, 2013). Research showed that PD patients in treatment have a seven-fold faster recovery than patients without treatment (Perry et al., 1999) and that psychotherapeutic treatment can lead to remission of symptoms (Zanarini, Frankenburg, Reich, & Fitzmaurice, 2012). In a recent meta-analysis it was concluded that psychotherapy should aim on behaviour change, instead of global personality change, as the change of fundamental personality structures may be unrealistic (Ferguson, 2010). We do not know whether the underlying personality pathology in our study changed as our outcome variables relied on certain assessment instruments and on self-report. However, it seems unlikely that the large effect sizes on outcome found in the current study can occur when the underlying pathology still exists in its original form.

The high level of comorbidity in our patient group made it difficult to differentiate between the clusters A, B, and C PDs. However, as this study was not a RCT with stringent exclusion criteria, the high level of comorbidity demonstrates a representative patient group. Indeed the overlap found in clinical practice between clusters of PD has initiated a discussion to the change of the classification model to a more dimensional model, as was suggested during the design of the DSM-5 (Skodol et al., 2011). The results of the current study only showed small differences between patients with and without comorbidity on the DSM-IV axis 2. Only in patients with a cluster C PD in (long-term) day hospital treatments, comorbidity on axis 2 was related to more progress compared to patients without comorbidity. The presence of comorbidity on axis 2 did therefore not seem to have a large influence on treatment effectiveness in the current study.

Research and clinical implications

As we found little differences in effectiveness when corrected for initial differences, we cannot state that one treatment modality is better than the other for cluster A and B PD. Within cluster A and B, the results of this study suggest counterintuitive clinical implications: more intensive, inpatient treatments are not superior to less intensive, outpatient treatments. When choosing a treatment, other characteristics than effectiveness, like costs of the treatment and the burden of treatment on everyday life, might therefore become more important. Policy makers might be inclined to further reduce more expensive day hospital and inpatient treatments and to further invest in cheaper outpatient treatments. Such line of reasoning ignores that other studies show that high costs of intensive treatments in PD can be compensated by savings in other health care areas and work-related costs (e.g. Soeteman et al., 2010; Soeteman et al., 2011; van Asselt et al., 2008). The current study showed superiority of short-term inpatient treatments to most other treatments in cluster C PDs. As it was not possible to split up the treatments of cluster A and B patients in terms of setting and duration, we do not know whether a short-term inpatient treatment would be most successful in these treatment groups, too. However, not all patients might be able to handle the high pressure which is associated with this treatment. Gullestad, Johansen, Hoglend, Karterud, and Wilberg (2013) reported that day hospital treatments resulted in worse treatment effects in patients with low reflective functioning (i.e. a low level of mentalizing) compared to individual outpatient therapy. Van Manen, Horn, Stijnen, van Busschbach, and Verheul (2015) on the other hand, studied the effect of stabilizing and destabilizing treatments in PD patients with high or low psychological strengths and found slightly better outcomes in patients who followed destabilizing therapies, independent of their psychological strengths. More studies on the influence of intensity of treatment are already planned or executed. Two notable examples are the dosage trial on MBT (MBT-DOS), in which intensive outpatient MBT is compared to day hospital MBT in BPD patients (Laurenssen et al., 2014) and a study on short-term (12 weeks) inpatient DBT followed by six months of outpatient DBT compared to outpatient DBT only for BPD patients (van den Bosch, Sinnaeve, Hakkaart-van Roijen, & van Furth, 2014).

As mentioned in the introduction, most studies in PD aim to compare different theoretical treatment models, with its own theoretical factors underlying the disorder and which are the focus of treatment, such as mentalization in MBT (Bateman & Fonagy, 2004) or schemas in SFT (Young, Klosko, & Weishaar, 2003). However, differences in treatment effectiveness are rather small (Budge et al., 2013). Our results suggest that treatment modality in terms of setting does not make much difference in treatment effectiveness. Which mechanisms of change are then present in PD patients and have to be addressed in an effective treatment? A recent review on potential mechanisms of change in psy-

chotherapeutic treatments for personality disorders found that most evidence points to the importance of the therapeutic alliance and the resolution of ruptures in the alliance in the process of change (Forster, Berthollier, & Rawlinson, 2014). Further research should rather focus on the effective ingredients of psychotherapy instead of comparing different theoretical orientations of treatments and to take into account patient characteristics as predictors for treatment effectiveness (Livesley, 2012).

CONCLUSION

After five year follow-up of psychotherapy, PD patients still maintained a better level of symptoms, psychosocial functioning, and health-related quality of life than before treatment. Little differences were found between the treatment modalities. The mainly small differences which were found in the short-term were still present at the long-term, notably in cluster C, where the greater superiority of short-term inpatient treatments consisted.

ACKNOWLEDGMENTS

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

REFERENCES

Arrindell, W. A., & Ettema, J. H. M. (2003). *SCL-90. Manual for a Multidimensional Indicator of Psychopathology* [SCL-90. Handleiding bij een multidimensionele psychopathologie-indicator]. Lisse, The Netherlands: Swets & Zeitlinger.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Thunnissen, M., Ziegler, U. M., Dekker, J., … Emmelkamp, P. M. G. (2011a). Patients with cluster A Personality Disorders in Psychotherapy: An Effectiveness Study. *Psychotherapy and Psychosomatics, 80*, 28-38.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Ziegler, U. M., Dekker, J., Rossum, B. V., … Emmelkamp, P. M. G. (2011b). Effectiveness of Outpatient, Day Hospital, and Inpatient Psychotherapeutic Treatment for Patients with Cluster B Personality Disorder. *Psychotherapy and Psychosomatics, 80*, 88-99.

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., … 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., Emmelkamp, P. M. G. (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., & Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorder: Mentalisation Based Treatment*. Oxford.

Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry, 166(12),* 1355-1364.

Bateman, A. W., & Fonagy, P. (2000). Effectiveness of Psychotherapeutic Treatment of Personality Disorder. *British Journal of Psychiatry, 177,* 138-143.

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.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (rev. ed.). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.

National Health and Medical Research Council (2012). *Clinical Practice Guideline for the Management of Borderline Personality Disorder.* Melbourne: National Health and Medical Research Council.

De Beurs, E., & Zitman, F. G. (2006). The Brief Symptom Inventory (BSI): the reliability and validity of a brief alternative of the SCL-90 [De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de 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). *Structured Interview for DSM-IV personality disorders (SIDP-IV) (Dutch version)* [Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen (SIDP-IV)]. Sint Oedenrode, The Netherlands: Stichting Verslavingszorg Oost Brabant.

Derogatis, L. (1983). *SCL-90-R: Administration, Scoring and procedures manual-II.* Towson, MD: Clinical Psychometric Research.

Derogatis, L. (2011). *BSI and BSI 18. Revised Dutch edition* [Brief Symptom Inventory Herziene Nederlandse Editie]. Amsterdam, The Netherlands: Pearson.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory report. *Psychological Medicine, 13(3)*, 595-605.

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, NL: Sanders Institute EUR.

Falkenstrom, F., Granstrom, F., & Holmqvist, R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology, 60(3),* 317-328.

Ferguson, C. J. (2010). A Meta-Analysis of Normal and Disordered Personality Across the Life Span. *Journal of Personality and Social Psychology, 98(4),* 659-667.

Forster, C., Berthollier, N., & Rawlinson, D. (2014). A Systematic Review of Potential Mechanisms of Change in Psychotherapeutic Interventions for Personality Disorder. *Journal of Psychology and Psychotherapy, 4(1),* 133.

Gibbon, S., Duggan, C., Stoffers, J., Huband, N., Vollm, B. A., Ferriter, M., … Lieb, K. (2010). Psychological interventions for antisocial personality disorder. *Cochrane Database of Systematic Reviews*(6), CD007668.

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., … Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of General Psychiatry, 63(6),* 649-658.

Glenn, A. L., Johnson, A. K., & Raine, A. (2013). Antisocial Personality Disorder: A Current Review. *Current Psychiatry Reports, 15(12)*.

Gude, T., & Vaglum, P. (2001). One-year follow-up of patients with cluster C personality disorders: a prospective study comparing patients with "pure" and comorbid conditions within cluster C, and "pure" C with "pure" cluster A or B conditions. *Journal of Personality Disorders, 15(3),* 216-228.

Gunderson, J. G., Stout, R. L. , McGlashan, T. H., Shea, M. T., Morey, L. C., Grilo, C. M. … Skodol, A. (2011). Ten-year course of borderline personality disorder: Psychopathology and function from the collaborative longitudinal personality disorders study. *Archives of General Psychiatry, 68(8):* p. 827-837.

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.

Hadjipavlou, G., & Ogrodniczuk, J. S. (2010). Promising psychotherapies for personality disorders. *Canadian Journal of Psychiatry, 55(4),* 202-210.

Horn, E. K., Bartak, A., Meerman, A. M., Rossum, B. V., Ziegler, U. M., Thunnissen, M., … Verheul, R. (2014). Effectiveness of Psychotherapy in Personality Disorders Not Otherwise Specified: A Comparison of Different Treatment Modalities. *Clinical Psychology and Psychotherapy.*

Institute of Medicine (2009). *Initial National Priorities for Comparative Effectiveness Research*. Washington, D.C.: The National Academies Press.

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.

Khalifa, N., Duggan, C., Stoffers, J., Huband, N., Vollm, B. A., Ferriter, M., … Lieb, K. (2010). Pharmacological interventions for antisocial personality disorder. *Cochrane Database of Systematic Reviews*(8), CD007667.

Kvarstein, E. H., & Karterud, S. (2012). Large variations of global functioning over five years in treated patients with personality traits and disorders. *Journal of Personality Disorders, 26(2),* 141-161.

Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. 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). Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff [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). *Multidisciplinary Clinical Guideline for Personality Disorders* [Multidisciplinaire Richtlijn Persoonlijkheidsstoornissen]. Utrecht, The Netherlands: Trimbos-instituut.

Laurenssen, E., Smits, M. L., Bales, D. L., Feenstra, D. J., Eeren, H. V., Noom, M. J., … Verheul, R. (2014). Day hospital Mentalization-based treatment versus intensive outpatient Mentalization-based treatment for patients with severe borderline personality disorder: protocol of a multicentre randomized clinical trial. *BMC Psychiatry, 14(1),* 301.

Leichsenring, F., & Leibing, 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.

Lieb, K., Vollm, B., Rucker, G., Timmer, A., & Stoffers, J. M. (2010). Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. *British Journal of Psychiatry, 196(1),* 4-12.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical Analysis with Missing Data.* New York, NY, USA: John Wiley & Sons, Inc.

Livesley, 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.

Morey, L. C., & Hopwood, C. J. (2013). Stability and change in personality disorders. *Annual Review of Clinical Psychology, 9,* 499-528.

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.

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.

Samuels, J. (2011). Personality disorders: Epidemiology and public health issues. [Article]. *International Review of Psychiatry, 23(3),* 223-233.

Simon, W. (2009). Follow-up psychotherapy outcome of patients with dependent, avoidant and obsessive-compulsive personality disorders: A meta-analytic review. *International Journal of Psychiatry in Clinical Practice, 2009. 13(2)*: p. 153-165.

Skodol, A. E., Bender, D. S., Morey, L. C., Clark, L. A., Oldham, J. M., Alarcon, R. D., … Siever, L. J. (2011). Personality Disorder Types Proposed For DSM-5. *Journal of Personality Disorders, 25(2)*, 136-169.

Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2008). The burden of disease in personality disorders: diagnosis-specific quality of life. *Journal of Personality Disorders, 22(3),* 259-268.

Soeteman, D. I., Verheul, R., Delimon, J., Meerman, A. M. M. A., van den Eijnden, E., Rossum, B. V., … Kim, J. J. (2010). Cost-effectiveness of psychotherapy for cluster B personality disorders. *British Journal of Psychiatry, 196*(5), 396-403.

Soeteman, D. I., Verheul, R., Meerman, A. M. M. A., Ziegler, U. M., Rossum, B. V., Delimon, 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.

Spreeuwenberg, M. D., Bartak, A., Croon, M. A., Hagenaars, J. A., Busschbach, J. J. V., Andrea, H., … Stijnen, T. (2010). The Multiple Propensity Score as Control for Bias in the Comparison of More Than Two Treatment Arms An Introduction From a Case Study in Mental Health. *Medical Care, 48*(2), 166-174.

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.

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). Retrieved from http://repub.eur.nl/pub/8322.

van Asselt, A. D. I., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., van Dyck, R., Spinhoven, P. … Severens, J. L. (2008). Out-patient psychotherapy for borderline personality disorder: cost-effectiveness of schema-focused therapy v. transference-focused psychotherapy. *British Journal of Psychiatry, 192(6),* 450-457.

van Beek, N., & Verheul, R. (2008). Motivation for treatment in patients with personality disorders. *Journal of Personality Disorders, 22(1),* 89-100.

van den Bosch, L. M., Sinnaeve, R., Hakkaart-van Roijen, L., & van Furth, E. F. (2014). *Efficacy and cost-effectiveness of an experimental short-term inpatient Dialectical Behavior Therapy (DBT) program: study protocol for a randomized controlled trial.* Trials, 15, 152.

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.

Van Manen, J. G., Horn, E. K., Stijnen, T., Timman, R., Busschbach, J. J. V., and Verheul, R. (2015), Tailoring psychotherapy in patients with personality disorders: Matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health, 9*, 133–149.

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J., … Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability, and validity. *Psychological Assessment, 20(1),* 23-34.

Young, J. E., Klosko, J. S., & Weishaar, E. (2003). *Schema Therapy: A Practitioner's Guide*. New York, NY, USA: Guilford Press.

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2010). The 10-year course of psychosocial functioning among patients with borderline personality disorder and axis II comparison subjects. *Acta Psychiatrica Scandinavica, 122(2)*: p. 103-109.

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2012). Attainment and Stability, of Sustained Symptomatic Remission and Recovery Among Patients With Borderline Personality Disorder and Axis II Comparison Subjects: A 16-Year Prospective Follow-Up Study. *American Journal of Psychiatry, 169(5),* 476-483.

Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The Prevalence of DSM-IV Personality Disorders in Psychiatric Outpatients. *American Journal of Psychiatry, 162(10),* 1911-1918.

SUPPLEMENTARY MATERIAL

Supplementary material 2.1.
Effect sizes and clinical significant change rates at 60 months for Cluster C split up in five treatment modalities

Variable	Treatment group		N (modelling)	Mean ±SD		Effect size (Cohen's d)	N (CSC)	% CSC[1]
				Baseline	60 months			
GSI	Outpatient	long	81	1.47 ±.69	.93 ±.63	.82	56	32%*
	Day hospital	short	79	1.44 ±.64	.81 ±.69	.95	50	46%*
		long	83	1.69 ±.62	.78 ±.68	1.41	54	52%*
	Inpatient	short	61	1.74 ±.51	.60 ±.50	2.28	40	70%*
		long	94	1.79 ±.72	.88 ±.76	1.24	62	53%*

GSI=Global Severity Index, CSC=Clinical Significant Change.
*significant differences between treatment groups (p≤.05)
[1] patients both moved from a dysfunctional range to a normative range and also demonstrated reliable change

Supplementary material 2.2.
Difference scores at 60 months for five treatment groups split up in Cluster C (corrected).

Variable	Treatment group		N	Outpatient	Day hospital		Inpatient
				long	short	long	short
GSI	Outpatient	long	81				
	Day hospital	short	79	.05			
		long	83	.17	.12		
	Inpatient	short	61	.39**	.33*	.31	
		long	94	.06	.01	-.10	-.35*

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

CHAPTER 3

Effectiveness of Psychotherapy in Personality Disorders Not Otherwise Specified (PDNOS): A Comparison of Different Treatment Modalities

Eva K. Horn[a, b], Anna Bartak[c, d], Anke M. M. A. Meerman[e], Bert V. Rossum[f], Uli M. Ziegler[g], Moniek Thunnissen[h, i], Mirjam Soons[j], Helene Andrea[k, l], Elisabeth F. M. Hamers[m], Paul M. G. Emmelkamp[c, n], Theo Stijnen[o], Jan J. V. Busschbach[a, b], Roel Verheul[c, m]

[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] Department of Clinical Psychology, University of Amsterdam, Amsterdam; [d] Bos en Lommer Private Practice, Amsterdam; [e] Centre of Psychotherapy Pro Persona, Lunteren; [f] Altrecht, Zeist; [g] Zaans Medical Centre, Zaandam; [h] GGZ WNB, Bergen op Zoom; [i] Private Practice, Bergen op Zoom; [j] Arkin, Amsterdam; [k] GGZ Breburg, Tilburg; [l] Tranzo Scientific Centre for Care and Welfare, Faculty of Social and Behavioral Sciences, Tilburg University, The Netherlands; [m] De Viersprong, Netherlands Institute for Personality Disorders, Halsteren; [n] The Center for Social and Humanities Research, King Abdulaziz University, Jeddah, Saudi Arabia; [o] Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands

ABSTRACT

Objective

Although personality disorder not otherwise specified (PDNOS) is highly prevalent and associated with a high burden of disease, only a few treatment studies in this patient group exist. This study is the first to investigate the effectiveness of different modalities of psychotherapy in patients with PDNOS, i.e. short-term (up to six months) and long-term (more than six months) outpatient, day hospital, and inpatient psychotherapy.

Method

A total of 205 patients with PDNOS were assigned to one of six treatment modalities. Effectiveness was assessed over 60 months after baseline. The primary outcome measure was symptom severity, and the secondary outcome measures included psychosocial functioning and quality of life. The study design was quasi-experimental and the multiple propensity score was used to control for initial differences between treatment groups.

Results

All treatment modalities showed positive outcomes, especially in terms of improvements of symptom severity and social role functioning. At 12-month follow-up, after adjustment for initial differences between the treatment groups, short-term outpatient psychotherapy and short-term inpatient psychotherapy showed most improvement and generally outperformed the other modalities concerning symptom severity. At 60 months after baseline, effectiveness remained but observed differences between modalities mostly diminished.

Conclusion

Patients with PDNOS benefit from psychotherapy both at short-term and long-term follow-up. Short-term outpatient psychotherapy and short-term inpatient psychotherapy seem to be superior to the other treatment modalities at 12-month follow-up. At 60-month follow-up, treatments showed mostly comparable effectiveness.

INTRODUCTION

According to the DSM-IV-TR, the category of personality disorder not otherwise specified (PDNOS) can be used for "*disorders of personality functioning (…) that do not meet criteria for any one personality disorder (…), but that together cause clinically significant distress or impairment in one or more important areas of functioning*" (American Psychiatric Association, 2000). Numerous studies showed that PDNOS is one of the most prevalent mental disorders in clinical practice (Coid, Yang, Tyrer, Roberts, & Ullrich, 2006; Verheul, Bartak, & Widiger, 2007; Wilberg, Hummelen, Pedersen, & Karterud, 2008; Zimmerman & Coryell, 1989; Zimmerman, Rothschild, & Chelminski, 2005). A meta-analysis on the prevalence and use of PDNOS diagnoses showed that 3-6% of the general population and 8-13% of clinical samples met the diagnostic criteria for a PDNOS diagnosis (Verheul & Widiger, 2004). The relative prevalence, defined as the prevalence of PDNOS divided by the overall axis II percentage without PDNOS, was estimated at 21-49% (Verheul & Widiger, 2004). As is the case for patients with specific PD, the burden of disease of patients with PDNOS is high (Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008a; Soeteman, Verheul, & Busschbach, 2008b; Verheul & Widiger, 2004). In terms of quality of life, PDNOS patients report a quality of life score on the EuroQol (EQ-5D) of between .42 (negativistic PD) and .62 (mixed PD). Such EQ-5D scores represent a range comparable to patients with haemodialysis, rheumatic disease, lung cancer, Parkinson's disease, or diabetes type II (Soeteman et al., 2008b). In terms of the economic burden, a PDNOS diagnosis is associated with high costs for society (Soeteman et al., 2008a). As is the case for PD patients in general, patients with PDNOS show a wide range of problems, such as substance use or self-mutilation (Johnson et al., 2005; Verheul et al., 2007; Wilberg et al., 2008). In terms of severity of symptoms, personality problems, and relational problems, patients with PDNOS seem to fall within an intermediate position between patients without personality disorder (PD) and patients meeting the full criteria for one of the ten formal PDs in DSM-IV-TR (Verheul et al., 2007; Wilberg et al., 2008). However, patients with a diagnosis of PDNOS in addition to a formal PD diagnosis have typically higher symptom severity, more personality problems, and more relational problems than those with a formal PD diagnosis alone (Verheul et al., 2007).

Since 2000, researchers and clinicians have become increasingly interested in psychotherapy for PD patients. This has resulted in numerous studies and the development of new evidence-based therapies. There is now sufficient evidence that psychotherapy is the treatment of choice in PDs (Karterud et al., 2003; Leichsenring & Leibing, 2003). This is reflected in treatment guidelines (e.g., National Institute for Health and Clinical Excellence, 2009a, 2009b; Trimbos-institute, 2008). Specialized psychotherapeutic treatments have proven to be effective for PD in general (Gabbard, 2000; Leichsenring & Leibing,

2003), more effective than being on a waiting list or general psychiatric outpatient care, and to be associated with faster recovery rates compared to natural recovery (Bateman & Fonagy, 2008; Perry, Banon, & Ianni, 1999; Petersen et al., 2008). Approximately 75 years ago, Rosenzweig suggested that common factors were responsible for the effectiveness of psychotherapies making them equally effective in outcome. This effect was later called the 'dodo bird effect' by Luborsky et al. (e.g. Wampold et al., 1997). However, comparative trials have not yet provided conclusive evidence for the superiority of one theoretical orientation over another. Furthermore, the available studies suggest that other treatment characteristics, such as dosage (Bartak, Soeteman, Verheul, & Busschbach, 2007) or level of destabilization (van Manen, Horn, Stijnen, Busschbach, & Verheul, 2015), might be more important determinants of (cost-) effectiveness than theoretical orientation per se (e.g.Leichsenring & Leibing, 2003).

However according to our knowledge, there are no explicit treatment studies on PDNOS patient groups, despite their high prevalence and high burden of disease. Treatment studies on PD populations typically focus on formal PDs and do not report results for the PDNOS group separately (e.g., Gabbard, 2000; Winston et al., 1994; Zanarini, 2009; Zanarini, Frankenburg, Reich, & Fitzmaurice, 2010). We found one general PD treatment study reporting results separately for PDNOS patients following short-term and long-term day hospital treatment (Karterud et al., 2003) and three case studies of patients with a PDNOS diagnosis (Newman, 2002; Savoja et al., 2011; Warren, 2012). In the study of Karterud et al. (2003) the best treatment results in a day hospital were found in PD patients with a diagnosis of PDNOS, a cluster C PD, or borderline PD. Furthermore, few studies have focused on the differences in effectiveness of different treatment modalities, despite the fact that they account for large differences in costs. These differences in costs are likely to have a high impact on cost-effectiveness ratios and become still more important given current restricted health care budgets (Bartak et al., 2007). Recently, a study on the effectiveness of five treatment modalities for cluster C PDs was conducted and showed that the effectiveness of psychotherapy differed substantially across the various modalities at 12 months after baseline (Bartak et al., 2010). A similar pattern, although not statistically significant, was found in cluster B patients (Bartak et al., 2011b). This implies that treatment modality does matter, at least in the treatment of cluster B and C PD patients at relatively short-term follow-up. These findings suggest that a similar short-term effect might be present in patients with PDNOS. Since PDs are characterized by pervasive and persistent patterns of experiences and behaviors, long-term follow-up studies should also generate more information on the changeability of symptoms. For this reason, we looked for evidence that treatment modalities matter in the large study used by Bartak et al., (2010, 2011a, 2011b), selecting PDNOS patients and extending

the follow up period to 60 months. We expected a similar pattern of results compared to cluster C PDs, as the problem severity in both groups is similar (Verheul et al., 2007).

Aim of the study

The present study aimed to extend the evidence of the effectiveness of psychotherapy to patients with PDNOS. Effectiveness of six treatment modalities in patients with PDNOS, i.e. short-term (up to six months) and long-term (more than six months) outpatient, day hospital, and inpatient psychotherapy, was investigated over 60 months after baseline.

The two questions we wanted to address in this study were:

- Did PDNOS patients profit from psychotherapy in terms of severity of symptoms, relational functioning, and quality of life, and did the results remain stable over time?
- Were there differences in effectiveness between different treatment modalities in PDNOS patients, and did these differences remain stable over time?

The paper reported on the results obtained from the Study on Cost-Effectiveness of Personality Disorder Treatment (SCEPTRE), which is a large multi-centre study in the Netherlands (trial register ISRCTN: 73817429).

METHODS

Study population and design

During a 3-year period (2003-2006), 1,379 patients completed the intake procedure in six mental health centers in the Netherlands and were selected for treatment (De Viersprong Netherlands Institute for Personality Disorders, Halsteren; GGZ WNB, Bergen op Zoom and Roosendaal; Centre of Psychotherapy Pro Persona, Lunteren; Altrecht, Utrecht; Zaans Medical Centre, Zaandam; Centre of Psychotherapy Arkin, Amsterdam). The six centers offer outpatient, day hospital, and/or inpatient psychotherapeutic treatments, tailored to PD patients. Patients had to go through an extensive intake procedure prior to treatment allocation. The intake procedure consisted of one or two assessment sessions with an intake clinician, a semi-structured interview for PD diagnosis and questionnaires (see 'Assessments'), after which, allocation to treatment was discussed by an intake team and afterwards with the patient (van Manen et al., 2011). Of the 1,379 patients who participated in the intake procedure, 959 (70%) were enrolled in the SCEPTRE study (see figure 3.1). To avoid overlap with earlier studies on the effectiveness of different treatment modalities in PD patients (Bartak et al., 2010, 2011a, 2011b), patients with a diagnosis of one of the cluster PDs were excluded, which left 219 patients (23%) with a PDNOS diagnosis only who were allocated to treatment. Of these, 205 (94%) had

completed at least one follow-up assessment, had received a "*minimally effective dosage*" of psychotherapy (defined as at least two sessions of outpatient psychotherapy or at least two treatment days of day hospital or inpatient psychotherapy) and were included in the present study. All patients provided written informed consent prior to inclusion.

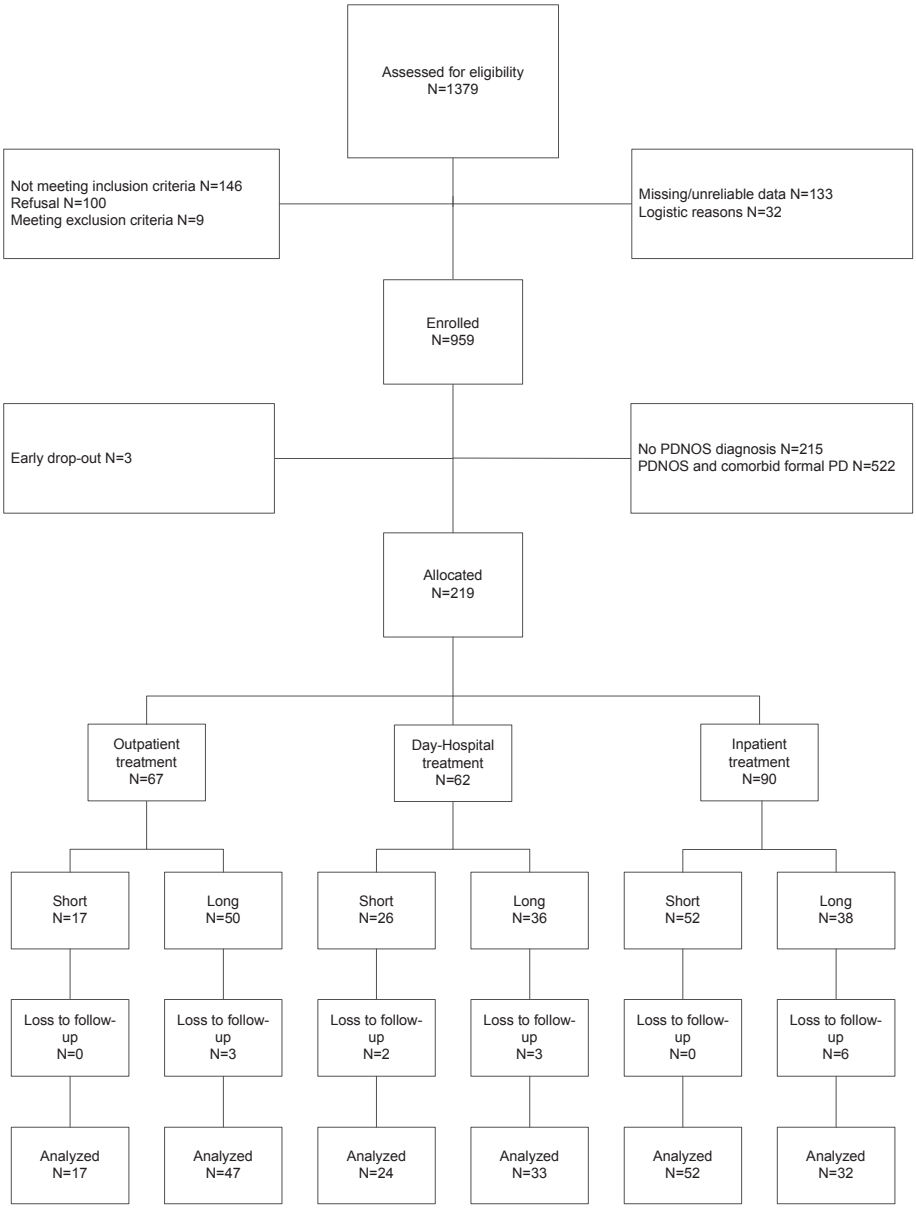


Figure 3.1. Patient Flow.

Treatments

The six mental health care centers offer a variety of psychotherapeutic treatments tailored to a PD patient population. The treatments under study can be considered highly representative of regular clinical practice for PD in the Netherlands, as therapists did not receive specific training for this study and treatment integrity was not monitored (Bartak et al., 2010). One hundred-and-one psychotherapists who were all licensed psychiatrists or psychologists participated in the study. On average, they had 14.9 ±10.1 years of postgraduate clinical experience. All treatment centers offered treatments with varying theoretical orientations, such as a psychodynamic orientation (27% of all given treatments), a cognitive-behavioral orientation (21% of all given treatments), or an integrative orientation (combining different theoretical frameworks, 52% of all given treatments).

The study focused on different treatment modalities in terms of setting and duration. The duration of treatments was defined as 'short-term' for treatments which lasted up to six months, and 'long-term' for treatments which lasted more than six months. The settings were defined as outpatient psychotherapy (i.e. individual or group psychotherapy sessions, up to two sessions per week), day hospital psychotherapy (i.e. at least one morning/afternoon per week, various forms of psychotherapeutic and psychosocial treatments, where patients sleep at home), and inpatient psychotherapy (i.e. stay at the institutions five days a week, various forms of psychotherapeutic and psychosocial treatments, where patients sleep in the institutions). 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, individual psychotherapy, milieu therapy, coaching for social problems, discussions about household tasks and living together, community meetings, and/or pharmacological treatment. Before the start of the treatment, psychotherapists were asked to register the intended treatment in terms of setting and duration. When the treatment was finished, therapists were asked once more to register the setting and duration of the actual treatment received. Since this study conformed to the intention-to-treat principle, the treatments described are the intended treatments and are not necessarily the actual treatments received.

The six treatment modality groups in this study had the following characteristics (intended treatments, mean (SD)):

1. short-term outpatient psychotherapy (N=17): sessions per week: .84 (.26), duration: 4.82 (1.27) months, 57% group therapy, 43% individual therapy;
2. long-term outpatient psychotherapy (N=47): sessions per week: .88 (.46), duration: 16.34 (7.22) months, 32% group therapy, 68% individual therapy;

3. short-term day hospital psychotherapy (N=24): days per week: 3.06 (1.31), duration: 5.75 (.74) months;
4. long-term day hospital psychotherapy (N=33): days per week: 3.02 (1.39), duration: 12.30 (2.56) months;
5. short-term inpatient psychotherapy (N=52): days per week: 5.00 (0.00), duration: 4.06 (1.43) months; and
6. long-term inpatient psychotherapy (N=32): days per week: 4.94 (0.21), duration: 11.22 (3.12) months.

Assessments

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

Diagnosis of PD

PDs were assessed 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 10 formal DSM-IV axis II diagnoses, as well as the two appendix diagnoses, i.e. depressive and negativistic PD. Furthermore, the self-defeating PD was assessed (American Psychiatric Association, 1987). Interviewers were thoroughly trained masters-level psychologists who received monthly booster sessions to avoid drift from the interviewer guidelines. Inter-rater reliability was evaluated in 25 videotaped interviews which were rated by three observer-raters. Percentage of agreement between observer-raters ranged from 84% (avoidant PD) to 100% (schizoid) (median 95%). To estimate the intraclass correlation coefficients (ICC(2,1)) for the sum of DSM-IV PD traits present (i.e. scores '2' or '3'), 25 videotaped interviews were rated by three (out of 25) random observers, which resulted in 75 observations. With analysis of variance the between patients, the between observers and the residual variance components were calculated. The ICCs were calculated as the between patients variance divided by the total variance. The ICC ranged from 0.31 (schizotypal PD) through 0.88 (depressive PD) (median 0.66). A diagnosis of PDNOS can be obtained in two ways: (1) meeting the diagnostic criteria for an appendix PD (DSM-IV-TR depressive or negativistic PD, DSM-III-R self-defeating PD) but not for any specific cluster PD, or (2) meeting the criteria for a mixed PD: meeting ten or more diagnostic criteria of various PDs but not for any specific PD. The definition of mixed PD in the current study is in agreement with some earlier studies (Pagan, Oltmanns, Whitmore, & Turkheimer, 2005; Wilberg et al., 2008) but not with other studies in which a cut-off of only five criteria was suggested (Coccaro, Nayyer, & McCloskey, 2012; Verheul et al., 2007). Since in clinical practice mostly ten or more criteria have to be met, we chose to use this definition.

Outcome measures

The primary outcome measure was symptom severity. This was measured using the Dutch version of the Brief Symptom Inventory (De Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983), a validated self-report scale derived from the revised Symptom Checklist-90 (Arrindell & Ettema, 2003; Derogatis, 1983). 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 zero to four. Cronbach's Alpha was α=.96. The secondary outcome measures included psychosocial functioning and quality of life. Psychosocial functioning was measured using two subscales of the Outcome Questionnaire-45 (OQ-45), i.e. Interpersonal Relations and Social Role (Lambert et al., 2004). Cronbach's Alpha was α=.74 for Interpersonal Relations and α=.58 for Social Role. Furthermore, health-related quality of life was measured using the EQ-5D (EuroQolGroup, 1995). A recent study in the Netherlands elicited valuations for 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). Cronbach's was α=.50. All outcome measures were assessed at baseline and several follow-up points. Three treatment centers conducted their assessments at baseline, end of treatment, at six and 12 months after the end of treatment, and at 36 and 60 months after baseline. Three other centers conducted their assessments at baseline and at 12, 24, 36, and 60 months after baseline. Different assessment points were used due to logistic reasons.

Additional baseline measures

The assessment battery at baseline included three supplementary instruments to assess variables that were used as potential confounders for the propensity score (PS) estimation. First, the Dutch version of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ) was used to measure type and degree of personality pathology in four domains, i.e. emotional dysregulation, dissocial behavior, inhibition, and compulsivity (Livesley, 2002; van Kampen, 2002). 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 core components of personality pathology were measured using the Severity Indices of Personality Problems-Short Form-118 (SIPP-118), a 118-item questionnaire aimed to measure five core domains of personality pathology, i.e. Self Control, Identity Integration, Responsibility, Relational Functioning, and Social Concordance (Verheul et al., 2008).

Data completeness

Follow-up response was high, thereby enhancing the robustness of the multi-level analyses. We included all patients with at least one follow-up measure as multi-level models make optimal use of incomplete repeated measures data and are robust for

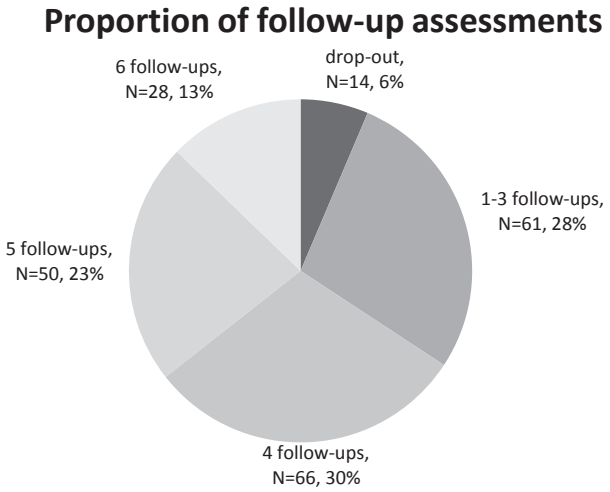


Figure 3.2. Follow-up ratios.

selective drop-out when the drop-out is missing at random (Little & Rubin, 1987). Figure 3.2 illustrates the proportion of patients who provided different numbers of follow-up assessments.

Drop-outs (6%) and completers (94%) differed significantly on only one baseline measure of the four assessed outcome variables (OQ-45 Social Role: $t(213)=2.54$, $p=.012$). Drop-outs had a higher mean score ($M=18.00$, $SD=3.36$) compared to completers ($M=14.59$ $SD=4.75$), which implies that they experienced more difficulties adjusting to their social role.

Statistical analyses

We used multilevel modeling to deal with: (1) the dependency of the repeated measures within the same subject in time, and (2) the longitudinal data with observations unequally spaced in time (see 'Outcome measures').

First, we examined the uncorrected results on all outcome measures over 60 months after baseline in 12 month intervals. To estimate the uncorrected treatment effects over 60 months after baseline, we used a random intercept and random slope model 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: time, a spline (knot point at three years), dummy variables indicating group membership, interaction between time and group membership, and interaction between spline and group membership. Subsequently, we calculated within-group effect sizes (ES, Cohen's d) using the estimated pooled standard deviations from the models to describe change from baseline over 60 months in each group (Cohen, 1988).

Second, since this is a non-randomized study, the comparisons of the groups had to be corrected for the influence of potential confounders, i.e. initial patient differences between treatment groups that are related to outcome. We included a multiple propensity score in our analyses to adjust for these differences and avoid bias in effect estimation. The classic propensity score is defined as the conditional probability of assignment to one of two treatment groups given a set of observed pre-treatment variables (Bartak et al., 2009; Rosenbaum & Rubin, 1983). The PS is designed to reduce selection bias by equating groups based on their pre-treatment variables and to reduce the number of covariates by combining these pre-treatment variables into one PS. By using the PS, the random assignment to one of two treatments is imitated as the PS adjusts for (observed) baseline differences between patients in different treatment groups. After adjustment, the distribution of these pre-treatment variables is similar in the two treatment groups (for an illustration see Bartak et al., 2009). The multiple PS is an extension of the classic propensity score to more than two treatment groups and its feasibility in mental health research has been illustrated earlier (Spreeuwenberg et al., 2010). To identify relevant confounders, we considered a broad list of social, economic, and diagnostic variables carefully selected by both clinicians and researchers, which were based on the literature and clinical knowledge (Bartak et al., 2009). Only pre-treatment characteristics significantly related to the studied outcome variables were used to estimate the multiple propensity scores. A simulation study of Brookhart et al. (2006) showed that the addition of variables which were related to treatment but which were not related to the outcome led to an increase of the variance of the estimated treatment effect and removed only a small amount of bias. The multiple propensity scores were achieved by a multinomial regression analysis, with group membership as the dependent variable and pre-treatment variables significantly related to the outcome as independent variables (see appendix table). The propensity score method allows the possibility of visualization and judgment of the overlap in propensity score distributions (and thus the overlap in relevant variables) between treatment groups. The overlap on the PS distributions showed that these were sufficient for the comparison of the six treatment groups. Furthermore, significance testing on variables related to the outcome showed that by correcting for the multiple PS initial significant differences between treatment groups

disappeared. In earlier investigations in patients with cluster A, B, and C PDs, this visual judgment led to a less firm interpretation of the results (cluster A) or to the exclusion of treatment groups (cluster C), as the propensity score distributions were too far apart (Bartak et al., 2010, 2011a, 2011b).

After the analyses of the uncorrected results with the first multi-level model, a second multilevel model was used to compare change in outcome variables across treatment groups. This model included the multiple propensity scores to correct for initial patient differences. Dependent variables were the change scores (from baseline) observed during follow-up for each of the outcome measures. Independent variables were time, a spline (knot point at three years), dummy variables indicating group membership, the multiple propensity scores, interaction between time and group membership, interaction between spline and group membership, and the mutual interactions of the multiple propensity scores. This model was used to estimate differences in change scores over 60 months after baseline in pair wise comparisons of the six treatment groups. Unstandardized coefficients (b) were used to describe the size of the effect.

The analyses of outcomes were based on the intention-to-treat (ITT) principle. The analyses were performed using IBM SPSS Statistics 20.0 for data preparation and the estimation of the multiple propensity scores, and SAS 9.2 for multilevel modeling. The study protocol was approved by the Medical Ethics Committee of the Erasmus University Medical Centre in Rotterdam, the Netherlands.

RESULTS

Sample

The sample consisted of 205 patients with a diagnosis of PDNOS and at least one follow-up measurement (see figure 3.1). The mean age of the sample was 35.1 (SD=10.3) years, and 72% of the sample was female. Educational level was high (European Qualifications Framework [EQF]≥6) for 33%, medium (EQF 3 to 5) for 43%, and low (EQF≤2) for 23% (van der Sanden, Smit, & Dashorst, 2012). About one quarter (24%) of the patients were married or were in steady relationships, and about one quarter (27%) of the sample lived together with children. Two-thirds (66%) of the sample was employed or was studying (see table 3.1). One hundred and thirty-four patients (65%) met the criteria for a mixed PD only, 34 patients (17%) met the criteria for an appendix diagnosis only, and 37 patients (18%) met the criteria for both a mixed and an appendix PD. On average, 14.6 (6-29) diagnostic criteria were fulfilled. For patients with a mixed PD, traits from one or more of the appendix diagnoses were most common (M=5.4 SD=2.4), followed by

Table 3.1. Sociodemographic and clinical characteristics of patients.

Sociodemographic and clinical variables	Short Outpatient	Long Outpatient	Short Day hospital	Long Day hospital	Short Inpatient	Long Inpatient	Overall
	N=17	N=47	N=24	N=33	N=52	N=32	N=205
	Mean (SD)						
Age*	35.6 (8.9)	36.9 (10.4)	36.2 (10.0)	30.9 (10.2)	39.5 (9.6)	28.4 (7.7)	35.1 (10.3)
	N (%)						
Gender							
Female	13 (77)	32 (68)	19 (79)	23 (70)	36 (69)	24 (75)	147 (72)
Education							
High ($EQF^+ \geq 6$)	2 (12)	14 (30)	9 (38)	9 (27)	24 (46)	10 (31)	68 (33)
Medium (EQF^+ 3-5)	8 (47)	20 (43)	10 (42)	14 (42)	19 (37)	18 (56)	89 (43)
Low ($EQF^+ \leq 2$)	7 (41)	13 (28)	5 (21)	10 (30)	9 (17)	4 (13)	48 (23)
Living alone or together*							
Alone	6 (35)	18 (38)	7 (29)	20 (61)	26 (50)	17 (53)	94 (46)
With partner	9 (53)	25 (53)	15 (63)	6 (18)	24 (46)	2 (6)	81 (40)
With parent(s)	1 (6)	4 (9)	—	7 (21)	—	8 (25)	20 (10)
With other people	1 (6)	—	2 (8)	—	2 (4)	5 (16)	10 (5)
Children*							
Care for child(ren)	8 (47)	20 (43)	7 (29)	4 (12)	15 (29)	2 (6)	56 (27)
Civil status*							
Married/steady relationship	6 (35)	14 (30)	9 (38)	3 (9)	14 (27)	3 (9)	49 (24)
Divorced/widowed	5 (29)	14 (30)	2 (8)	4 (12)	11 (21)	—	36 (18)
Never married	6 (35)	19 (40)	13 (54)	26 (79)	27 (52)	29 (91)	120 (59)

Table 3.1. Sociodemographic and clinical characteristics of patients. (continued)

Sociodemographic and clinical variables	Short Outpatient	Long Outpatient	Short Day hospital	Long Day hospital	Short Inpatient	Long Inpatient	Overall
	N=17	N=47	N=24	N=33	N=52	N=32	N=205
Mode of employment							
Employed/student	14 (82)	33 (70)	18 (75)	16 (49)	34 (65)	20 (63)	135 (66)
Previous treatment							
Outpatient*	12 (71)	30 (64)	20 (83)	26 (79)	47 (90)	28 (88)	163 (80)
Inpatient	2 (12)	6 (13)	2 (8)	4 (12)	9 (17)	8 (25)	31 (15)
Medication*	3 (18)	16 (34)	9 (38)	20 (61)	23 (44)	19 (59)	90 (44)
PDNOS diagnosis							
PD mixed only	14 (82)	32 (68)	18 (75)	23 (70)	30 (58)	17 (53)	134 (65)
Appendix PD only	2 (12)	8 (17)	1 (4)	7 (21)	10 (19)	6 (19)	34 (17)
PD mixed & appendix PD	1 (6)	7 (15)	5 (21)	3 (9)	12 (23)	9 (28)	37 (18)
Appendix PD diagnosis							
Depressive PD	2 (12)	13 (28)	6 (25)	9 (27)	19 (37)	15 (47)	64 (31)
Negativistic PD	1 (6)	3 (6)	—	—	1 (2)	—	5 (2)
Self-defeating PD	—	—	—	1 (3)	4 (8)	—	5 (2)
	Mean (SD)						
Comorbidity in PD mixed							
Cluster A traits	2.0 (2.1)	2.0 (2.0)	1.4 (1.3)	1.8 (2.2)	0.9 (1.6)	1.3 (1.3)	1.5 (1.8)
Cluster B traits	2.1 (2.2)	2.9 (2.5)	4.2 (2.7)	3.4 (2.5)	3.1 (2.4)	3.5 (2.2)	3.2 (2.5)
Cluster C traits	4.3 (2.1)	4.6 (1.9)	5.0 (2.2)	4.3 (1.8)	5.2 (1.7)	5.3 (2.0)	4.8 (1.9)
Appendix traits	4.0 (2.0)	5.6 (2.2)	5.3 (2.6)	5.3 (2.1)	5.7 (2.6)	5.8 (2.3)	5.4 (2.4)
Number of traits	12.3 (2.5)	14.9 (3.9)	15.8 (4.9)	14.2 (3.9)	14.3 (3.8)	15.2 (3.5)	14.6 (3.9)

*significant differences between six treatment groups ($p<.05$)

[+] European Qualifications Framework

traits from cluster C (M=4.8 SD=1.9) and traits from cluster B (M=3.2 SD=2.5). The least number of traits came from cluster A (M=1.5 SD=1.8). Of the patients with an appendix diagnosis, the majority (64 patients, 90%) had a depressive PD, five patients (7%) had a self-defeating PD, and five patients (7%) had a negativistic PD.[1] A significant difference was found on one of the assessed outcome scales at baseline (OQ Interpersonal Relations) between patients with a mixed PD only (M=18.38 SD=5.79), patients with an appendix PD only (M=17.70 SD=6.16), and patients with both a mixed PD and an appendix PD (M=21.44 SD=5.87; F(2,202)=4.68, p=.01). Patients with both diagnoses reported a more dysfunctional interpersonal functioning compared to patients with a mixed PD (p=.006) or appendix PD only (p=.008).

Treatment adherence

Forty-five patients (22%) changed their intended treatment group: Of these, 38 patients stayed in the same setting, but for a longer (N=16) or shorter (N=22) period than planned. Three patients changed their treatment setting and their treatment lengths and four patients changed their treatment setting only. Of these, five patients followed a less intensive treatment and two patients a more intensive treatment (in terms of setting). When the intended treatment was stopped earlier than planned, this was mostly done in agreement with the treatment staff [N=13 (48%)], or because of the patient dropping out [N=13 (48%)]. One patient was forced to leave earlier by the treatment staff (4%).

Effectiveness results

Uncorrected outcomes

Sixty months after baseline, within-group effect sizes of the uncorrected scores of symptom severity (GSI) ranged from .91 (large effect, short-term outpatient) to 1.42 (very large effect, short-term day hospital; see table 3.2). A positive significant change was found for all treatment modalities. See figure 3.3 for the course of the GSI scores.

Improvements also appeared in terms of psychosocial functioning and quality of life. Effect sizes for these outcome measures were somewhat smaller compared to symptom severity, but a positive significant change was evident in most measures at 60 months (except for OQ-45 Interpersonal Relations in short-term day hospital, and EQ-5D in short-term outpatient and short-term day hospital). Effect sizes ranged from .30 (small effect, OQ-45 Interpersonal Relations, short-term day hospital) to 1.46 (very large effect, OQ-45 Interpersonal Relations, long-term inpatient) (see table 3.3).

1 Because patients could have more than one appendix diagnosis, the sum of the prevalence is higher than 100%.

Table 3.2. Within-group effect sizes over 60 months, primary outcome GSI (uncorrected).

Treatment group		N	Mean (SD)						Within-group effect size (Cohen's d)				
			Baseline*	12m	24m	36m	48m	60m	12m	24m	36m	48m	60m
Outpatient	short	17	1.05 (.44)	0.49 (0.43)	0.48 (0.47)	0.47 (0.52)	0.52 (0.57)	0.57 (0.63)	1.33	1.29	1.24	1.07	0.91
	long	47	1.21 (.59)	0.72 (0.67)	0.70 (0.64)	0.67 (0.60)	0.63 (0.56)	0.59 (0.51)	0.78	0.84	0.92	1.02	1.14
Day hospital	short	24	1.37 (.57)	0.75 (0.59)	0.68 (0.58)	0.61 (0.57)	0.59 (0.56)	0.58 (0.57)	1.09	1.23	1.36	1.41	1.42
	long	33	1.53 (.62)	0.83 (0.60)	0.82 (0.64)	0.80 (0.68)	0.76 (0.73)	0.71 (0.79)	1.17	1.14	1.14	1.15	1.17
Inpatient	short	52	1.40 (.53)	0.54 (0.48)	0.63 (0.52)	0.72 (0.57)	0.68 (0.61)	0.63 (0.65)	1.72	1.48	1.25	1.27	1.31
	long	32	1.59 (.54)	0.95 (0.63)	0.92 (0.65)	0.89 (0.67)	0.82 (0.69)	0.75 (0.71)	1.11	1.14	1.17	1.26	1.35

*means differ significantly at baseline (p<.05)

Table 3.3. Within-group effect sizes at 60 months, secondary outcomes (uncorrected).

Variable	Treatment group		N	Mean (SD)		Within-group effect size (Cohen's d)
				Baseline	60months	
OQ-45 Social Role	Outpatient	short	17	13.27 (5.01)	8.14 (5.56)	1.00
		long	47	13.90 (4.37)	10.08 (3.11)	1.02
	Day hospital	short	24	13.56 (4.80)	9.58 (5.39)	0.80
		long	33	14.35 (4.66)	10.15 (5.09)	0.87
	Inpatient	short	52	15.90 (4.80)	10.53 (5.26)	1.08
		long	32	15.16 (4.55)	9.69 (4.51)	1.23
OQ-45 Interpersonal Relations*	Outpatient	short	17	15.61 (3.85)	10.78 (4.15)	1.24
		long	47	18.55 (6.52)	12.83 (5.65)	0.95
	Day hospital	short	24	15.64 (6.01)	13.80 (6.59)	0.30
		long	33	19.50 (6.35)	14.31 (5.33)	0.90
	Inpatient	short	52	19.39 (5.71)	13.51 (6.40)	0.98
		long	32	21.69 (4.31)	13.45 (6.86)	1.46
EQ-5D	Outpatient	short	17	0.68 (0.20)	0.79 (0.18)	0.60
		long	47	0.67 (0.21)	0.82 (0.16)	0.81
	Day hospital	short	24	0.57 (0.28)	0.68 (0.21)	0.45
		long	33	0.58 (0.28)	0.75 (0.30)	0.59
	Inpatient	short	52	0.57 (0.28)	0.73 (0.27)	0.59
		long	32	0.53 (0.25)	0.75 (0.23)	0.93

*means differ significantly at baseline (p<.05)

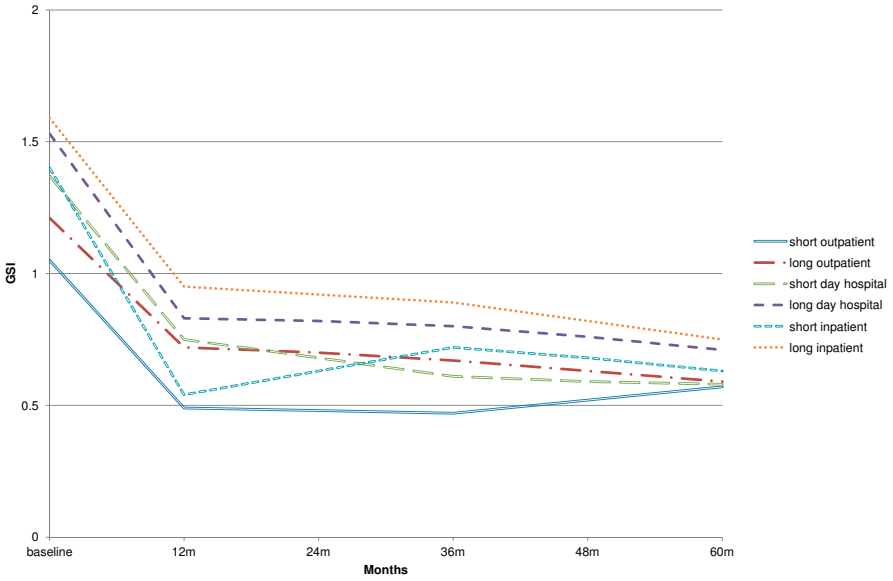


Figure 3.3. Within-group results on the GSI.

Corrected outcomes

We concluded that the overlap in the propensity scores was sufficient for all between-group comparisons, indicating that the propensity score was capable to correct for initial baseline differences (see 'Methods'). After correction for all relevant pre-treatment differences with the multiple propensity score, no significant differences were found between treatment modalities concerning either of the outcome measures at 60-month follow-up (see table 3.4). However, analyses on the intervening time points on the primary outcome measure did show significant differences up to 36-month follow-up. Strongest differences were found at 12-month follow-up (less at 24 months), mainly due to the low gains of long-term inpatient psychotherapy. Significant differences were found in favor of short-term outpatient psychotherapy compared to long-term inpatient psychotherapy (b=-.57, p=.01), long-term outpatient psychotherapy compared to long-term inpatient psychotherapy (b=-.34, p=.05), short-term inpatient psychotherapy compared to short-term day hospital (b=.33, p=.02), and finally short-term inpatient psychotherapy compared to long-term inpatient psychotherapy (b=-.60, p=.00). Over time, these differences diminished and from 48 month follow-up onwards, all significant differences had disappeared. Concerning the secondary outcome measures, no significant differences at 60 months after baseline were found.

Table 3.4. Difference scores (b) at 60 months (corrected).

Variable	Treatment group		N	Outpatient		Day hospital		Inpatient
				short	long	short	long	short
GSI	Outpatient	short	17					
		long	47	-0.04				
	Day hospital	short	24	-0.08	-0.04			
		long	33	-0.08	-0.04	0.00		
	Inpatient	short	52	-0.01	0.03	0.06	0.06	
		long	32	-0.30	-0.27	-0.23	-0.23	-0.29
OQ-45 Social Role	Outpatient	short	17					
		long	47	-2.67				
	Day hospital	short	24	-2.42	0.25			
		long	33	-2.05	0.62	0.37		
	Inpatient	short	52	-2.30	0.37	0.13	0.51	
		long	32	-1.74	0.93	0.68	-1.34	0.56
OQ-45 Interpersonal Relations	Outpatient	short	17					
		long	47	-1.72				
	Day hospital	short	24	-4.74	-3.01			
		long	33	-3.68	-1.96	1.06		
	Inpatient	short	52	-2.04	-0.32	2.69	3.40	
		long	32	-3.63	-1.91	1.10	-5.14	-1.59
EQ-5D	Outpatient	short	17					
		long	47	0.04				
	Day hospital	short	24	-0.07	-0.12			
		long	33	-0.02	-0.06	0.05		
	Inpatient	short	52	-0.01	-0.05	0.06	0.11	
		long	32	-0.05	-0.09	0.03	-0.19	-0.04

GSI=Global Severity Index. OQ-45=Outcome Questionnaire-45. EQ-5D=EuroQol-5D.
b= unstandardized beta coefficients. Positive coefficients indicate that the treatment group shown in the left column is superior, negative coefficients indicate that the treatment group in the above row is superior.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The results of this study were based on the classification of treatments on the basis of intention to treat (see methods section). When completion of the intended treatment was introduced as a covariate, no significant differences were found between patients who followed their intended treatment and the ones who did not on the primary outcome measure (GSI). There was a trend that the initial (statistical insignificant) differences between these two groups became even smaller over time (b=.17, p=.07 at 12-month follow-up, b=-.01, p=.95 at 60-month follow-up). In the secondary outcome measures, the same pattern and insignificant differences were found.

DISCUSSION

This is the first study to compare the effectiveness of various treatment modalities in patients with PDNOS. Patients in all treatment modalities showed positive outcomes at short- and long-term follow-up, especially in terms of improvements of symptom severity and social role functioning. This study provides evidence that PDNOS patients profit from psychotherapy and that accomplished changes are maintained over time. Strongest differences were found at 12-month follow-up. In terms of symptom severity, short-term outpatient psychotherapy, long-term outpatient and short-term inpatient psychotherapy was found to be superior to long-term inpatient psychotherapy. Short-term inpatient psychotherapy was found to be superior to short-term day hospital treatment. At 60-month follow-up, the observed differences between modalities were diminished. The higher effect sizes of the short-term outpatient and short-term inpatient psychotherapies decreased, while the lower effect sizes of the long-term outpatient and long-term inpatient psychotherapies increased. Concerning day hospital treatment, we found a deviant course over time, i.e. an increase in effect size of the short-term modality, and a stable effectiveness over time for the long-term modality.

Embedding of previous studies

To the best of our knowledge, the study of Karterud et al. (2003) on day hospital treatment in PD patients is the only study reporting treatment results separately for PDNOS patients (Karterud et al., 2003). The admission GSI score of PDNOS patients in their study (M=1.4, sd=.6) was comparable to the baseline GSI score in our day hospital patients (short-term day hospital M=1.37, sd=.57, long-term day hospital M=1.53, sd=.62). However mean GSI scores at one year follow-up were lower in the current sample (short-term day hospital M=.75, sd=.59/long-term day hospital M=.83, sd=.60) and effect sizes were therefore higher than in the sample of Karterud et al (M=1.1, sd=.7). It is difficult to explain where these differences came from, given the same GSI score at baseline and while other variables were unknown.

Within the same SCEPTRE project, three studies were published on the effectiveness of different modalities of psychotherapy in patients of the DSM-IV-TR cluster A, B, and C PDs, respectively. These studies revealed invariably positive effects at 12 or 18 months after baseline (long-term follow-up is not available, yet, (Bartak et al., 2010, 2011a, 2011b). A comparison with these populations showed that PDNOS patients reported the healthiest scores on all outcome measures at baseline, thus reporting the least psychological distress, a better psychosocial functioning, and a higher quality of life, which is in accordance with earlier studies (Verheul et al., 2007; Wilberg et al., 2008). Given the different follow-up time points (12 and 18 months) and treatment groups (three, five and

six groups), it was possible to compare the effectiveness between PDNOS and cluster A, B, and C patient groups in a general way only. This comparison demonstrated that concerning symptom severity, PDNOS patients in outpatient and inpatient psychotherapies showed higher effect sizes compared to cluster A, B, and C PD patients. PDNOS patients in day hospital psychotherapies, however, seemed to benefit more compared to cluster A and C patients, but benefitted less compared to cluster B patients. Furthermore, three long-term studies on the effectiveness of psychotherapeutic treatments lasting one to three years in PDs showed results comparable to the current study. First, in the study of Chiesa & Fonagy (2003), and Chiesa, Fonagy, & Holmes (2003), a steady drop in the GSI score and therefore a steady increase in effectiveness was found in a group of PD patients of which more than half met the criteria for a PDNOS diagnosis (Chiesa & Fonagy, 2003; Chiesa, Fonagy, & Holmes, 2003). Treatments consisted of an inpatient and a step-down program (consisting of inpatient treatment which was followed by outpatient treatment) and results were presented at three and six years follow-up. Second, in studies on mainly BPD patients, Giesen-Bloo et al. (2006) showed stable differences for outpatient schema-focused therapy vs. transference-focused therapy at three years follow-up, and finally Bateman & Fonagy (2008) showed stable effectiveness for day hospital MBT vs. treatment as usual (general psychiatric services) at eight years follow-up. These studies support the notion that PDNOS patients have a similar development in psychotherapy compared to other PDs.

Possible explanations of findings

Generally, we found higher but decreasing effects in short-term outpatient and short-term inpatient psychotherapies, and lower but increasing effects in long-term outpatient and long-term inpatient psychotherapies. This led to insignificant differences between treatment groups at long-term follow-up which is in line with earlier effectiveness studies on psychotherapeutical treatments and supports the conclusion that differences in effectiveness of active treatments in PD are negligible, also called "*equivalence effect*" or "*dodo-bird effect*" (Budge et al., 2013; Wampold et al., 1997). Five possible explanations for the results found are set out hereafter.

First, the axis-I study of Knekt et al. (2008) on the effectiveness of short-term (i.e., five to six months) and long-term (i.e., up to three years) psychodynamic psychotherapy for patients with mood and anxiety disorders showed similar results. Within the first year after the start of the treatment, the short-term psychotherapy group showed better results than the long-term group in terms of psychiatric symptoms, while at three year follow-up the long-term group even exceeded the short-term group in terms of treatment success. Knekt et al. explained this difference as being due to the difference in duration of treatment which influences the orientation and aim of therapy. In short-term

therapies, an active and problem-based orientation is applied, which patients can make use of, and profit from, in a short period of time. On the other hand, long-term therapies aim for "*more global changes by affecting the patient's long-term vulnerability to stressors*", requiring more time to achieve good results.

Second, explanations for the superiority of the short-term psychotherapies in the short-term are that these typically have a highly structured format, a consistently applied theoretic orientation, and are focused with respect to treatment targets right from the start of treatment. Such aspects are typically less pronounced in long-term treatments. Both the level of structure and the consistent application of a comprehensible, coherent theoretical orientation can be considered key components of effective psychotherapies (Bateman & Fonagy, 2000).

Third, high pressure in the therapy might accelerate the therapy effect. However, high pressure is difficult to imagine bearable in long-term treatment. Short-term treatments are therefore in a better position to use high pressure, facilitating changes to come more quickly.

Fourth, since this PDNOS population is characterized by slightly milder symptoms compared to patients with formal PD diagnoses, potential iatrogenic effects could occur in long-term hospitalization which could account for the poorer short-term effects for the long-term inpatient psychotherapy (Fonagy & Bateman, 2006).

Fifth, evidence suggested that the best outcome is usually apparent a few weeks or months after the termination of treatment, followed by some relapse and a consolidation of results (Perry et al., 1999). This would give the short-term treatments an advantage on short-term follow-up, as these treatments had already finished and could explain part of the results found.

Strengths and limitations

The most important strength is the inclusion of a large number of patients as well as the rather equal division of patients over various treatment modalities. Another strength is its representativeness and therefore high external validity due to the naturalistic design of the study, a minimal set of exclusion criteria, and a long-term follow-up period of five years.

This study has also several limitations. First, the study is limited by the fact that patients were not randomized. However, comparisons between markedly different treatment modalities or dosages, as were made in this study, are not feasible in a randomized

study. Most patients would probably refuse to be randomly assigned to a condition which consists of either three months outpatient psychotherapy or 12 months inpatient psychotherapy (an example of the greatest contrast). However, this limitation is mitigated by the fact that we controlled as rigorously as possible for initial patient differences as potential confounders by means of the multiple propensity score. It is important to acknowledge that this score corrects for observed differences only, and it cannot correct for unobserved differences. Therefore it is possible that the treatment groups differed on aspects that we did not measure and therefore could not control for. However, we controlled for a substantial number of social and diagnostic variables, which minimizes the possibility that important variables were overlooked. Second, another limitation is the difference of loss to follow-up which was higher in the long-term psychotherapies compared to the short-term psychotherapies. This concern is somewhat mitigated since we found no significant differences in the outcome measures at baseline between patients with and without follow-up, with one exception, and the general low loss to follow-up. Third, this study focused on treatment dosage and did not take into account other treatment attributes such as the potential impact of theoretical orientation and medication use or patient attributes, such as Axis I comorbidity. However, we came across two complications to investigate the impact of theoretical orientation. First, more than 50% of the given treatments were considered integrative treatments (e.g., Schema-Focused Therapy (SFT), Dialective Behavior Therapy (DBT) and Mentalization Based Treatment (MBT)). Second, as theoretical orientation was associated with setting and duration, a different research design would have been necessary to explore the effect of theoretical orientations in addition to treatment modalities. Furthermore, studies for PD typically show that theoretical orientations only account for small differences in effectiveness (Bartak et al., 2007). Nevertheless, we consider this study to be a starting point for further research. Future studies are recommended to take these other treatment and patient attributes into account. Fourth, the absence of a control condition without psychotherapy or a placebo condition makes it less straightforward to conclude that psychotherapy works in patients with PDNOS. Part of the effectiveness could also be due to natural recovery or regression towards the mean. Nevertheless, in the review of Perry et al. (1999) it was concluded that psychotherapeutic treatments may lead to a sevenfold increase in speed of recovery in comparison with natural recovery. Fifth, a limitation is that effectiveness is determined by self-report. From the present study we know that patients report less complaints and a better functioning, but we do not have information whether patients still meet criteria for a PD diagnosis after five years. Sixth and finally, sites overlapped only partially in terms of the (equal) availability of the six modalities. Disentangling the site effect would reduce statistical power. The reason why we did not do so is that we could not think of any valid reason why, given the specialist

expertise of the six sites involved in this study, sites would have a clinically important and interpretable impact on treatment effectiveness.

Clinical and scientific implications

As this is the first large-scale treatment study in patients with PDNOS, it is difficult to translate the results into definitive conclusions and strong recommendations for clinical practice. It might be that short-term psychotherapeutic treatments are preferred for this patient population since patients in these treatment modalities improved quickly. Furthermore, since all treatment modalities seemed to be equally effective in the long run, it might be preferable to make treatment choices based on pragmatic reasons. Therefore a treatment with the least impact on every-day life and - for economic reasons - the least costly treatment could be the way forward.

Our results have several implications for future research. First, a cost-effectiveness study is essential to inform policy makers in a budget-constrained health care system. The cheapest treatment is not necessarily the most cost-effective treatment. Patients who have followed an intensive – and therefore expensive – but effective treatment might function better and consume less additional care after treatment compared to patients who followed less intensive and less effective treatments. In other words, intensive and expensive treatments might provide more long term benefits thereby ultimately proving superior to cheaper and less intensive therapies. Therefore, a state-of-the art cost-effectiveness analysis should be performed, as has already been carried out for cluster B and cluster C PDs (Soeteman et al., 2010; Soeteman et al., 2011). Since the costs of short-term treatments are generally lower compared to their longer counterparts (especially compared to long-term inpatient psychotherapy) and they show a rather high effectiveness, it might be expected that short-term outpatient and inpatient psychotherapies in PDNOS are more cost-effective than longer-term alternatives. Second, subgroup analyses directed at 'what works for whom' could give more valuable information for clinical practice about which treatments work best for which category of patients instead for which category of diagnoses. This is even more important in this patient group since various definitions of PDNOS are used in clinical practice and across studies, limiting the comparability and generalizability of study findings (Coccaro et al., 2012; Verheul & Widiger, 2004; Wilberg et al., 2008). Third, more research about the effect of dosage would be valuable. In the current study, dosage was defined as a product of setting in three and duration in two stratifications. The study of a further division in smaller units of time could give more useful information for clinical practice.

Conclusions

In conclusion, based on long-term follow-up, all treatment modalities were shown to be approximately equally effective treatments in a PDNOS population. However, short-term outpatient and short-term inpatient psychotherapy, limited to a duration of six months, displayed higher effect sizes on short-term follow-up. This study provides the first evidence for the benefit of different treatment modalities for one of the most prevalent mental disorders in clinical practice: PDNOS. Further state-of-the-art cost-effectiveness studies will inform us about the most cost-effective treatments.

ACKNOWLEDGMENTS

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

REFERENCES

Arrindell, W. A., & Ettema, J. H. M. (2003). *SCL-90. Manual for a Multidimensional Indicator of Psychopathology* [SCL-90. Handleiding bij een multidimensionele psychopathologie-indicator]. Lisse, The Netherlands: Swets & Zeitlinger.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. Washington, DC: American Psychiatric Association.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Thunnissen, M., Ziegler, U. M., Dekker, J., … Emmelkamp, P. M. G. (2011a). Patients with cluster A Personality Disorders in Psychotherapy: An Effectiveness Study. *Psychotherapy and Psychosomatics, 80*, 28-38.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Ziegler, U. M., Dekker, J., Rossum, B. V., … Emmelkamp, P. M. G. (2011b). Effectiveness of Outpatient, Day Hospital, and Inpatient Psychotherapeutic Treatment for Patients with Cluster B Personality Disorder. *Psychotherapy and Psychosomatics, 80*, 88-99.

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., … 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., Emmelkamp, P. M. G. (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., & Fonagy, P. (2008). 8-year Follow-up of Patients treated for Borderline Personality Disorder: Mentalization-based Treatment versus Treatment as Usual. *American Journal of Psychiatry, 165(5),* 631-638.

Bateman, A. W., & Fonagy, P. (2000). Effectiveness of Psychotherapeutic Treatment of Personality Disorder. *British Journal of Psychiatry, 177*, 138-143.

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.

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., & Holmes, J. (2003). When less is more: An exploration of psychoanalytically oriented hospital-based treatment for severe personality disorder. *The International Journal of Psychoanalysis, 84(3),* 637-650.

Coccaro, E. F., Nayyer, H., & McCloskey, M. S. (2012). Personality disorder-not otherwise specified evidence of validity and consideration for DSM-5. *Comprehensive Psychiatry, 53(7),* 907-914.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (rev. ed.). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.

Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry, 188,* 423-431.

De Beurs, E., & Zitman, F. G. (2006). The Brief Symptom Inventory (BSI): the reliability and validity of a brief alternative of the SCL-90 [De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de 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). *Structured Interview for DSM-IV personality disorders (SIDP-IV) (Dutch version)* [Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen (SIDP-IV)]. Sint Oedenrode, The Netherlands: Stichting Verslavingszorg Oost Brabant.

Derogatis, L. (1983). *SCL-90-R: Administration, Scoring and procedures manual-II.* Towson, MD: Clinical Psychometric Research.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory report. *Psychological Medicine, 13(3),* 595-605.

EuroQolGroup (1995). *EQ-5D user guide.* Rotterdam, NL: Sanders Institute EUR.

Fonagy, P., & Bateman, A. (2006). Progress in the treatment of borderline personality disorder. *British Journal of Psychiatry, 188,* 1-3.

Gabbard, G. O. (2000). Psychotherapy of personality disorders. *The Journal of Psychotherapy Practice and Research, 9(1),* 1-6.

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., … Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of General Psychiatry, 63(6),* 649-658.

Johnson, J. G., First, M. B., Cohen, P., Skodol, A. E., Kasen, S., & Brook, J. S. (2005). Adverse outcomes associated with personality disorder not otherwise specified in a community sample. *American Journal of Psychiatry, 162(10),* 1926-1932.

Karterud, S., Pedersen, G., Bjordal, E., Brabrand, J., Friis, S., Haaseth, O. … Urnes, O. (2003). Day treatment of patients with personality disorders: experiences from a Norwegian treatment research network. *Journal of Personality Disorders, 17*(3), 243-262.

Knekt, P., Lindfors, O., Harkanen, T., Valikoski, M., Virtala, E., Laaksonen, M. A., … Helsinki Psychotherapy Study (2008). Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during 3-year follow-up. *Psychological Medicine, 38(5),* 689-703.

Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. 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). Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff [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). *Multidisciplinary Clinical Guideline for Personality Disorders* [Multidisciplinaire Richtlijn Persoonlijkheidsstoornissen]. Utrecht, The Netherlands: Trimbos-instituut.

Leichsenring, F., & Leibing, 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.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical Analysis with Missing Data.* New York, NY, USA: John Wiley & Sons, Inc.

Livesley, W. J. (2002). *Manual for the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ)*. Port Huron, MI: Sigma Press.

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.

Newman, C. F. (2002). A case illustration of resistance from a cognitive perspective. *Journal of Clinical Psychology, 58(2),* 145-149.

Pagan, J. L., Oltmanns, T. F., Whitmore, M. J., & Turkheimer, E. (2005). Personality disorder not otherwise specified: Searching for an empirically based diagnostic threshold. *Journal of Personality Disorders, 19(6),* 674-689.

Perry, J. C., Banon, E., & Ianni, F. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry, 156(9),* 1312-1321.

Petersen, B., Toft, J., Christensen, N. B., Foldager, L., Munk-Jorgensen, P., Lien, K., … Valbak, K. (2008). Outcome of a psychotherapeutic programme for patients with severe personality disorders. *Nordic Journal of Psychiatry, 62(6),* 450-456.

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.

Savoja, V., Sani, G., Kotzalidis, G. D., De Rossi, P., Stefani, S., Pancheri, L., … Girardi, P. (2011). Bipolar Disorder Presenting as Stalking - A case report. *Psychiatria Danubina, 23(1),* 69-72.

Soeteman, D. I., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. (2008a). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry, 69(2)*, 259-265.

Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2008b). The burden of disease in personality disorders: diagnosis-specific quality of life. *Journal of Personality Disorders, 22(3)*, 259-268.

Soeteman, D. I., Verheul, R., Delimon, J., Meerman, A. M. M. A., van den Eijnden, E., Rossum, B. V., … Kim, J. J. (2010). Cost-effectiveness of psychotherapy for cluster B personality disorders. *British Journal of Psychiatry, 196*(5), 396-403.

Soeteman, D. I., Verheul, R., Meerman, A. M. M. A., Ziegler, U. M., Rossum, B. V., Delimon, 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.

Spreeuwenberg, M. D., Bartak, A., Croon, M. A., Hagenaars, J. A., Busschbach, J. J. V., Andrea, H., … Stijnen, T. (2010). The Multiple Propensity Score as Control for Bias in the Comparison of More Than Two Treatment Arms An Introduction From a Case Study in Mental Health. *Medical Care, 48*(2), 166-174.

van Beek, N., & Verheul, R. (2008). Motivation for treatment in patients with personality disorders. *Journal of Personality Disorders, 22(1),* 89-100.

van der Sanden, K., Smit, W., & Dashorst, M. (2012). *The referencing document of The Dutch National Qualification Framework to the European Qualification Framework*. Brussels, Belgium: European Commission.

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.

van Manen, J. G., Andrea, H., van den Eijnden, E., Meerman, A. M., Thunnissen, M. M., Hamers, E. F., … Verheul, R. (2011). Relationship between patient characteristics and treatment allocation for patients with personality disorders. *Journal of Personality Disorders, 25*(5), 656-667.

Van Manen, J. G., Horn, E. K., Stijnen, T., Timman, R., Busschbach, J. J. V., and Verheul, R. (2015), Tailoring psychotherapy in patients with personality disorders: Matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health, 9*, 133–149.

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J., … Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability, and validity. *Psychological Assessment, 20(1),* 23-34.

Verheul, R., Bartak, A., & Widiger, T. (2007). Prevalence and construct validity of Personality Disorder Not Otherwise Specified (PDNOS). *Journal of Personality Disorders, 21(4),* 359-370.

Verheul, R., & Widiger, T. A. (2004). A meta-analysis of the prevalence and usage of the personality disorder not otherwise specified (PDNOS) diagnosis. *Journal of Personality Disorders, 18(4),* 309-319.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially, "all must have prizes." *Psychological Bulletin, 122*(3), 203-215.

Warren, R. (2012). A cognitive-behavioural perspective on Personality disorders with over-regulation of emotions and poor self-reflectivity: The case of a man with avoidant and not-otherwise specified personality disorder, social phobia and dysthymia treated with Metacognitive Interpersonal Therapy. *Personality and Mental Health, 6(2),* 170-173.

Wilberg, T., Hummelen, B., Pedersen, G., & Karterud, S. (2008). A study of patients with personality disorder not otherwise specified. *Comprehensive Psychiatry, 49(5),* 460-468.

Winston, A., Laikin, M., Pollack, J., Samstag, L. W., McCullough, L., & Muran, J. C. (1994). Short-term psychotherapy of personality disorders. *American Journal of Psychiatry, 151(2)*, 190-194.

Zanarini, M. C. (2009). Psychotherapy of borderline personality disorder. *Acta Psychiatrica Scandinavica, 120(5),* 373-377.

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2010). The 10-year course of psychosocial functioning among patients with borderline personality disorder and axis II comparison subjects. *Acta Psychiatrica Scandinavica, 122(2),* 103-109.

Zimmerman, M., & Coryell, W. (1989). DSM-III personality disorder diagnoses in a nonpatient sample. Demographic correlates and comorbidity. *Archives of General Psychiatry, 46(8),* 682-689.

Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The Prevalence of DSM-IV Personality Disorders in Psychiatric Outpatients. *American Journal of Psychiatry, 162(10),* 1911-1918.

CHAPTER 4

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

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

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.

INTRODUCTION

There is now compelling evidence to support the hypothesis that personality disorders (PDs) 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).Based on the evidence, various systematic literature reviews and clinical guidelines conclude that psychotherapy is 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 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 reported for time-limited psychodynamic psychotherapy and cognitive behavioural therapy (Leichsenring & Leibing, 2003; Matusiewicz, Hopwood, Banducci, & Lejuez, 2010; Town, Abbass, & Hardy, 2011; Verheul & Herbrink, 2007), whereas 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 et al. (2010) and Soeteman et al. (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 et al. (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 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 multicenter project in the Netherlands focusing on the effectiveness and cost-effectiveness of psychotherapeutic treatments in PD patients (Bartak et al., 2010). After informed consent, 38 patients dropped out prematurely. This left 764 patients of which 76 received STIP-TA at De Viersprong, Netherlands Institute for Personality Disorders, and 688 patients receiving other specialized psychotherapies (OP). Five patients of STIP-TA and 100 patients of OP were lost to follow-up. It was not possible to find a good match for four STIP-TA patients, therefore, 67 STIP-TA patients were matched to 67 OP patients 1-to-1 by the logit of the propensity score, which left 134 patients for this trial (see figure 4.1).

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, 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 et al., 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 et al., 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.

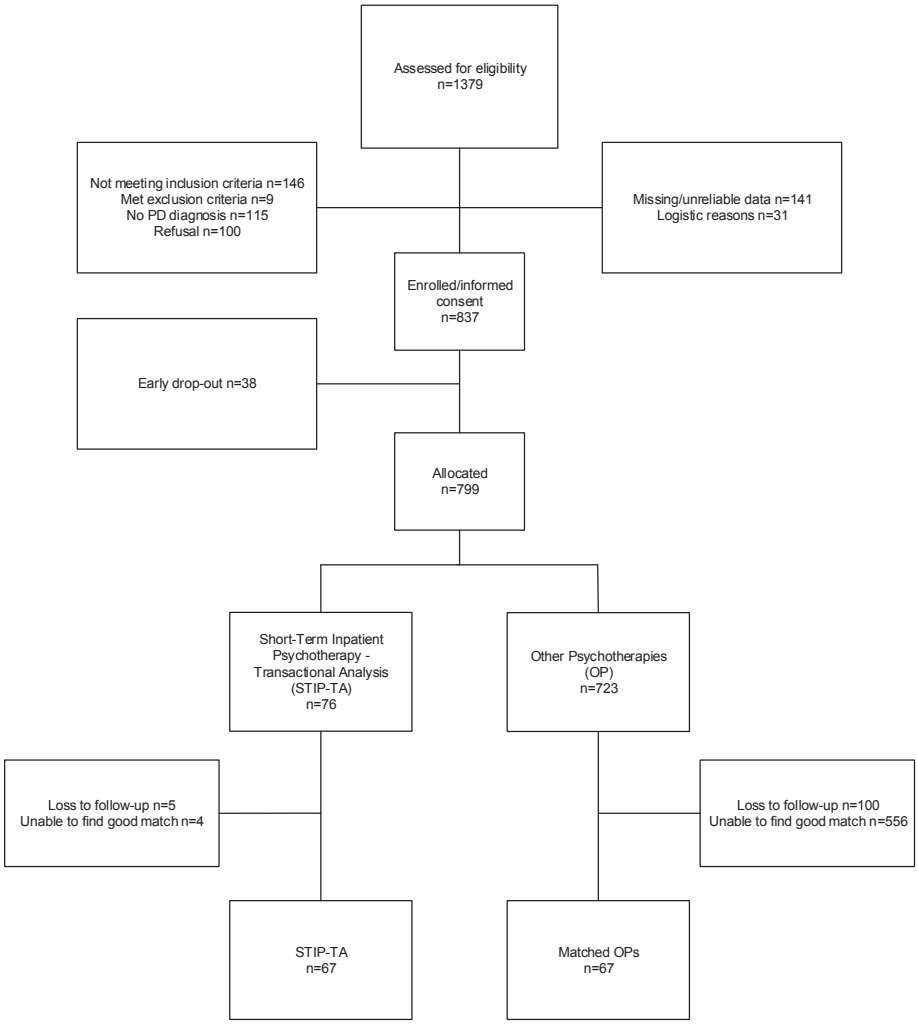


Figure 4.1. Patient flowchart.

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 the smallest distance of the PS logit with the STIP-TA subject and whose 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. STIP-TA subjects for whom no match could be found with a PS logit within the caliper, were also 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.

Outcome analyses are based on the intention-to-treat (ITT) principle. All patients who received a 'minimal exposed dose 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. Differences of $p<0.05$ were considered statistically significant.

Assessments

Baseline measures

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). The PDNOS diagnosis applied to 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. Percentage of agreement between observer-raters ranged from 84% (avoidant PD) to 100% (schizoid) (median 95%). To estimate the intraclass correlation coefficients (ICC(2,1)) for the sum of DSM-IV PD traits present (i.e. scores '2' or '3'), 25 videotaped interviews were rated by three (out of 25) random observers, which resulted in 75 observations. The intraclass correlation coefficients (ICC(2,1)) ranged from 0.31 (schizotypal PD) through 0.88 (depressive PD) (median 0.66).

Self-report

Besides the three outcome measures GSI, OQ-45 and EQ-5D (discussed below), 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 (Livesley, 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 et al., 2008).

Main outcome measures

The primary outcome measure was general psychiatric symptomatology, as measured by the Global Severity Index (GSI) of 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 et al., 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 clinically 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 et al., 2007), respectively, were used. Differences between treatment groups were analyzed using McNemar's test.

Interventions

STIP-TA

Transactional analysis (TA) 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, 1996, 2007). The aim of TA is to think, feel, and act more adaptively by changing old and dysfunctional patterns of behavior. Although TA does not explicitly state to work on mentalizing and metacognition, these concepts are extensively addressed in TA and described in Hawkes (2011).

STIP-TA had incorporated the ideas of transactional analysis in a very short (13 weeks) inpatient psychotherapy. It was specifically designed for and tailored to the needs of patients with various PDs, particularly cluster C PD and PDNOS. The program typically includes patients with childhood traumatic experiences, such as severe illness, disability or death (sometimes by suicide) of a parent, emotional or physical neglect, and sexual or physical abuse. The treatment is guided by an individual treatment contract. In this

contract, patients state in which way they want to change their patterns of thinking, feeling, and behavior. The TA program 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.

Specific for STIP-TA are the following eight characteristics: One shared vision which is carried out by the staff; one shared language; shared responsibility of patients and staff; 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. Non-verbal therapies are included as these are better entrances for many patients to explore their dysfunctional patterns compared to verbal therapies. As therapists (including the psychotherapist, psychosocial nurses, psychiatrist, and non-verbal therapist) work very closely together and share their experiences in working with the patient group two times a day, patients are followed carefully in their treatment process and progress (Delimon, 1999; Thunnissen, 2007).

Other psychotherapies

The specialized psychotherapies consisted of treatments varying widely in terms of setting (i.e. outpatient, day hospital, and inpatient), duration (i.e. ranging from 3-month to 36-month programs), and theoretical orientation (e.g. cognitive-behavioral or psychodynamic). The 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 hospital up to five days a week, including 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-behavioral orientation, 30% had a psychodynamic orientation, 42% had an integrative orientation, and the remaining 5% had an unspecified 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.

RESULTS

Sample

The sociodemographic and clinical characteristics of STIP-TA and OP patients are displayed in table 4.1. As expected, differences between the two groups were not significant (t-tests and χ^2-tests). Mean age of patients in STIP-TA was 39.4 years, in OP 39.3. Thirty-three percent of STIP-TA and 27% of OP were male. Forty-nine percent of STIP-TA and OP patients were living together with a partner. Fifty-eight percent of STIP-TA and fifty-seven percent of OP worked or was studying. Most patients (91% STIP-TA, 88% OP) were diagnosed with either 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.

Figure 4.2 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, t(66)=.26, p=.797).

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, a significant difference (χ^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.

Treatment Outcome

At all time points (12, 24, and 36 months), both groups showed symptomatic improvements and large effect sizes (ES) (Cohen, 1992). Nevertheless, STIP-TA outperformed the OP group at all time points, especially at twelve months after start of treatment (b=.35, p<.001; see table 4.2). 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 4.2) with effect sizes of 1.93 for STIP-TA and 1.39 for OP. Figure 4.3 shows the course of general psychiatric symptomatology over 36 months.

Table 4.1. Sociodemographic and clinical characteristics of STIP-TA and OP patients.

Sociodemographic and clinical characteristics	STIP-TA	OP
	n=67	n=67
	mean (s.d.)	
Age, years	39.4 (9.8)	39.3 (10.2)
	n (%)	
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)
	mean (s.d.)	
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)

[1]European Qualifications Framework

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

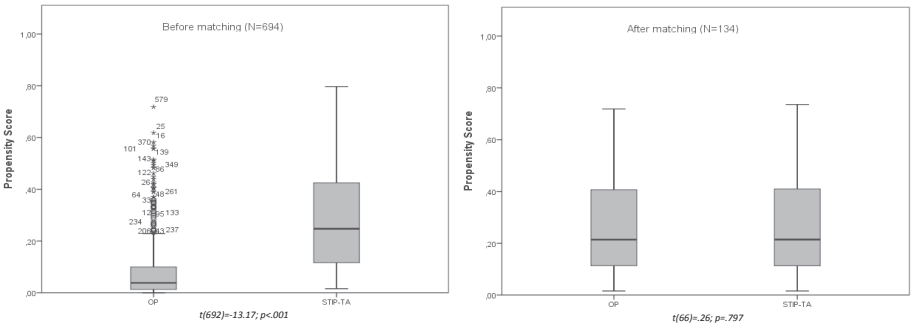


Figure 4.2. Distribution of estimated probabilities of assignment to STIP-TA treatment (propensity scores) before and after matching.

Table 4.2. 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 $p<0.05$.

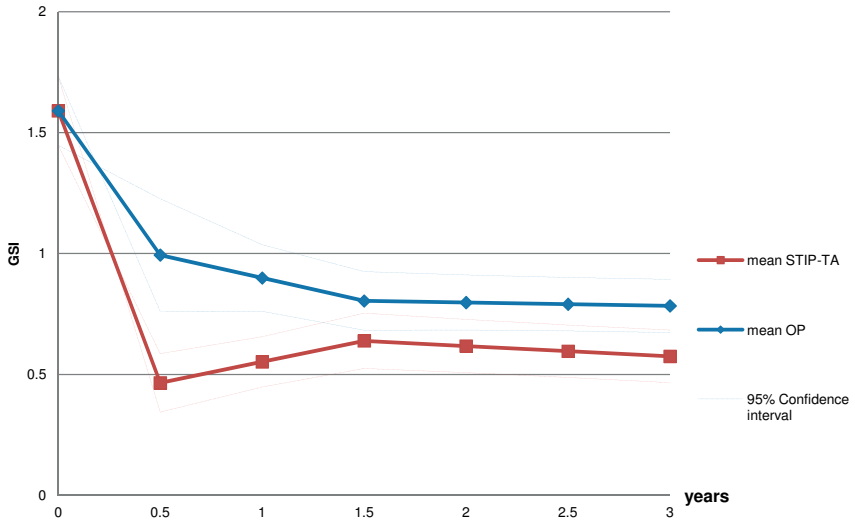


Figure 4.3. Symptom severity over the course of three years.

At 36 months after baseline, 44 pairs of patients (66%) had a 36-month follow-up measurement and were compared in terms of clinically significant change. Ninety percent 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 in terms of psychiatric symptomatology (McNemar's test can be carried out only for a pxp table, where p must be greater than 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 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, analyses were done to compare STIP-TA with different dosages in the OP group, separately. These indicated that part of the superior effect of STIP-TA could be attributed to the lower effectiveness of inpatient treatments in the OP sample.

With respect to the secondary outcomes, i.e. psychosocial functioning (OQ-45) and health-related quality of life (EQ-5D), we found medium to large improvements from baseline to 36 months in both groups as well.

Regarding psychosocial functioning, STIP-TA showed larger ESs than OP on all scales at all time points, except for one (i.e. OQ-45 Interpersonal Relations, 24 months; STIP-TA: M=15.67, sd=4.37, OP: M=15.10, sd=4.68). However, differences between both groups were non-significant. Regarding the OQ-45 Interpersonal relations, at 36 months after baseline, 47 pairs of patients (70%) had a 36-month follow-up measurement. Fifty-one percent 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 4.2). 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. Sensitivity analyses showed that our findings were (relatively) independent of baseline severity and dosage in the control condition. 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), thereby strengthening the conclusion that STIP-TA is a very promising treatment program for patients with cluster C PD or PDNOS.

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 of follow-up. Our assessment of clinically meaningful change at 36 months after baseline showed a small but insignificant advantage in favor of STIP-TA with 68% of STIP-TA patients and 48% of OP patients being symptomatically recovered. Other studies on reliable change or recovery rates of PD patients showed varying results. In Cluster C, 10-54% of patients showed reliable change or recovery (Muran, Safran, Samstag, & Winston, 2005; Svartberg, Stiles, & Seltzer, 2004). 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 recovery rates than other published effect studies of cluster C patients, and were within the range of rates found in studies of cluster B patients. However, it is difficult to interpret these differences as they might be accounted for by a variety of reasons, such as differences in (1) effectiveness of treatments, (2) the follow-up period, (3) definition of response or recovery, and (4) patient characteristics such as the severity of personality pathology.

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: Significant differences in quality of life between STIP-TA and OP patients vanished at the long-term follow-up. 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. 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 et al., 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 "i*ntrinsically 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

Several methodological considerations are worth mentioning. 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 as 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, 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.

Third, information about the treatment fidelity and adherence was not collected. However, it is reasonable to assume that treatments delivered in specialized treatment centers by experienced psychotherapists are of relatively high quality. Perhaps more

importantly, several ingredients of STIP-TA are specifically designed to maintain treatment integrity, e.g. staff training, ongoing supervision, and multidisciplinary meetings twice a day.

Fourth, the interpretation of the results is limited by the variation of treatment modalities in the OP condition. 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). 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 patients are generally very effective and effect sizes of STIP-TA are even larger compared to those typically observed in previous studies (e.g. Arnevik et al., 2010; Leichsenring & Leibing, 2003). Since STIP-TA patients are hospitalized during 13 weeks, it is also a relatively expensive treatment, which 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. 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.

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 et al., 2013; Dimaggio, 2014), 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). Instead of further comparing different treatments, research should be concentrated on the active ingredients of treatments (Clarkin, 2012). Recent studies point to generic effective principles of change in psychotherapeutic treatments of PD. Three factors which

seem to be potentially related to outcome were participant characteristics, therapeutic relationship variables, and technical factors (Castonguay & Beutler, 2006). Elements such as the therapeutic alliance, the ability of the therapist to repair ruptures in the alliance, and the cohesion in group therapy might 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. The abovementioned articles highlight the need for dismantling studies, research on the principles of change and on 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. In these more severe PD patients, this treatment may be too destabilizing, 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. Cluster C PD patients might be able to handle the pressure and intensity of these treatments when provided in a safe and holding environment. As these patients often show rigid patterns of behavior, motivated by anxiety, an intensive, inpatient treatment is pre-eminently suited for these patients. 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. It is further of practical relevance to study whether more seriously disordered patients can 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 ulleval 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* [SCL-90. Handleiding bij een multidimensionele psychopathologie-indicator]. 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., Emmelkamp, P. M. G. (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., Meerding, 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),* 155-159.

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). *Structured Interview for DSM-IV personality disorders (SIDP-IV) (Dutch version)* [Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen (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). Short inpatient psychotherapy [Kortdurende klinische psychotherapie]. 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 and BSI 18. Revised Dutch edition* [Brief Symptom Inventory Herziene Nederlandse Editie]. 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. (2014). Hitting the Bull's Eye in Personality Disorders Psychotherapy. *Journal of Contemporary Psychotherapy, 44(2),* 65-70.

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). Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff [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). *Multidisciplinary Clinical Guideline for Personality Disorders* [Multidisciplinaire Richtlijn Persoonlijkheidsstoornissen]. Utrecht, The Netherlands: Trimbos-instituut.

Leichsenring, F., & Leibing, 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., … 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. M. A., Ziegler, U. M., Rossum, B. V., Delimon, 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). Retrieved from http://repub.eur.nl/pub/8322.

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., … 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.

CHAPTER 5

Cost-Effectiveness of Short-Term Inpatient Psychotherapy based on Transactional Analysis in Patients with Personality Disorders

Eva K. Horn[a, b], Roel Verheul[a, c], Moniek Thunnissen[d, e], Jos Delimon[a], Maartje Goorden[f], Leona Hakkaart-van Roijen[f], Mirjam Soons[g], Anke M. M. A. Meerman[h], Uli M. Ziegler[i], Bert V. Rossum[j], Theo Stijnen[k], Paul M. G. Emmelkamp[c, l], 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 MC, Rotterdam; [c] Department of Clinical Psychology, University of Amsterdam, Amsterdam; [d] GGZ WNB, Bergen op Zoom; [e] Private Practice, Bergen op Zoom; [f] institute for Medical Technology Assessment (iMTA), Erasmus University, Rotterdam; [g] PuntP, Amsterdam; [h] Centre of Psychotherapy Pro Persona, Lunteren; [i] Zaans Medical Centre, Zaandam; [j] Altrecht, Zeist; [k] Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands; [l] The Center for Social and Humanities Research, King Abdulaziz University, Jeddah, Saudi Arabia

SUMMARY

Background

Short-term inpatient psychotherapy with transactional analysis (STIP-TA) in patients with personality disorders (PD) has shown to be more effective than comparable other specialised psychotherapies (OP).

Aims

To assess whether the higher effectiveness of STIP-TA also results in a better cost-effectiveness.

Method

Patients treated with STIP-TA were matched with patients treated with OP by the propensity score. Healthcare costs and lost productivity costs were measured over three years and from the societal perspective. Cost-effectiveness was represented by costs per quality adjusted life years (QALYs). Uncertainty was assessed using bootstrapping.

Results

Mean three year costs were €59,834 for STIP-TA and €69,337 for OP, a difference of €-9,503 (95% CI -32,561 to 15,726). QALYs were 2.29 for STIP-TA and 2.05 for OP, a difference of .24 (95% CI .05 to .44). STIP-TA is a dominant treatment compared to OP: less costly and more effective.

Conclusions

STIP-TA is a cost-effective treatment in PD patients.

Declaration of Interest

EKH, RV, MT, JD, and JJVB have worked or are still working for the 'Viersprong', Netherlands Institute for Personality Disorders. The 'Viersprong' offers STIP-TA. MT is a trainer and supervisor of TA. MT, JD, MS, AMMAM, UMZ, and BVR have been or are still involved with the control conditions.

INTRODUCTION

Personality disordered (PD) patients are a highly prevalent group in the community (Coid, Yang, Tyrer, Roberts, & Ullrich, 2006) with a high individual and societal burden (Soeteman, Verheul, & Busschbach, 2008). PD patients report a quality of life compared to patients with a serious somatic illness, such as rheumatic disease or lung cancer (Soeteman, Verheul, & Busschbach, 2008), and the presence of a PD is associated with high costs for society (Maclean, Xu, French, & Ettner, 2013; Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008). Psychotherapy is considered to be the first choice in treatment based on evidence showing the effectiveness of psychotherapy for PD (Dixon-Gordon, Turner, & Chapman, 2011; Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ, 2008; National Institute for Health and Clinical Excellence, 2009a, 2009b).

State-of-the-art cost-effectiveness analyses (CEAs) on psychotherapeutic treatments in PD populations revealed differences in cost-effectiveness between treatments (Bamelis, 2013; Berghout, Zevalkink, & Hakkaart-van Roijen, 2010; Palmer et al., 2006; Priebe et al., 2012; Soeteman et al., 2010; Soeteman et al., 2011; van Asselt et al., 2008). A large cost-effectiveness study on different modalities in PD patients showed that short-term inpatient psychotherapy (STIP) in general was most cost-effective compared to other treatment modalities in a cluster C PD population (Soeteman et al., 2011). The initial higher costs of STIP were offset by savings in other parts of the health care system (Soeteman et al., 2011). STIP based on Transactional Analysis (STIP-TA) was the shortest treatment of the short-term inpatient modality of this study. As all modalities of STIP were combined, the cost-effectiveness of STIP-TA, specifically, was not determined. However, the evidence of cost-effective treatments in a PD population is still scarce given the large number of different modalities of psychotherapy involved in these studies. Furthermore, more evidence is essential to make deliberate decisions for reimbursed treatments in restricted healthcare budgets (Soeteman & Kim, 2013).

A trial on different types of aftercare following a STIP-TA treatment found a substantial decrease of symptoms at the end of the treatment (Thunnissen et al., 2008). In a matched controlled effectiveness study in which STIP-TA was compared to other specialized psychotherapies (OP) in PD patients, STIP-TA proved to be an effective treatment option over a three year follow-up in terms of general psychiatric symptomatology (Brief Symptom Inventory [BSI]), psychosocial functioning (Outcome Questionnaire-45 [OQ-45]) and quality of life (EuroQol; Horn et al., 2014). As the more effective STIP-TA is also likely to be more costly then less intensive alternatives, a cost-effectiveness study on this comparison is warranted. The current study extends the matched controlled

effectiveness study mentioned above with a cost-effectiveness analysis. This analysis is performed from the societal perspective, with the outcome expressed in costs per quality-adjusted life years (QALY).

METHODS

Study population and design

We compared the costs and effects of Short-Term Inpatient Psychotherapy based on Transactional Analysis (STIP-TA) with Other Psychotherapies (OP) in a matched controlled study. The matching was done on the basis of the propensity score (Bartak et al., 2009). The data used was collected as part of the SCEPTRE study, a large naturalistic study about the cost-effectiveness of psychotherapeutic treatments in PD patients which took place from 2003 to 2011 (Bartak et al., 2010). Six mental health centres participated in the SCEPTRE study which all offered psychotherapeutic treatments for PD patients. They offered a wide range of treatments that varied regarding duration (i.e. 3 to 36 months), setting (i.e. outpatient, day hospital, or inpatient treatment), format (i.e. individual or group meetings), as well as theoretical orientation (e.g. cognitive-behavioural, psychodynamic). Eight hundred thirty-seven patients gave informed consent for the study and were enrolled. Of these, 38 patients dropped out prematurely. Five patients of STIP-TA and 100 patients of OP were lost to follow-up. This left 694 patients of which 71 received STIP-TA at De Viersprong, Netherlands Institute for Personality Disorders (one of the six mental health centres), and 623 patients who received other specialized psychotherapies (OP) at one of the six mental health centres. The PS is defined as the conditional probability of assignment to one of two treatment groups given a set of observed pre-treatment variables (Bartak et al., 2009; Rosenbaum & Rubin, 1983). To estimate the PS, we fitted one logistic regression model with relevant confounders (see Horn et al., 2014) for more details), and medical and productivity costs at baseline as independent variables, and group membership (STIP-TA or OP) as outcome variable (Brookhart et al., 2006). The relevant confounders consisted of sociodemographic variables (e.g. age, sex, level of education) as well as clinical variables (e.g. motivation, baseline scores of the outcome measures, SIPP-118 scales, PD diagnoses). A full list of these variables is provided in the appendix. Nearest neighbourhood matching using a caliper distance of .2 of the standard deviation of the logit of the PS was used (Austin, 2011). The construction and characteristics of this dataset were described in more detail elsewhere (Horn et al., 2014).

In this study, STIP-TA is compared to other psychotherapy (OP). STIP-TA is based on the ideas of transactional analysis (e.g. Barnes, 2007; Berne, 1996) and has integrated

these ideas in a short (13 weeks) inpatient treatment (Delimon, 1999; Horn et al., 2014; Thunnissen, 2007). STIP-TA includes psychotherapy, psychomotor and art therapy, sociotherapy and milieu therapy. This program is designed for patients with PDs, mainly cluster C PD or personality disorder not otherwise specified (PDNOS; see for more details Horn et al., 2014).

The results are based on intention to treat analyses. The study protocol was approved by the Medical Ethics Committee of the Erasmus University Medical Centre in Rotterdam, the Netherlands.

Assessments

Costs

The economic evaluation was undertaken from the societal perspective, the preferred perspective in health economics (Drummond et al., 2005). This perspective describes that all costs related to medical resource utilization, and all costs due to productivity losses should be included. To collect data on costs in other part of health care than the initial intervention and to collect costs related to work (productivity losses), the Trimbos and Institute for Medical Technology Assessment (iMTA) Questionnaire on Costs Associated with Psychiatric Illness (TiC-P) was used (Hakkaart-van Roijen, 2002). A study about the feasibility, reliability, and validity of the TiC-P in patients with a psychiatric disorder showed the questionnaire to be a feasible and a reliable instrument in this patient group (Bouwmans et al., 2013).The TiC-P consists of two sections:

The first section covers the medical costs, which were split up in somatic and psychiatric healthcare costs, medication costs and other costs (e.g. alternative healers, domiciliary care). For mean somatic and psychiatric healthcare costs, the total number of contacts with medical and psychological/psychiatric healthcare providers (e.g. outpatient visits, length of stay in hospital) was asked and multiplied by unit costs of the corresponding health care services (Health Care Insurance Board, 2014; Hakkaart-van Roijen, Tan, & Bouwmans, 2010). The costs of medication were calculated as the cost price per medication multiplied by the daily dose multiplied by the number of prescription days, plus the pharmacist's dispensing costs of €7.50 per prescription. Other costs were valued according to prices reported in the Dutch manual for cost research (Hakkaart-van Roijen et al., 2010).

The second section includes the short form Health and Labour questionnaire (SF-HLQ) for collecting data on productivity losses. Productivity losses could be due to absence from work (absenteeism) or due to reduced efficiency at work (presenteeism; Roijen,

Essink-Bot, Koopmanschap, Bonsel, & Rutten, 1996). To value long-term absence from work, we applied the friction-cost method, which takes into account that a formerly unemployed person may replace a person who becomes disabled (Koopmanschap & Rutten, 1996). The period needed to replace a worker (the so-called friction period) in 2011 was estimated to be 115 days; maximum productivity costs to society were therefore limited to productivity losses during a period of 115 days.

The costs of the primary treatment were estimated using a micro costing approach and the hospital information system. The assessment at the end of treatment did not consist of the TiC-P, as it was expected that patients would confuse questions about possible other treatments in the TiC-P with their primary treatment. This would lead to double counting and therefore overestimated costs. As a result of that decision, the TiC-P at the six months after start of treatment was missing for all patients. To create plausible estimates of the direct costs of other treatments and productivity costs at that measurement point, several methods were used:

Direct medical costs. First, for the time outpatients were in treatment, medical costs were set to the level of the 12 months measure. Second, it was assumed that patients in day hospital and inpatient psychotherapy did not generate any additional medical costs during their psychotherapeutic treatments. Medical costs additional to their primary treatment costs were therefore set to zero. Third, medical costs during the waiting list period were set to the baseline level for all patients, since we did not expect any alterations in medical costs during the waiting list.

Productivity costs. First, patients were not able to work during the hours or days being in treatment. These productivity costs due to absence from work were calculated taking into account the number of days and hours of treatment corrected for numbers of paid employment of the patient per week. These productivity costs were considered to be a good representation of the productivity costs at six months follow-up. Second, costs related to presenteeism were set to the level of the 12 months measure.

Reference unit prices of health care services were applied and adjusted to the year of the study (2011) according to the consumer price index. For medical and productivity costs, we assumed that the recall period of the TiC-P was representative of the six months prior to the assessment. Costs were discounted at an annual rate of 4%, effectiveness as measured by QALYs were discounted at a rate of 1.5%, as recommended by the Dutch guidelines for cost research (Hakkaart-van Roijen et al., 2010). The uncertainty around the mean costs and QALY scores was assessed with bootstrap simulations with 1000 replications for the imputed datasets. The uncertainty interval was represented by the

2.5th and 97.5th percentiles. The results were presented in a cost-effectiveness plane and an acceptability curve (Briggs, Wonderling, & Mooney, 1997; Van Hout, Al, Gordon, & Rutten, 1994).

Effects

For the economic evaluation, the effects were measured in health-related quality of life years (QALYs) using the EQ-5D-3L (EuroQolGroup, 1995). The EQ-5D is a standardised instrument and has shown to be sensitive to change in PD patients (Papaioannou, Brazier, & Parry, 2013). Five health states were measured: 'mobility', 'self-care', 'usual activities', 'pain/discomfort', and 'anxiety/depression'. Possible response levels were 'no problems', 'some or moderate problems' and 'extreme problems or complete inability'. This resulted in 243 different possible health states, which were weighted to obtain a single index score between −0.33 (worst imaginable health state) and 1.00 (best imaginable health state). QALYs were further calculated as the area under the curve for the three year follow-up period. To calculate the mean EQ-5D index values, the Dutch norm scores were used (Brooks, Rabin, & de Charro, 2003; EuroQolGroup, 1995; Lamers, Stalmeier, McDonnell, Krabbe, & van Busschbach, 2005).

Analyses

All outcome measures were assessed at baseline and several follow-up points up to five years after baseline. Due to logistic reasons, three treatment centres 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 centres had their assessments at baseline, at 12, 24, 36, and 60 months after baseline. Since the period between the last two follow-up points was rather long (two years), and the number of missing data increased in time, we limited the analyses to a time horizon of three years after baseline. The recall period for the utilization of medical resources was four weeks prior to completion of the questionnaire. Costs were multiplied by 6.5 to calculate the 6-month costs. The recall period for productivity losses was two weeks prior to completion of the questionnaire. Costs were multiplied by 13 to calculate the 6-month costs.

The expectation-maximization algorithm was used to impute missing quantitative baseline data. For missing categorical data, multinomial logistic regression was used. Since the distribution of the costs was skewed, multiple imputation with ten imputations was used in the case of intermittent missing data (Rubin, 1996). In this multiple imputation, cost data are considered semi-continuous data, which is characterized by a mixture of a considerable proportion of zero values and a skewed distribution of positive values. Using these assumptions, the predictive mean matching (PMM) approach was used to impute the missing values. PMM can account for the semi-continuous nature and

therefore skewed distribution of costs and ensures that imputed values are plausible (Yu, Burton, & Rivero-Arias, 2007). In PMM, the imputed variable takes on the value of one of a set of nearest observed values in the dataset (Horton & Kleinman, 2007). The imputed datasets were further analyzed conforming to the rules established by Rubin (Rubin, 1987; Wayman, 2003).

To compare treatments in terms of costs and effects, the incremental cost-effectiveness ratio (ICER) is determined. The ICER is calculated as the difference in costs of treatments divided by the difference in effectiveness of treatments. With limited health-care resources, thresholds for the ICER's are used to inform decisions on allocation of these resources. In our study, we used a threshold of 40,000 euro/QALY which is diverted from recommendations by the Dutch Council for Public Health and Health Care (Council for Public Health and Health Care, 2006; Soeteman et al., 2011).

The analyses were performed using IBM SPSS 20.0 (IBM Corporation, USA) for data preparation, STATA 12.1 (StataCorp LP, USA) for imputing missing data, and SAS 9.2 (SAS Institute Inc, USA) for bootstrapping.

Sensitivity analyses

To assess the robustness of results under different assumptions, four sensitivity analyses were done. In a first sensitivity analysis, we studied the impact of applying a 3.5% discount rate for both costs and health outcomes as recommended by the UK guideline on health technology assessment of the National Institute for Health and Clinical Excellence (NICE, 2008). A second sensitivity analysis was undertaken using the healthcare perspective: only the medical costs were taken into account. Third, the cost-effectiveness was adjusted for differences in baseline costs. Four different types of regression-based methods were carried out (standard, split, trimmed, and replacement regression) with total costs as dependent variable and costs at baseline and treatment group as independent variables (van Asselt et al., 2009). Fourth, to study the influence of missing data, an analysis was done on the complete follow-up data only.

RESULTS

Sample

The matching procedure resulted in 67 pairs of patients, as 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 (PS), which left 134 patients for this trial (see figure 5.1). The sample consisted mainly of patients with a cluster C PD or a PDNOS diagnosis (see

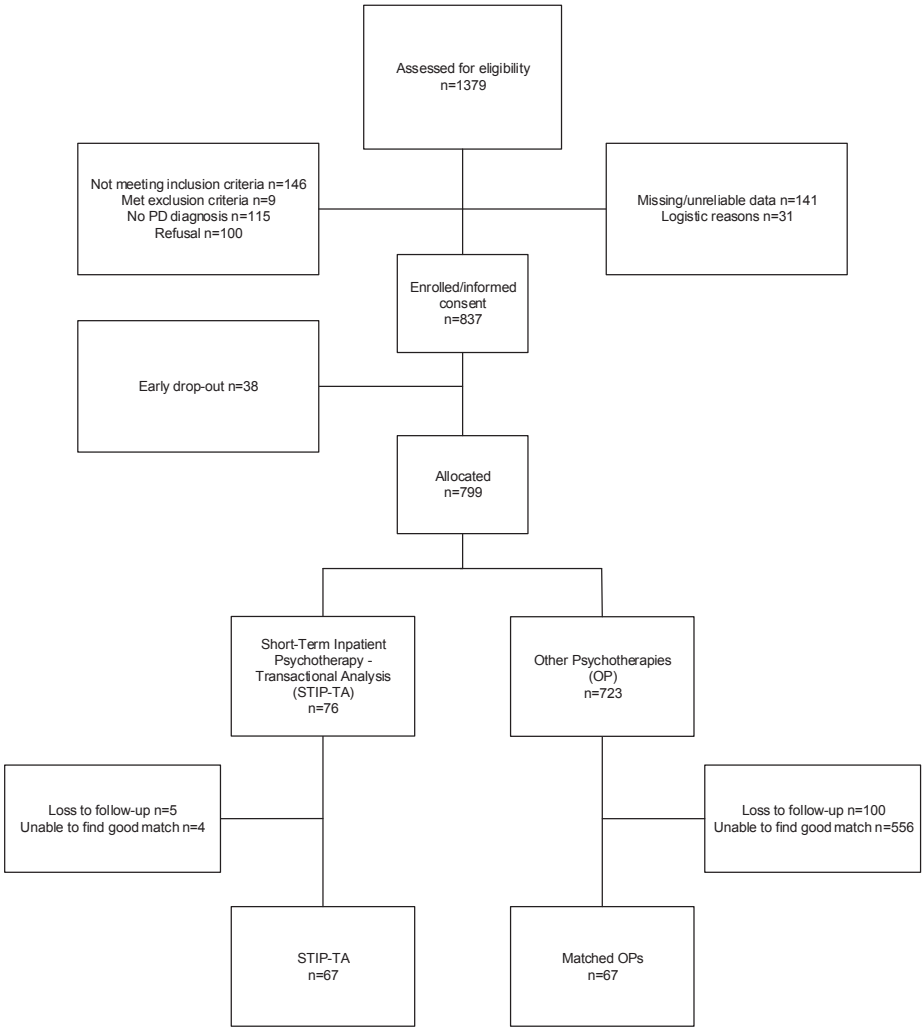


Figure 5.1. Patient flowchart.

Table 5.1). The majority of patients (93% of STIP-TA and 82% of OP) had a history of outpatient treatment and over half (61% of STIP-TA and 51% of OP) had a history of psychotropic drug treatment. Eighty-two percent of STIP-TA and 85% of OP patients reported that they suffered from psychological problems since more than five years. Educational level was high (European Qualifications Framework [EQF] ≥6; van der Sanden, Smit, & Dashorst, 2012) for two-thirds of both groups, and about 60% of both groups worked or was studying. The matching on the basis of the propensity score was successful in

Table 5.1. Characteristics of patients and costs per year before allocation to STIP-TA and OP (costs in Euros).

	STIP-TA N=67	OP N=67
[mean (SD)]		
Age, years	39.4 (9.8)	39.3 (10.2)
N (%)		
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%)
Mode of employment		
Study	3 (5%)	5 (8%)
Paid work	36 (54%)	33 (49%)
Unemployed/other	28 (42%)	29 (43%)
Presence PD[2]		
Cluster A	1 (2%)	2 (3%)
Cluster B	8 (12%)	7 (11%)
Cluster C	33 (49%)	26 (39%)
PD NOS	28 (42%)	33 (49%)

	Mean (SD)	Median	Mean (SD)	Median
Healthcare Costs				
Somatic care	5,620 (24,647)	758	1,962 (8,403)	379
Psychiatric care	3,606 (8,499)	2,303	8,272 (18,897)	2,336
Other care	868 (1,670)	377	1,280 (3,787)	377
Medication	303 (332)	209	331 (326)	195
Total direct costs	10,398 (30,260)	4,471	11,845 (20,427)	5,430
Lost Production Costs				
Absenteeism	9,750 (16,435)	2,869	7,164 (13,163)	0
Presenteeism	3.868 (10,116)	0	1,595 (5,096)	0
Total productivity costs[3]	13,618 (19,234)	6,698	8,759 (13,988)	0
Total costs	24,016 (37,302)	11,747	20,603 (22,518)	14,055
EQ-5D	.54 (.27)	.69	.53 (.26)	.65

STIP-TA, Short-Term Inpatient Psychotherapy - Transactional Analysis; OP, Other Psychotherapies
[1]European Qualifications Framework.
[2]Since it is possible to have more than one diagnosis, the sum of the prevalence is higher than 100%.
[3]One-sample sign test on medians significant between groups.

reducing all differences between background variables as presented in table 5.1, except for the overall costs due to productivity losses (see below).

The matched OP group (see above) consisted of the following short-term (up to six months) and long-term (longer than six months) treatments [mean(SD)]:

- short-term outpatient: N=6 (9%), mean length of treatment 5.4 (1.0) months
- long-term outpatient: N=18 (27%), mean length of treatment 16.9 (5.4) months
- short-term day hospital: N=7 (10%), mean length of treatment 6 (0.0) months
- long-term day hospital: N=11 (16%), mean length of treatment 11.5 (1.0) months
- short-term inpatient: N=9 (13%), mean length of treatment 3.3 (0.9) months
- long-term inpatient: N=16 (24%), mean length of treatment 10.5 (2.3) months.

Day hospital and inpatient programs typically consisted of group psychotherapy as a core element, mostly in combination with one or more nonverbal 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, and 42% consisted of integrative treatments (e.g. cognitive-behavioural/psychodynamic). The orientation of the remaining 5% was not specified (Horn et al., 2014).

Costs and effects

Costs

Total costs at baseline for the past year were €24,016 (sd=€37,302) for STIP-TA and €20,603 (sd=€22,518) for OP, a difference of €3,413. High somatic and psychiatric costs were primarily due to hospitalizations prior to completion of the questionnaires. Costs were not distributed normally (Kolmogorov–Smirnov test on total costs: df=134, p<.000, see table 5.1). The overall medical and total societal costs at baseline were not significantly different between the two conditions (one-sample sign test). The overall costs due to productivity losses, however, differed significantly between both groups (p=.021).

After bootstrap, mean three year total costs were €59,829 (95% CI 45,544 to 82,797) for STIP-TA and €71,007 (95% CI 54,965 to 90,137) for OP, resulting in a non-significant difference in costs of €-9,503 (95% CI -36,719 to 16,597; see table 5.2). There was a trend that STIP-TA patients generated lower costs associated with healthcare and costs associated with productivity losses. Within the healthcare costs, we found a trend that STIP-TA patients generated lower costs in psychiatric and other care but more costs in somatic

Table 5.2. Costs and QALYs for patients in the STIP-TA and OP group over three years follow-up (costs in Euros).

	STIP-TA, €	OP, €	Incremental costs, €	2.5-97.5 percentile
Healthcare costs				
Somatic care, mean(SE)	11,981 (7,978)	7,352 (4,862)	4,715	-12,079 to 24,744
Psychiatric care, mean(SE)	5,579 (1,745)	12,963 (3,510)	-7,317	-14,962 to -634
Intervention, mean(SD)	19,420 (3,048)	25,459 (20,395)	6,040	1,050 to 11,029
Other care, mean(SE)	1,122 (220)	1,586 (368)	-483	-1,309 to 268
Medication, mean(SE)	738 (125)	714 (132)	11	-331 to 338
Subtotal, mean, mean(SE)	38,839 (9,123)	48,075 (4,269)		
Bootstrapped subtotal, mean(SD)	39,087 (9,059)	48,178 (6,882)	-9,091	-29,838 to 15,380
Lost production costs				
Absenteeism, mean(SE)	15,289 (1,880)	18,294 (3,354)	-2,921	-10,434 to 4,417
Presenteeism, mean(SE)	5,405 (1,212)	4,611 (1,599)	834	-3,354 to 4,256
Subtotal, mean, mean(SE)	20,694 (2,542)	22,905 (4,068)		
Bootstrapped subtotal, mean(SD)	20,741 (2,508)	22,828 (3,953)	-2,087	-36,719 to 7,143
Total societal costs, mean(SE)	59,533 (9,789)	70,981 (9,131)		
Bootstrapped total costs, mean(SD)	59,829 (9,733)	71,007 (8,907)	-11,178	-36,719 to 16,597
QALY, mean(SE)	2.30 (.06)	2.03 (.07)		
Bootstrapped QALY, mean(SD)	2.29 (.06)	2.02 (.07)	.27	.08 to .45

STIP-TA, Short-Term Inpatient Psychotherapy - Transactional Analysis; OP, Other Psychotherapies

healthcare compared to OP. High somatic and psychiatric costs were mainly due to hospitalizations. In terms of productivity losses, there was a trend that STIP-TA patients reported being more often absent from work, while OP patients reported less efficiency during work. The STIP-TA treatment was significantly cheaper (mean €19,420, 95% CI €18,676 to €20,163) compared to the OP treatments (mean €25,459, 95% CI €20,458 to €30,434), a difference of 6,040 (95% CI €1,050 to 11,029). Figure 5.2 shows the distribution of cost categories in STIP-TA and OP patients.

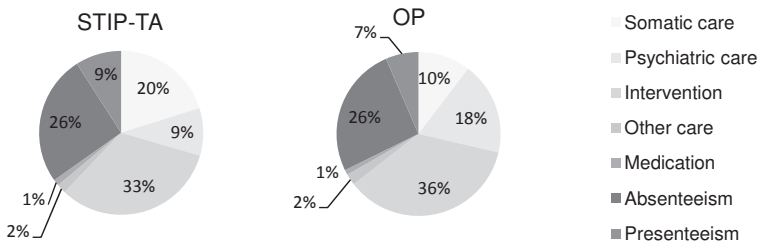


Figure 5.2. Distribution of costs for patients in the STIP-TA and OP group over three years follow-up.

Effects

Mean three year bootstrapped QALYs were 2.29 for STIP-TA and 2.05 for OP meaning a significant difference in QALYs of .24 (95% CI .05 to .43). Figure 5.3 shows the course of the QALYs over 36 months.

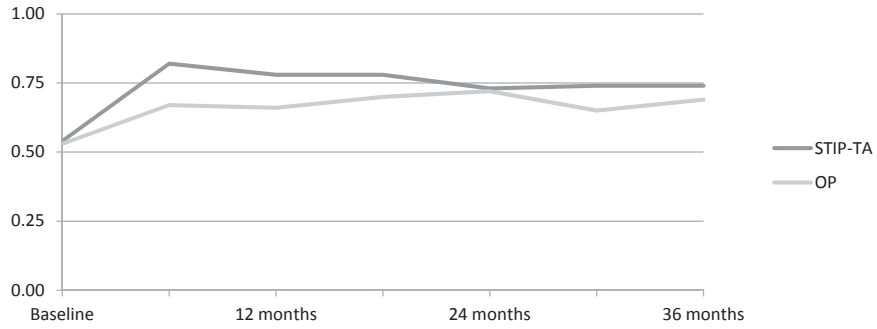


Figure 5.3. Course of the QALY for patients in the STIP-TA and OP group over three years follow-up.

Cost-Effectiveness

Mean three year costs were €59,834 for STIP-TA and €69,337 for OP, a difference of €-9,503 (95% CI -32,561 to 15,726). QALYs were 2.29 for STIP-TA and 2.05 for OP, a difference of .24 (95% CI .05 to .44). The trend towards lower costs and the significant better effects of STIP-TA, resulted in STIP-TA being a '*dominant intervention*' compared to OP: less costly and more effective. To display the impact of uncertainty around the estimated mean costs and QALYs, a cost-effectiveness plane with the bootstrapped results was constructed. Eighty-three percent of the cost-effectiveness pairs lie in the south-east quadrant of the plane, implying lower costs and higher effectiveness, and 17% in the north-east quadrant, implying both higher costs as well as higher effectiveness (see figure 5.4).

The cost-effectiveness acceptability curve (CEAC) indicates that the probability that STIP-TA is cost-effective is between 80% and 90% until a threshold of approximately €20,000/QALY. The probability that STIP-TA is cost-effective increases up to 93% with a ceiling ratio of 40,000/QALY (see figure 5.5).

Sensitivity analyses

Although sensitivity analyses displayed minor differences, the dominance of STIP-TA over OP remained. Costs for STIP-TA ranged from 37,750 (health care perspective) to 61,732 (replacement regression; see table 5.3 for the difference split up per sensitivity

analyses). In OP patients, costs ranged from 44,908 (health care perspective) to 72,298 (replacement regression). QALYs ranged from 1.97 (OP, 3.5% discount rate) to 2.36 (STIP-TA, complete data). ICERs ranged from -23,717 (trimmed regression) to -49,146 (standard regression).

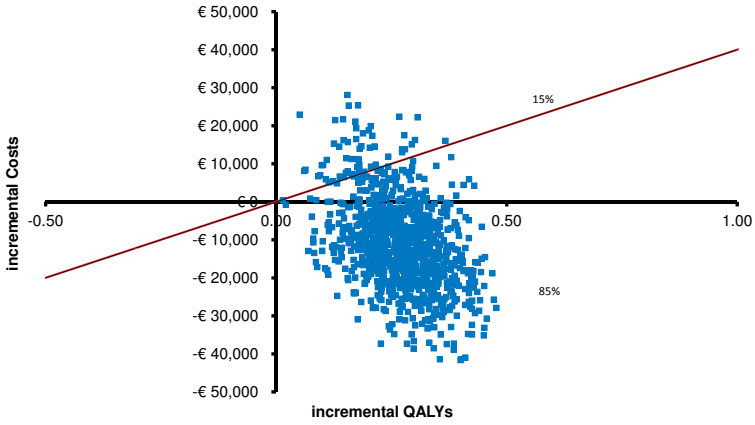


Figure 5.4. Cost-effectiveness plane of STIP-TA versus OP.

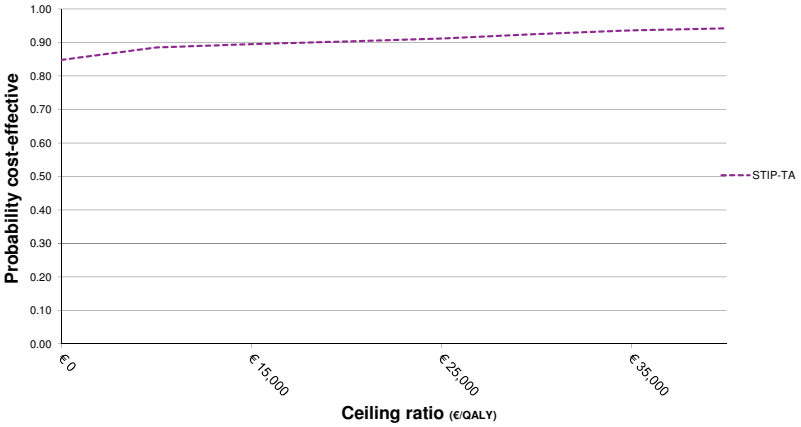


Figure 5.5. Acceptability curve of STIP-TA versus OP.

Table 5.3. Sensitivity analyses on costs and QALYs for patients in the STIP-TA and OP group.

			Costs	QALYs	ICER (Costs per QALY)
3.5% discount rate		STIP-TA	59,785	2.21	
		OP	69,802	1.97	-41,738
Payer perspective		STIP-TA	38,839	2.29	
		OP	48,075	2.05	-38,483
Correction baseline differences	Standard regression	STIP-TA	58,688	2.29	
		OP	71,826	2.05	-54,742
	Split regression	STIP-TA	58,428	2.29	
		OP	71,254	2.05	-53,442
	Trimmed regression	STIP-TA	53,611	2.29	
		OP	63,920	2.05	-42,954
	Replacement regression	STIP-TA	61,723	2.29	
		OP	73,333	2.05	-48,375
Complete data only		STIP-TA	53,623	2.43	
		OP	61,034	2.11	-23,159

DISCUSSION

The current study compared the cost-effectiveness of an intensive short-term inpatient treatment based on Transactional Analysis (STIP-TA) to other specialized psychotherapies (OP) in patients with PD over a three-year time horizon. OP was dominated by STIP-TA: STIP-TA patients generated lower costs and reported a significantly better quality of life, in terms of QALYs gained. With a ceiling ratio of 40,000 Euro/QALY, the probability that STIP-TA is cost-effective compared to OP is 93%. Sensitivity analyses on different discount rates, analyses from the health care perspective, corrections for baseline cost differences, and complete case analysis showed similar results. In both groups, about one third of all costs over 36 months were due to the intervention itself, and about one quarter were due to absenteeism.

Strengths and limitations

A strength of this study is its representativeness and therefore high external validity due to the naturalistic nature and a minimal set of exclusion criteria. The OP group was found to have a great variation in treatments which supports the assumption that OP is representative for psychotherapeutic treatments for many cluster C PD and PDNOS patients in the Netherlands.

Abovementioned also leads to the first limitation: patients were not randomized but were assigned to treatment by clinical knowledge which limits the internal validity. However, a comparison of such different dosages of treatment by means of a RCT is

difficult or even impossible. This is reflected in an earlier attempt to compare STIP-TA to outpatient treatment in cluster C PD patients in a RCT, and which failed due to patient preferences (PSILO trial, National Academic Research and Collaborations Information System). Furthermore, in economic evaluations, controlled but nevertheless nonrandomized designs can even be considered an advantage, since the external validity of these studies is often higher than that of a RCT as these often make use of artificial research settings, strict in- and exclusion criteria, and mostly a limited time horizon (Soeteman et al., 2011). Economic evaluations are supposed to be done on estimates of the actual costs and effects in practice (Wells, 1999). To use data collected in RCT's for health economic studies, the wider inclusion criteria and the longer time horizon relevant for practice has to be modelled which leads to additional uncertainty. In the present study, data is collected in regular clinical practice. Thus, analyses are already based on the wide inclusion criteria and long-term follow-up, and the introduction of this uncertainty is avoided. At the same time, the generalizability to a PD population at large is limited: to the OP condition is not representative for all patients who have the same pathology as the patients in the STIP-TA group, neither is the OP group representative for all specialized psychotherapies in general. We can therefore not conclude in absolute terms that STIP-TA is the most cost-effective option for all cluster C PD and PDNOS patients. The external validity of this study is limited to the included alternative treatments. Moreover, given that data on treatments were collected from 2003 to 2007, due to changes in psychiatry and reimbursements one could question whether the OP treatments studied are still similar to contemporary PD treatments. The trend towards less (long-term) inpatient treatments, inpatient treatments which are less often open-ended, and treatment durations which are limited in general are, among others, important changes.

We used the TiC-P to estimate costs which had a recall period of two to four weeks and extrapolated these to estimate the costs of patients for six months. The assumption that these weeks are representative can be questioned, however we assumed that the over– as well as the under-estimation of costs will balance each other out on average. The long-term follow-up of three years and varying measurement points led to a high amount of missing data. This made assumptions about the missing healthcare costs and productivity losses necessary and led to uncertainty about the results. However, we handled missing data with sophisticated analyses (multiple imputation) and carried out a secondary analysis on the complete data to confirm the robustness of results.

Implications

The current study compared STIP-TA to a variety of other psychotherapies (OP), heterogeneous in method, duration and dosage and adds evidence to the knowledge of

cost-effective treatment options in PD patients. (Dixon-Gordon et al., 2011; Soeteman & Kim, 2013). Soeteman et al. have studied the value of further research on cost-effective psychotherapies in cluster B and C PD patients in the Netherlands and concluded that the societal value of additional research is considerable (Soeteman et al., 2011).

Currently, mental health care for PD patients mainly consists of long-term outpatient psychotherapy. Inpatient treatments are generally not the first choice of treatment but are mainly used for crisis intervention and stabilization, especially in BPD patients. A hospitalization of three months, although a relatively short period of time, can be an obstacle for many patients. In the current economic climate, less radical and (on the short-term) cheaper outpatient treatments are preferred. However, although STIP-TA is viewed as an expensive treatment, these treatment costs were generally lower than the costs for OP. This was due to two-thirds of OP patients following (rather expensive) long-term treatments of more than a year on average. The current study further points to savings elsewhere in society and even when considering only health care costs, our estimates are that STIP-TA remains less costly compared to OP. A study on whether the dominance of STIP-TA holds when compared to evidence-based outpatient treatments gives relevant information for decision making and clinical practice and adds evidence to make well-informed decisions about the preservation or dispensation of inpatient treatments for this patient group.

Moreover, more knowledge about the efficacious ingredients of this specific treatment is needed. With the knowledge of these factors, it might be possible to add some of these factors to other, less expensive and less intense outpatient treatments. This would be in line with the deinstitutionalisation of mental health care and the emphasis on patient empowerment. The question remains whether it is possible to keep the high pressure used in the inpatient treatment which seems to be one of the efficacious ingredients of treatments (van Manen, Horn, Stijnen, van Busschbach, Verheul, 2015).

An extension of the current study with a subgroup analysis would furthermore provide more evidence to give a strong and evidence-based advice. Since STIP-TA is known as a high pressure and sometimes destabilizing treatment, mainly rather stable cluster C PD and PDNOS patients are admitted to this treatment. A study of van Manen et al. (2015) showed that destabilizing treatments (which use techniques as interpretation and confrontation) in general seemed more effective than stabilizing treatments (which use techniques as empathic validation and advice) in psychological strong as well as in psychological rather weak patients. These results point to the importance of a study on the expansion of STIP-TA to (relatively mild) cluster A and B PD patients.

Conclusion

The present study showed that STIP-TA is a cost-effective treatment when compared to other specialized psychotherapies (OP). As studies like the present one showed that investments in expensive treatments can be earned back on the long-term, these cost-effective treatments should be continued in spite of the high initial expenses. These treatments should be preserved as an alternative treatment for outpatient psychotherapies and should be adopted in clinical guidelines. The present study invites clinicians, patients and payers to reconsider the direct medical costs of treatments in the light of favourable effects and future savings.

REFERENCES

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.

Bamelis, L. L. (2013). *Considering care: a clinical and economic evaluation of schema therapy for personality disorders* (Doctoral dissertation). Retrieved from http://www.narcis.nl/publication/RecordID/oai:publications.maastrichtuniversity.nl:27-M276040.

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., Spreeuwenberg, M. D., Andrea, H., Busschbach, J. J., Croon, M. A., Verheul, R. … 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., … Emmelkamp, P.M.G. (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.

Berghout, C. C., Zevalkink, J., & Hakkaart-van Roijen, L. (2010). A Cost-utility Analysis of Psychoanalysis versus Psychoanalytic Psychotherapy. *International Journal of Technology Assessment in Health Care, 26(1),* 3-10.

Berne, E. (1996). Principles of transactional analysis. *Indian Journal of Psychiatry, 38(3),* 154-159.

Bouwmans, C., De Jong, K., Timman, R., Zijlstra-Vlasveld, M., Van der Feltz-Cornelis, C., Tan Swan, S., Hakkaart-van Roijen, L. (2013). Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). *BMC Health Services Research, 13,* 217.

Briggs, A. H., Wonderling, D. E., & Mooney, C. Z. (1997). Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation. *Health Economics, 6(4),* 327-340.

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.

Brooks, R., Rabin, R., & de Charro, F. (2003). *The Measurement and Valuation of Health Status using EQ-5D: A European Perspective*. Dordrecht, NL: Kluwer Academic Publishers.

Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. British *Journal of Psychiatry, 188*, 423-431.

Consumer Price Index. Central Bureau for Statistics [in Dutch]. Website. http://statline.cbs.nl.

Council for Public Health and Health Care (RVZ*). Sensible and Sustainable Care* [in Dutch]. The Hague, The Netherlands. 2006. http://www.rvz.net/en/view/sensible-and-sustainable-care.

Delimon, J. (1999). Short inpatient psychotherapy [Kortdurende klinische psychotherapie]. 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.

Dixon-Gordon, K. L., Turner, B. J., & Chapman, A. L. (2011). Psychotherapy for personality disorders. *International Review of Psychiatry, 23(3),* 282-302.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B., & Stoddart, G. L. (2005). *Methods for the Economic Evaluation of Health Care Programmes, third Edition.* Oxford, UK: Oxford University Press.

EuroQolGroup. (1995). *EQ-5D user guide.* Rotterdam, the Netherlands: Sanders Institute EUR.

Hakkaart-van Roijen, L. (2002). *Manual Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness (TiC-P)* [Handleiding Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness (TiC-P)]. Rotterdam, the Netherlands: Institute for Medical Technology Assessment, Erasmus Universiteit Rotterdam.

Hakkaart-van Roijen, L., Tan, S. S., & Bouwmans, C. A. M. (2010). *Manual for Cost Research: Methods and Unit-prices for Economic Evaluations in Health Care* [Handleiding voor kostenonderzoek]. Diemen, The Netherlands: The Health Care Insurance Board (CVZ).

Health Care Insurance Board (CVZ). Pharmacotherapeutic Compass [in Dutch]. Website. http://www.fk.cvz.nl

Horn, E. K., Verheul, R., Thunnissen, M., Delimon, J., Soons, M., Meerman, A. M., … Busschbach, J.J.V. (2014). Effectiveness of Short-Term In-patient Psychotherapy using Transactional Analysis in Patients with Personality Disorders: A Matched Control Study using Propensity Score. Journal of Personality Disorders.

Horton, N. J., & Kleinman, K. P. (2007). Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician, 61(1),* 79-90.

Koopmanschap, M., & Rutten, F. (1996). A practical guide for calculating indirect costs of disease. *Pharmacoeconomics, 10,* 460-466.

Lamers, L. M., Stalmeier, P. F., McDonnell, J., Krabbe, P. F., & van Busschbach, J. J. (2005). Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff [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). *Multidisciplinary Clinical Guideline for Personality Disorders* [Multidisciplinaire Richtlijn Persoonlijkheidsstoornissen]. Utrecht, The Netherlands: Trimbos-instituut.

Maclean, J. C., Xu, H., French, M. T., & Ettner, S. L. (2014). Mental Health and High-Cost Health Care Utilization: New Evidence from Axis II Disorders. *Health Services Research, 49,* 683-704.

National Institute for Health and Clinical Excellence (2008). *Guide to the methods of technology appraisal*. London: National Institute for Health and Clinical Excellence.

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

Palmer, S., Davidson, K., Tyrer, P., Gumley, A., Tata, P., Norrie, J., … Seivewright, H. (2006). The cost-effectiveness of cognitive behavior therapy for borderline personality disorder: results from the BOSCOT trial. *Journal of Personality Disorders, 20(5),* 466-481.

Papaioannou, D., Brazier, J., & Parry, G. (2013). How to measure quality of life for cost-effectiveness analyses of personality disorders: a systematic review. *Journal of Personality Disorders, 27(3),* 383-401.

Priebe, S., Bhatti, N., Barnicot, K., Bremner, S., Gaglia, A., Katsakou, C., … Zinkler, M. (2012). Effectiveness and cost-effectiveness of dialectical behaviour therapy for self-harming patients with personality disorder: a pragmatic randomised controlled trial. *Psychotherapy and Psychosomatics, 81(6),* 356-365.

National Academic Research and Collaborations Information System. *PSILO trial. A randomized comparison of the efficacy and cost-effectiveness of Psychotherapy Short-term Inpatient versus Long-term Outpatient.* Retrieved 06-01-2014, 2014, from http://www.narcis.nl/research/RecordID/OND1324951

Roijen, L., Essink-Bot, M., Koopmanschap, M., Bonsel, G., & Rutten, F. (1996). Labor and health status in economic evaluation of health care: the Health and Labor Questionnaire. *International Journal of Technology Assessment in Health Care, 12,* 405-415.

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.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91(434),* 473-489.

Soeteman, D. I., Busschbach, J. J., Verheul, R., Hoomans, T., & Kim, J. J. (2011). Cost-effective psychotherapy for personality disorders in the Netherlands: the value of further research and active implementation. *Value in Health, 14(2),* 229-239.

Soeteman, D. I., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. (2008). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry, 69(2),* 259-265.

Soeteman, D. I., & Kim, J. J. (2013). Cost-effectiveness of psychotherapy for personality disorders: treatment recommendations and implementation. *Expert review of pharmacoeconomics & outcomes research, 13(1),* 73-81.

Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2008). The burden of disease in personality disorders: diagnosis-specific quality of life. *Journal of Personality Disorders, 22(3),* 259-268.

Soeteman, D. I., Verheul, R., Delimon, J., Meerman, A. M. M. A., van den Eijnden, E., Rossum, B. V., … Kim, J.J. (2010). Cost-effectiveness of psychotherapy for cluster B personality disorders. *British Journal of Psychiatry, 196(5),* 396-403.

Soeteman, D. I., Verheul, R., Meerman, A. M. M. A., Ziegler, U. M., Rossum, B. V., Delimon, 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.

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). Retrieved from http://repub.eur.nl/pub/8322.

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.

van Asselt, A. D., van Mastrigt, G. A., Dirksen, C. D., Arntz, A., Severens, J. L., & Kessels, A. G. (2009). How to deal with cost differences at baseline. *Pharmacoeconomics, 27(6),* 519-528.

van Asselt, A. D. I., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., van Dyck, R., Spinhoven, P., … Severens J.L. (2008). Out-patient psychotherapy for borderline personality disorder: cost-effectiveness of schema-focused therapy v. transference-focused psychotherapy. *British Journal of Psychiatry, 192(6),* 450-457.

van der Sanden, K., Smit, W., & Dashorst, M. (2012). *The referencing document of The Dutch National Qualification Framework to the European Qualification Framework*. Brussels, Belgium: European Commission.

van Hout, B. A., Al, M. J., Gordon, G. S., & Rutten, F. F. H. (1994). Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics, 3(5),* 309-319.

van Manen, J. G., Andrea, H., van den Eijnden, E., Meerman, A. M., Thunnissen, M. M., Hamers, E. F., … Verheul, R. (2011). Relationship between patient characteristics and treatment allocation for patients with personality disorders. *Journal of Personality Disorders, 25(5),* 656-667.

Van Manen, J. G., Horn, E. K., Stijnen, T., Timman, R., Busschbach, J. J. V., and Verheul, R. (2015), Tailoring psychotherapy in patients with personality disorders: Matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health, 9*, 133–149.

Wayman, J. C. (2003). *Multiple Imputation For Missing Data: What is it and How Can I Use it?* Paper presented at the 2003 Annual Meeting of the American Educational Research Association, Chicago, IL.

Wells, K.B. (1999). Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. *American Journal of Psychiatry, 156(1),* 5-10.

Yu, L. M., Burton, A., & Rivero-Arias, O. (2007). Evaluation of software for multiple imputation of semi-continuous data. *Statistical Methods in Medical Research, 16(3),* 243-258.

CHAPTER 6

Tailoring Psychotherapy in patients with Personality Disorders: Matching the level of psychological strengths to the level of stabilizing versus destabilizing Psychotherapy

Janine G. Van Manen[1,2], Eva K. Horn[1,2], Theo Stijnen[3], Reinier Timman[1,2], Jan J. V. Busschbach[1,2] And Roel Verheul[1,4]

[1] Viersprong Institute for Studies on Personality Disorders, Halsteren, The Netherlands; [2] Departement of Psychiatry, section Medical Psychology and Psychotherapy, Erasmus Medical Center, Rotterdam, The Netherlands; [3] Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands; [4] Programme group Clinical Psychology, Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

Background

Clinical evidence suggests that patients high on psychological strengths profit more from destabilizing psychotherapy, whereas patients low on strengths profit more from stabilizing psychotherapy. This matching hypothesis was tested.

Methods

This quasi-experimental study was conducted between 2003 and 2008 in 735 patients with personality disorders from 6 psychotherapy centers in the Netherlands. Patients were assigned to different levels of stabilizing and destabilizing psychotherapies. Levels of psychological strengths were measured. We used multilevel modeling to estimate outcome at 12 months after baseline. The propensity score controlled for initial differences at baseline.

Results

The findings show that destabilizing psychotherapies have slightly better outcomes than stabilizing psychotherapies. Patients high on psychological strengths improve slightly more than patients low on psychological strengths. The observed interaction effect contradicted our hypothesis.

Conclusion

The results imply that destabilizing psychotherapies can be considered as first treatment option for patients both high and low on psychological strengths.

INTRODUCTION

Personality disorders (PDs) are highly prevalent mental disorders with high individual, societal and economic burden of disease (Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008; Soeteman, Verheul, & Busschbach, 2008). Although PDs are relatively enduring conditions, amenability to psychological treatments has been established and documented (APA, 2001; Binks et al., 2006; Leichsenring & Leibing, 2003; Perry, Banon, & Ianni, 1999). Importantly, the efficacy of psychotherapy for PD is not primarily determined by the specific theoretical orientation, but rather by the consistent application of a coherent and – both to patient and to therapist – comprehensible therapeutic method (Verheul & Herbrink, 2007). In addition, efficacious treatments are typically characterized by a high level of structure, effort to enhance compliance, a clear focus, a long-term and powerful attachment relationship, an active stance, and integration with other services (Bateman & Fonagy, 2000).

An element that has received less attention but is nevertheless likely to be essential, is the optimal level of destabilizing in treatment. Patients with PD are typically characterized by persistent and pervasive patterns of cognition, emotion and behavior. From a dynamic systems theory perspective, it can be predicted that such patterns or '*attractor states*' need to be destabilized first. Then more functional patterns can be organized (Hayes & Strauss, 1998; Thelen & Smith, 1994). This prediction is in line with the principles of psychodynamic psychotherapy promoting the application of various interpretive or expressive techniques (Gabbard, 2005). Such techniques are focused on uncovering unconscious wishes, fears, conflicts and defenses, as opposed to supportive techniques that help the patients to adapt to stresses while avoiding insights. The broad spectrum of psychotherapeutic techniques can be placed on an expressive-supportive continuum, running from typically expressive or destabilizing categories such as interpretation and confrontation to typically supportive or stabilizing categories such as empathic validation, advice and praise, and affirmation (Horwitz et al., 1996). Psychodynamic psychotherapy explicitly encourages to "*be as expressive as you can be, and as supportive as you have to be*" (Wallerstein, 1986, p.688). In this study we defined three levels of destabilization. The focus in the '*stabilizing treatments*' is on acceptance and help patients to cope with his PD problems. Therapists typically work with supportive and structuring interventions, which results in relatively low stress levels during treatment. The focus in the '*destabilizing treatments*' is on change and help the patient to replace their dysfunctional patterns by adaptive ones. Therapists typically work with confrontative, expressive, insight-oriented interventions, which results in relatively high stress levels during treatment. In the intermediate variant therapists focus simultaneously on acceptance and change, and use both stabilizing and destabilizing interventions, resulting in changing stress levels in the patient.

To the best of our knowledge we are not aware of any empirical study focusing directly on the importance of stabilizing versus destabilizing in the treatment of PD. However, various studies provide pieces of evidence that are consistent with the psychodynamic literature which suggests that patients scoring high on psychological strengths or ego-adaptive capacities (e.g. capacity to relate, identity integration and the ability to mentalize) are better able to tolerate and profit from destabilizing techniques than patients scoring low on such psychological strengths. This '*matching hypothesis'* is for instance supported by various studies that have shown that patients with severe PD drop out prematurely from expressive psychotherapies more often than from supportive psychotherapies (Piper, Joyce, McCallum, & Azim, 1998; Piper, McCallum, Joyce, Azim, & Ogrodniczuk, 1999). Secondly, the studies of Bartak et al. (2011, 2010) have shown superiority of short-term inpatient psychotherapy in patients with cluster C but not with cluster B PD. Short-term inpatient treatments are characterized by a high level of therapeutic intensity and pressure. The authors suggest that "*patients with cluster C personality pathology might be able to handle the high pressure of this treatment modality better than (pure) cluster B PD patients, who probably have a lower tolerance for therapeutic pressure*" (Bartak et al., 2010, p. 28). Third and finally, the matching hypothesis is consistent with Gabbard et al. (2000) suggestion of patient characteristics that can help clinicians decide whether a predominantly expressive versus a predominantly supportive treatment focus is indicated. According to Gabbard, indications for a highly expressive modality are, for instance: a strong motivation, suffering, tolerance of frustration, psychological mindedness, and intact reality testing, whereas indications for a highly supportive modality are, for instance: low anxiety tolerance, poor frustration tolerance, poor impulse control, and little capacity for self-observation. Some research in a non-PD population supports the suggestion of Gabbard of a matching relation, i.e. matching between level of personality organization (Koelen et al., 2012) or different personality types (anaclitic/introjective) (Blatt, Zuroff, Hawley, & Auerbach, 2010) and type of intervention (interpretive versus supportive) (Piper et al., 1998; Piper et al., 1999).

The present study aims to explore the matching hypothesis outlined above in a large quasi-experimental, naturalistic study. In this population we study whether patients high on strengths profit more from predominantly destabilizing treatments, whereas patients low on strengths might profit more from predominantly stabilizing treatments. Research questions are focused on (1) the impact of psychological strengths on treatment outcome, (2) the impact of level of destabilization on treatment outcome, and (3) the interaction between the patient's psychological strength and the treatment's level of destabilization with respect to outcome.

METHOD

Participants

Participants (n=735) were recruited from a consecutive series of admissions to six mental health care centers in the Netherlands (i.e. de Viersprong, Netherlands Institute for Personality Disorders, Halsteren; Altrecht, Utrecht; Zaans Medical Centre, Zaandam; Pro Persona, Centre of Psychotherapy, Lunteren; GGZWNB, Halsteren; Arkin, Amsterdam). These centers offer specialist psychotherapy for adult patients with PDs. From March 2003 to March 2006, a total of 1,379 admissions completed the intake and screening procedure and were selected for treatment (Figure 6.1). The intake and screening procedure included self-report questionnaires and a semi-structured interview for diagnosing

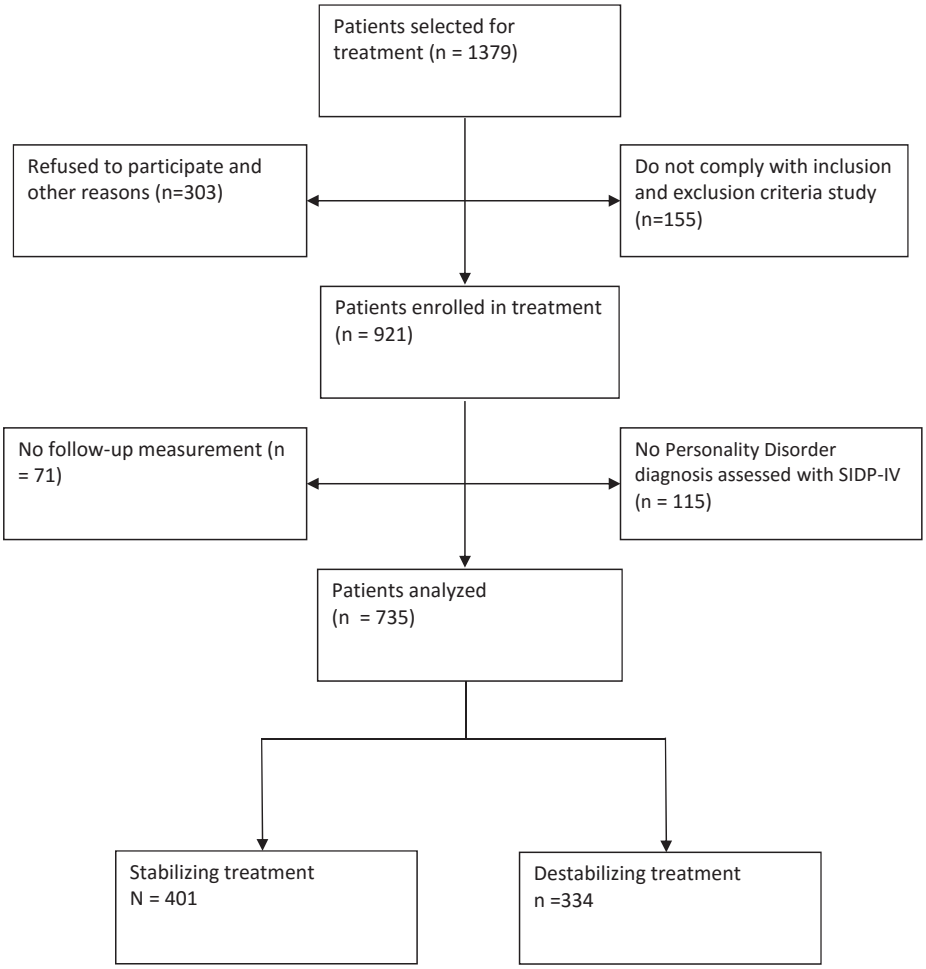


Figure 6.1. Patient flow.

PDs. The data obtained from this initial assessment served as baseline data for our study. As it was part of the standard screening procedure, and not involved additional risks or load, informed consent for the baseline data collection was not mandatory under Dutch law. The study was approved by the medical ethics committee of the Erasmus MC.

Of the 1,379 admissions, 146 were excluded from the study because of one of the following inclusion criteria: age between 18 and 70 years (n=13), personality pathology is primary psychiatric disorder (not eating disorder for example) (n=34), and referral for psychotherapeutic treatment aimed at personality problems (n=99). Nine patients 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 patients, of whom 100 refused to participate (i.e., did not provide informed consent) and 38 patients did not enter treatment (i.e., received less than two treatment sessions or less than two days of inpatient or day hospital psychotherapy). Another 31 patients could not participate due to logistic reasons (i.e., no appointment could be made to provide informed consent), and 134 patients were excluded due to missing or unreliable self-report questionnaires or semi-structured interview (mostly because of lack of interviewers at the start of the study, n=106).

The remaining 921 patients were informed about the study and its procedures, provided written informed consent for follow up data, and entered the study. Of those, 186 were post hoc excluded because they could either not be diagnosed with a PD (n=115) or the follow-up data were not available (n=71). There was no difference in psychiatric symptoms (BSI), their social role and relational functioning (OQ-45), their level of personality pathology (SIPP-118) and the socio-demographic variables age and sex at baseline between patients with follow-up data and those without. The final sample consisted of 735 patients who were included in this study.

Treatments and level of destabilization

Patients were assigned to the different psychotherapeutic treatments available in the six treatment centers in the local standard way, i.e. based on the available test results, expert opinion and clinical experience (for more information about the treatment selection: Van Manen et al., 2008; Van Manen et al., 2011; Van Manen et al., 2012). Treatments were delivered by licensed psychiatrists of psychologists. They had an average of 15 years of postgraduate clinical experience (SD =10.1).

The available treatments differ in terms of setting (i.e. outpatient, day-hospital and inpatient), duration (i.e., varying from three to 24 months), theoretical orientation (predominantly cognitive-behavioral and psychodynamic orientations) and level of destabilization. The latter characteristic is focused on in this study. The level of destabilization of all individual treatment programs in the six treatment centers were scored on a 3-point Likert scale (i.e., low, intermediate, and high level) at two times during the investigation. In 2002 (before the inclusion started), the intake clinicians of each center provided a consensus rating for each treatment program. As we were interested in the reliability and validity of this measurement, we repeated the scoring procedure in 2007 (after the inclusion was completed), but this time we asked the managers in the steering committee of the investigation, to independently provide scores. Both times we instructed the respondents to score the level of destabilization independent from the setting and the duration of the treatment. The three levels were described as follows:

(1) Low level of destabilization: Predominantly stabilizing psychotherapies focus on acceptance and help patients to cope with his PD problems. Therapists typically work with supportive and structuring interventions. Examples of therapeutic techniques are: giving advice, psycho-education and empathic validation. As a result the tension or stress in the patient is kept as low as possible.

(2) Intermediate level of destabilization: These psychotherapies focus simultaneously on acceptation of the PD problems as well as on helping patients to replace their dysfunctional patterns by adaptive ones. Therapists work both with confrontative, expressive, insight oriented interventions and with supportive and structuring interventions. Because of the flexibility in using both techniques, a therapist tailors his interventions to the tension and stress level of the patient, or by the psychic state of the patient at the specific moment in treatment.

(3) High level of destabilization: Predominantly destabilizing psychotherapies focus on change and help the patient to replace their dysfunctional patterns by adaptive ones. Therapists typically work with confrontative, expressive, insight-oriented interventions aiming at uncovering unconscious wishes, fears, conflicts and defenses. Examples of therapeutic techniques are: interpretation, confrontation and clarification. As a result the tension and stress level in a patient can increase to a high level.

The two measurements in 2002 and 2007 were highly correlated ($r=.69$, $p<.001$), supporting the reliability and construct validity of our operationalization of level of destabilization. In this study we used the level of destabilization scores by the managers in the steering committee of the investigation. Because only 36 out of the 735 patients

had a treatment with a low level of destabilization, we combined the low and intermediate level into a group with low level of destabilization (referred to as '*stabilizing psychotherapy')* and a group with high level of destabilization (referred to as '*destabilizing* psychotherapy').

Assessments

PD diagnosis

DSM-IV-TR PD diagnoses were measured using the Dutch version of the Structured Interview for DSM-IV Personality Disorders (SIDP-IV) (Jong, de Derks, Oel, & van Rinne, 1995; Pfohl, Blum, & Zimmerman, 1997). This interview covers the 11 formal DSM-IV-TR axis II diagnoses including PD not otherwise specified (PDNOS), two appendix diagnoses (i.e. depressive and negativistic PD), and self-defeating PD. Interviewers were Master 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. Inter-scorer reliability was evaluated in a convenience sample of 25 videotaped interviews, which were rated by three observer raters resulting in 75 observations. 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).

Strength measures

As there is no golden standard for measuring psychological strengths or ego adaptive capacities, we considered this variable a '*latent construct'* and used four operationalizations: severity of PD, adaptive personality functioning, overall defensive functioning, and motivation for treatment. These variables fit into the internal strength domain as revealed by a recent concept map study of patient characteristics relevant for treatment assignment (Van Manen et al., 2012). First, severity of PD was measured with the SIDPIV (describing of the administration is given above). To form mutually exclusive diagnostic groups, we clustered the formal DSM-IV-TR Axis II diagnoses hierarchically into: (a) Low strength group: at least one cluster A or B PD present (i.e., paranoid, schizoid, schizotypal, antisocial, borderline, histrionic, and/or narcissistic PD) versus (b) High strength group: at least one cluster C PD or PDNOS present (i.e., avoidant, dependent, obsessive-compulsive, depressive, passive aggressive, and/or mixed PD, but no cluster A or B PD). Second, adaptive personality functioning was measured using the Severity Indices of Personality Pathology (SIPP-SF) (Verheul et al., 2008). The SIPP-SF measures five domains of adaptive personality functioning; high scores reflect adaptive personality, whereas low scores reflect maladaptive personality. We computed a total score by adding all

items and applied a median split to distinguish high from low adaptivity. Third, overall defensive functioning was measured using the Dutch version of the Defense Style Questionnaire (DSQ-60). The DSQ-60 is designed to measure type and degree of the defensive style (Bond, Gardner, Christian, & Sigal, 1983; Thygesen, Drapeau, Trijsburg, Lecours, & de Roten, 2008), high scores reflect a more mature level of defensive functioning, whereas low scores reflect less mature level of defensive functioning. We applied a median split on the Overall Defensive Functioning (ODF) score, to form (a) a relatively mature group versus (b) a relatively immature group. Finally, motivation for treatment was measured using the 8-item Motivation for Treatment Questionnaire (MTQ) (Van Beek & Verheul, 2008). The MTQ consists of two subscales, i.e., Need for help and Readiness to change; high scores reflect high level of motivation, whereas low scores reflect a low level of motivation. A median split was applied on the total score of the 8 items and distinguishes high from low motivation.

Outcome measures

The primary outcome measures were psychiatric symptoms and psychosocial functioning. Psychiatric symptoms were measured using the Dutch version of the Brief Symptom Inventory (BSI) (De Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983), a validated self report scale derived from the revised Symptom Checklist-90 (SCL-90-R) (Arrindell & Ettema, 2003; Derogatis, 1986). In this study, we used the Global Severity Index (GSI) as the mean score of the 53 BSI items. The GSI ranges from 0-4, with higher scores indicating more problems. Psychosocial functioning was measured with two subscales of the Outcome Questionnaire-45 (OQ-45), i.e. Interpersonal relations and Social role functioning (Lambert et al., 1996). The subscale Interpersonal relations ranges from 0-44, the subscale Social role functioning ranges from 0-36, with higher scores indicating more problems. All three outcome measures were assessed at baseline and several follow-up points. Three treatment centers conducted follow-ups at approximately 12, 24, and 36months after baseline; the other three treatment centers conducted follow-ups at the end of treatment, subsequently after about 6 and 12months, and again at 36months after baseline. The use of different assessment points was due to logistic reasons and was taken into account by choosing multilevel modeling as the statistical method for the analyses.

Statistical analyses

Baseline differences between stabilizing groups were analyzed with t-tests for normally distributed variables, Mann-Whitney U tests for non-normal distributed variables and continuity corrected chi2 tests for categorical variables.

We used multilevel modeling to deal with the dependency of repeated measures on the same subject in time and longitudinal data with observations unequally spaced in time. First, we estimated the uncorrected treatment effect at 12 months after baseline using a random intercept and random slope model with time as level I and patient number as level II. Within-group effect sizes (Cohen's d) (Cohen, 1988) were calculated to describe changes from baseline to 12 months for each treatment group. Second, we estimated the treatment effects at 12 months corrected for baseline differences by means of the '*propensity* score' (for a detailed description of this method and its use in psychotherapy research, see Bartak et al. (2009) and Spreeuwenberg et al. (2010). Using the propensity score, we attempt to '*mimic'* random assignment (as in a randomized clinical trial) to psychotherapies with high and low levels of '*destabilization'.* To identify relevant confounders to be used to calculate the propensity score, we considered a list of social and economic variables. All variables significantly related to a specific outcome were used to estimate the univariate propensity scores in a regression analysis, with group membership (high versus low levels of destabilization) as a dependent variable. Diagnostic variables likely to be correlated with the psychological strengths, and the psychological strength variables themselves were not included in the propensity score, as including those would decrease the sensitivity of our design and diminish effects. To compare change in outcome variables across the treatment groups, a sophisticated multilevel model was used. Dependent variables were the change scores (follow up minus baseline) as observed during follow-up for each of the outcome measures. The following independent variables were entered in the model: time, outcome measure at baseline, the propensity score, group membership (high or low level of destabilization), the patient strength characteristic and the interaction between group membership and patient strength characteristic. This model estimated differences in change scores at 12 months after baseline between the two treatment groups.

All analyses were based on intention-to-treat (ITT). ITT is defined as assignment and a minimal exposure to the intended treatment modality. All patients completed at least one follow-up assessment, and received a 'minimally *effective* dosage' of psychotherapy (defined as at least two sessions of outpatient psychotherapy or at least two treatment days of day hospital or inpatient psychotherapy). The ITT analyses are based on the initial treatment assignment and not on the treatment eventually received. Drop-out and crossover between treatments are possible. However, dropout rage seems quite manageable; the proportion of dropout were 12.9% in stabilizing treatments and 19.5% in destabilizing treatments. Furthermore 79.2 percent of patients received the treatment setting they were allocated to. The analyses were performed using SPSS 21 for data preparation and baseline differences. Proc Mixed of SAS 9.3 was applied for multilevel modeling (SAS Institute Inc., Cary, N.C., USA).

RESULTS

Sample characteristics

Of the 735 patients, 69.9% were female, and 30.1% male. The mean age was 33.7 years (SD=9.7). Education was medium to high for 73.6% of the patients. Furthermore, 22.9% of the sample had a parental responsibility. The percentage of patients without a job was 35.2%. The percentage of patients that were married was 21.1%. In terms of PD diagnoses, 8.2% had a cluster A PD, and an additional 24.9% had a cluster B (but no cluster A) PD. Thus, 33.1% of the patients had a cluster A and/or B PD, constituting the 'low strength' group. Furthermore, 38.9% had a cluster C (but no cluster A and/or B) PD, and an additional 28.0% had a PDNOS (but no cluster A, B, and/or C) PD. Thus, 66.9% of the patients had a cluster C PD and/or PDNOS, constituting the *'high strength'* group.

Treatment characteristics

Table 6.1 shows that the average length of the destabilizing psychotherapies is somewhat shorter (7.6±4.8 months) than of stabilizing psychotherapies (11.7±5.3months). Furthermore, destabilizing psychotherapies are more likely to be executed in an inpatient setting than stabilizing psychotherapies (55.1% versus 25.4%), whereas stabilizing psychotherapies are more likely to be executed in a day hospital setting (39.4% versus 30.2%) or outpatient setting (35.2% versus 14.7%) than destabilizing psychotherapies. Higher mean scores for the strength operationalizations DSQ-odf, SIDP-IV and MTQ-total were observed for the destabilizing group. No baseline differences were found for the outcome variables.

Uncorrected outcome

Table 6.2 shows the uncorrected effect sizes for patients with low versus high psychological strengths, both in stabilizing and destabilizing psychotherapies, for each outcome variable and strength operationalization separately. One year after treatment all patients in destabilizing as well as in stabilizing psychotherapies showed improvements in terms of psychiatric symptoms, social role, and relational functioning (Table 6.2). Remarkably, we can observe a consistent pattern in the data, with substantially greater effect sizes in patients with low strengths (effect sizes range 0.8-2.0, median 1.3) than in those with high strengths (effect sizes range 0.0-1.0, median 0.5), both across outcome variables, levels of destabilization and across strength dichotomies (i.e., severity of PD, adaptive personality functioning, and overall defensive functioning), but not for motivation for treatment. With respect to motivation for treatment, we can observe a reversed pattern, with substantially greater effect sizes in patients with high motivation (effect sizes range 1.1-1.9, median 1.2) than in those with low motivation (effect sizes range 0.4-0.6, median 0.5), both across outcome variables and levels of destabilization.

Table 6.1. Socio-demographics, diagnostic and treatment characteristics of all 735 patients and of the patients in the two different psychotherapies

	Total population	Destabilizing psychotherapy	Stabilizing psychotherapy	p-value
N	735	334	401	
Socio-demographics				
Sex (% female)	69.9	64.4	74.6	0.004
Age (mean years ± SD)	33.7 (9.7)	34.7 (10.0)	32.8 (9.3)	0.008
Medium/high education (%)	73.6	77.5	70.3	0.027
Parental responsibility (%)	22.9	21.3	24.4	0.375
Unemployed (%)	35.2	33.2	39.9	0.337
Marital situation				
Never married (%)	67.5	67.4	67.6	0.120
Married (%)	21.1	23.7	19.0	0.057
Widowed or divorced (%)	11.4	9.0	13.5	0.950
Diagnostics [a]				
Cluster A (%)	8.2	8.1	8.2	1.000
Cluster B (%)	24.9	19.5	29.4	0.002
Cluster C (%)	38.9	43.3	35.2	0.027
Cluster NAO (%)	28.0	29.0	27.2	0.634
Strength operationalizations				
SIPP: total	2.6 (0.4)	2.6 (0.4)	2.6 (0.4)	0.168
DSQ: odf	3.6 (0.4)	3.7 (0.4)	3.6 (0.4)	0.006
SIDP-IV: AB vs CNOS (%)	38.9	43.4	35.2	0.027
MTQ: total	59.1 (8.5)	59.8 (7.8)	58.4 (8.9)	0.027
Outcome variables				
GSI	1.5 (0.7)	1.5 (0.6)	1.6 (0.7)	0.619
OQ-45 Interpersonal Relations	21.2 (6.2)	21.2 (6.0)	21.3 (6.3)	0.792
OQ-45 Social Role	15.6 (4.8)	15.8 (4.7)	15.5 (4.9)	0.473
Treatment characteristics				
Duration (mean months ± SD)	9.8 (5.5)	7.6 (4.8)	11.7 (5.3)	<0.001
Outpatient (%)	25.9	14.7	35.2	<0.001
Day-hospital (%)	35.2	30.2	39.4	0.009
Inpatient (%)	38.9	55.1	25.4	<0.001
Drop-out rate (%)	16.5	12.9	19.5	0.022

[a] Assessed with the SIDP-IV, a semi-structured interview for DSM-IV axis II diagnoses. Hierarchically ordered: cluster A (at least one cluster A PD present); cluster B (at least one cluster B PD present, but no cluster A PD), cluster C (at least one cluster C PD present,
but no cluster A or B PD) and cluster NAO (at least one mixed or appendix PD present, but no cluster A,B or C PD).
SIPP=Severity Indices of Personality Pathology, DSQ: odf=Overall Defensive Functioning scale of the Defense Style Questionnaire, SIDP-IV: cluster AB vs CNOS=hierarchically clustered PD groups measured with the Structured Interview for DSM-IV Personality Disorders, MTQ=Motivation for Treatment Questionnaire
GSI = Global Severity Index of the Brief Symptom Inventory, OQ-45 = Outcome questionnaire-45

Table 6.2. Uncorrected mean outcomes (SD) and effect sizes in the four patient-psychotherapy groups for all outcome variables estimated at 12 months after baseline

Outcome	Patient's strenght operationalizations[b]		Patient-psychotherapy groups[a]			
			Stabilizing psychotherapy		Destabilizing psychotherapy	
			Low strenghts[b]	High strenghts[b]	Low strenghts[b]	High strenghts[b]
GSI	SIPP: total	Baseline	1.90 (0.65)	1.20 (0.52)	1.82 (0.55)	1.27 (0.53)
		12 months	0.61 (0.72)	0.96 (0.62)	0.74 (0.80)	1.09 (0.60)
		ES	1.98	0.47	1.96	0.35
		n	200	195	158	169
	DSQ: odf	Baseline	1.84 (0.67)	1.21 (0.52)	1.77 (0.56)	1.33 (0.57)
		12 months	0.65 (0.73)	0.88 (0.64)	0.76 (0.78)	1.06 (0.65)
		ES	1.76	0.61	1.79	0.48
		n	220	178	151	182
	SIDP-IV: cluster AB vs CNOS	Baseline	1.80 (0.73)	1.41 (0.62)	1.61 (0.62)	1.50 (0.60)
		12 months	0.79 (0.72)	0.78 (0.68)	0.86 (0.81)	0.93 (0.68
		ES	1.38	1.02	1.22	0.95
		n	149	249	92	241
	MTQ: total	Baseline	1.33 (0.63)	1.79 (0.66)	1.38 (0.64)	1.68 (0.52)
		12 months	0.95 (0.69)	0.63 (0.66)	1.09 (0.63)	0.70 (0.78)
		ES	0.59	1.75	0.45	1.87
		n	202	190	162	170
OQ-45: Interpersonal relations	SIPP: total	Baseline	24.30 (5.34)	18.18 (5.71)	23.98 (4.95)	18.43 (5.72)
		12 months	14.21 (7.28)	17.95 (6.76)	15.54 (7.45)	17.86 (6.99)
		ES	1.89	0.04	1.71	0.10
		n	143	241	90	238
	DSQ: odf	Baseline	23.46 (5.93)	18.59 (5.70)	23.37 (5.55)	19.30 (5.77)
		12 months	14.87 (7.75)	16.98 (6.77)	15.93 (7.20)	17.39 (7.40)
		ES	1.45	0.28	1.34	0.33
		n	220	179	152	181
	SIDP-IV: cluster AB vs CNOS	Baseline	22.78 (6.37)	20.38 (6.10)	21.53 (5.59)	21.02 (6.17)
		12 months	16.13 (7.95)	16.04 (7.04)	16.87 (7.31)	16.48 (7.34)
		ES	1.04	0.71	0.83	0.74
		n	149	250	91	242
	MTQ: total	Baseline	20.30 (6.34)	22.33 (6.08)	20.18 (6.05)	22.09 (5.84)
		12 months	17.03 (7.02)	15.17 (7.40)	17.44 (7.28)	15.73 (7.29)
		Cohen's d	0.52	1.18	0.45	1.09
		n	202	191	162	171

Table 6.2. Uncorrected mean outcomes (SD) and effect sizes in the four patient-psychotherapy groups for all outcome variables estimated at 12 months after baseline (continued)

Outcome	Patient's strenght operationalizations[b]		Patient-psychotherapy groups[a]			
			Stabilizing psychotherapy		Destabilizing psychotherapy	
			Low strenghts[b]	High strenghts[b]	Low strenghts[b]	High strenghts[b]
OQ-45: Social role	SIPP: total	Baseline	16.95 (4.75)	14.07 (4.67)	17.21 (4.57)	14.49 (4.37)
		12 months	10.90 (5.76)	12.34 (5.07)	11.17 (6.32)	12.86 (5.63)
		ES	1.27	0.37	1.32	0.37
		n	195	188	155	166
	DSQ: odf	Baseline	16.65 (4.95)	14.14 (4.52)	17.18 (4.53)	14.63 (4.46)
		12 months	11.07 (5.93)	12.03 (5.08)	11.23 (6.32)	12.76 (5.67)
		ES	1.13	0.47	1.31	0.42
		n	212	173	148	179
	SIDP-IV: cluster AB vs CNOS	Baseline	16.52 (5.12)	14.95 (4.71)	16.22 (4.52)	15.62 (4.71)
		12 months	11.21 (5.84)	11.75 (5.34)	11.83 (6.44)	12.03 (5.84)
		ES	1.04	0.68	0.97	0.76
		n	141	244	89	238
	MTQ: total	Baseline	14.87 (4.57)	16.35 (5.14)	14.80 (4.70)	16.70 (4.45)
		12 months	12.75 (5.34)	10.58 (5.40)	12.91 (5.56)	10.91 (6.42)
		ES	0.46	1.12	0.40	1.30
		n	198	182	158	169

GSI = Global Severity Index of the Brief Symptom Inventory, OQ-45 = Outcome questionnaire-45
SIPP=Severity Indices of Personality Pathology, DSQ: odf=Overall Defensive Functioning scale of the Defense Style Questionnaire
SIDP-IV: cluster AB vs CNOS=hierarchically clustered PD groups measured with the Structured Interview for DSM-IV Personality Disorders
MTQ=Motivation for Treatment Questionnaire
ES = effect size calculated as Cohen's d
[a] Effect of stabilizing and destabilizing psychotherapy presented for the two levels of patient's psychological strenghts
[b] The high versus low psychological strengths are operationalized with four different measures: SIPP, DSQ, SIDP-IV and MTQ, as presented in the second column.

Corrected outcome

Table 6.3 shows the corrected effect sizes for patients with low versus high psychological strengths, both in stabilizing and destabilizing psychotherapies, for each outcome variable and strength characteristic separately. Furthermore, the main effects of level of destabilization (low versus high), psychological strengths (low versus high), and the interaction effect between level of destabilization and psychological strengths on treatment outcome are shown. Regarding the main effect of level of destabilization, destabilizing psychotherapies showed significantly more improvement on psychiatric symptoms than stabilizing treatments, for the strength variables: '*severity of PD*' (SIDP-IV) and '*motivation*

Table 6.3. Corrected effects and effect sizes in the four patient-psychotherapy groups for all outcome variables estimated at 12 months after baseline

		Patient-psychotherapy groups[a]				Effects					
		Stabilizing		Destabilizing		Treatment		Strengths		Treatment * Strengths	
		Low strenghts[b]	High strenghts[b]	Low strenghts[b]	High strenghts[b]						
Outcome	Psychological strengths[b]	ES	ES	ES	ES	b[c]	p	b[d]	p	b[e]	p
GSI	SIPP: total	0.85	0.94	1.15	1.08	-0.01	.88	-0.14	<.01	0.11	.17
	DSQ: odf	0.82	0.97	1.08	1.14	-0.07	.08	-0.14	<.01	0.06	.45
	SIDP-IV: cluster AB vs CNOS	0.72	0.99	1.02	1.15	-0.13	<.01	-0.15	<.01	0.08	.32
	MTQ: Total	0.79	1.00	0.97	1.23	-0.16	<.01	-0.14	<.01	-0.02	.75
OQ-45: Interpersonal relations	SIPP: total	0.54	0.70	0.85	0.69	0.00	.99	-0.92	.02	1.95	.02
	DSQ: odf	0.50	0.74	0.77	0.77	-0.76	.08	-0.92	.02	1.47	.07
	SIDP-IV: cluster AB vs CNOS	0.42	0.73	0.70	0.80	-1.26	<.01	-1.08	.01	1.30	.14
	MTQ: Total	0.59	0.65	0.68	0.85	-0.72	.08	-0.91	.03	-0.62	.45
OQ-45: Social role	SIPP: total	0.63	0.76	0.73	0.80	-0.48	.11	-0.35	.23	0.28	.63
	DSQ: odf	0.64	0.74	0.65	0.85	-0.70	.02	-0.29	.33	-0.47	.42
	SIDP-IV: cluster AB vs CNOS	0.54	0.77	0.74	0.78	-0.63	.05	-0.50	.12	0.97	.13
	MTQ: Total	0.61	0.77	0.66	0.86	-0.86	.00	-0.32	.28	-0.17	.76

GSI = Global Severity Index of the Brief Symptom Inventory, OQ-45 = Outcome questionnaire-45
SIPP=Severity Indices of Personality Pathology, DSQ: odf=Overall Defensive Functioning scale of the Defense Style Questionnaire
SIDP-IV: cluster AB vs CNOS=hierarchically clustered PD groups measured with the Structured Interview for DSM-IV Personality Disorders
MTQ=Motivation for Treatment Questionnaire
ES = effect size calculated as Cohen's d

[a] Effect of stabilizing and destabilizing psychotherapy presented for the two levels of patient's psychological strengths

[b] The high versus low psychological strengths are operationalized with four different measures: SIPP, DSQ, SIDP-IV and MTQ, as presented in the second column.

[c] This beta shows the average effect for treatment calculated over the stabilzing and destabilzing treatments, expressed as number of standard deviations decrease (-) on the outcome scales (GSI, OQ-45)

[d] this beta shows the average effect for strength calculated over the high and low strength patients , expressed as number of standard deviations decrease (-) on the outcome scales (GSI, OQ-45)

[e] this beta shows the interaction effect (high strength patients in destabilizing treatment) , expressed as number of standard deviations decrease (-) or increase ( ) on the outcome scales (GSI, OQ-45)

for treatment' (MTQ). Furthermore, destabilizing treatments were superior to stabilizing treatments in terms of their impact on relational functioning, only for the psychological strength '*severity of PD*' (SIDP-IV). For social role functioning, we observe the superiority of destabilizing psychotherapies for the psychological strengths *'defensive functioning'* (DSQ), '*severity of PD'* (SIDP-IV) and motivation for treatment (MTQ).

Regarding the main effect of psychological strengths, patients high on psychological strengths show generally significantly better outcomes than patients low on psychological strength. This pattern is most obvious with respect to psychiatric symptoms and interpersonal relational outcome, and least obvious with respect to social role functioning.

Regarding the interaction effect between level of destabilization and psychological strengths, only one significant effect occurred. Patients low on adaptive personality functioning (SIPP) profit more from destabilizing than from stabilizing psychotherapy (which is the opposite towards our hypothesis), whereas patients high on adaptive personality functioning (SIPP) do equally well in both levels of destabilization (also not according to our hypothesis). This matching effect was observed for the improvement in terms of relational functioning, but not for the other outcome variables.

DISCUSSION

In this study we investigated whether patients high on psychological strengths profit more from predominantly destabilizing treatments, whereas patients low on psychological strengths profit more from predominantly stabilizing treatments. This hypothesis is often stated in psychodynamic clinical literature (e.g. Gabbard (2005) and Winston, Rosenthal, and Pinsker (2004)) and used in clinical practice when matching patients to psychotherapies (Van Manen et al., 2012). However, in this large quasi experimental naturalistic study we cannot confirm this matching hypothesis. The findings do show main effects for the level of destabilization (i.e., high level of destabilization is associated with better outcomes) and psychological strengths (i.e., patients high on strengths have better outcomes than those low on strengths), but no interaction effects in line with the matching hypothesis. The only interaction effect that emerged, was opposite to our hypothesis.

Main findings

This study shows a positive impact of a high level of destabilization on treatment outcome, irrespective of psychological strengths and specific outcome variable. Fur-

thermore, to some extent this finding is in contrast with the prevailing view that too much pressure on vulnerable patients increases the risk of drop-out, difficulties to form a stable working alliance, and even psychotic decompensation (Horwitz et al., 1996). Our finding suggests that even vulnerable patients profit from confrontative, expressive, and insight-oriented interventions. Moreover, we found a higher drop-out rate in the stabilizing therapy group. This finding is consistent with the dynamic systems theory perspective as described in the introduction (Hayes & Strauss, 1998; Thelen & Smith, 1994). We suspect that the majority of destabilizing treatments included in our sample, which were predominantly executed in an day-hospital or inpatient setting (86.6%), provide a highly structured and safe environment for patients to have corrective social-emotional experiences, to let go of their old dysfunctional patterns, and to experiment with and adopt new functional patterns. In other words, we suggest that these settings can provide the necessary positive holding environment patients need to work through the high anxiety levels that can occur in an insight-oriented treatment (Bateman & Fonagy, 2001; Lorentzen & Hoglend, 2008).

Our finding that destabilizing psychotherapy has a more positive impact on treatment outcome then stabilizing psychotherapy contrasts with the results of the study of Piper et al. (Piper et al., 1998; Piper et al., 1999). They found in a randomised clinical trial, in an outpatient patient population with a majority suffering from PD, that interpretive psychotherapy provided the same effectiveness as the supportive psychotherapies. The differences in outcome between our study and the study of Piper could be explained by the more intensive setting of the destabilizing treatments in our study. Our hypothesis is that PD patients can only profit fully from a high pressure, destabilizing psychotherapy if the setting provides enough safety, which is for example in a day hospital or inpatient setting. In the study of Piper and colleagues the expressive therapy was (even as the supportive variant) in an outpatient setting.

Furthermore, this study revealed that patients high on psychological strengths, for instance, overall mature defensive functioning, benefit more from psychotherapy than patients low on psychological strengths, irrespective of the level of destabilization and specific outcome variable. This finding is in line with previous research indicating that healthier patients tend to do better in psychotherapy than more severely ill patients (Luborsky et al., 1980). Possibly, healthier patients have psychological resources that enables them to profit from psychotherapy more than severely ill patients. Note that in our study '*healthier'* does not mean *'less psychiatric symptoms, and healthy interpersonal relations and social role'* as we entered these outcome measures at baseline in our multi-level model. The term healthier in this study is restricted to '*psychological strengths',* e.g. motivation and overall defensive functioning.

The matching effect found in this study revealed that patients low on personal strengths profit more from a destabilizing treatment, and patients high on psychological strengths profit equally from destabilizing and stabilizing psychotherapies. This finding is opposite to our hypothesis. Perhaps a consistent reasoning according to the dynamic systems theory can help us interpret this interaction effect: patients high on psychological strengths only require a limited adjustment within the same pattern or attractor state, whereas those low on psychological strengths require a major change including replacing dysfunctional patterns or attractor state by functional ones. Thus, destabilization is not necessary in those high on psychological strengths, while it is in their low-scoring counterparts.

Clinical and scientific implications

Our findings have two important clinical implications. First, our findings discourage clinical practice to routinely match patients low on psychological strengths to supportive or stabilizing variants of psychotherapy. Second, the overall positive effect of destabilizing psychotherapies in a PD population and the lack of evidence for a matching hypothesis strengthens the position of predominantly destabilizing psychotherapies or, at least, the application of expressive and confrontative techniques within psychotherapeutic treatments. Destabilization seems to be beneficial for both the more vulnerable and the relatively healthier PD patient. However, our results do not preclude the possibility that destabilization can involve safety risks and thus iatrogenic effects for patients such as premature drop-out and difficulties in forming a stable working alliance. We would therefore recommend to apply destabilizing techniques in a well-structured, safe, and holding therapeutic environment. An approach to safety in psychotherapeutic environments is offered by Hutsebaut and colleagues, who distinguish between organizational, team and therapist adherence to a treatment model as necessary components of treatment integrity in the implementation of complex interventions for PD patients (Hutsebaut, Bales, Busschbach, & Verheul, 2012).

It is important to note that this study is the first study of treatment matching in PD, which is a highly complex domain of research. Replication of this study will help to build further on a clinically useful evidence base for practitioners, but only a replication of the results in this study in a randomized clinical trial will give enough evidence to implement the results in daily practice. Therefore RCTs are recommended. Furthermore, we would recommend future studies to elaborate on the potentially moderating role of the level of structure, safety and holding in the therapeutic environment, with a beneficial impact of destabilization in safe environments and a negative impact in unsafe environments.

Strengths and limitations

A clear strength of this study is its relatively high external validity. The study is conducted in clinical practice and not under stringent experimental conditions. Nevertheless, it should be recognized that all patients were referred and admitted to specialist psychotherapy. It can therefore not be precluded that our results are not applicable to PD patients who are not referred and admitted to specialist psychotherapy. A second strength is the large number of patients enabling the search for a matching effect. Despite these strengths the present findings have to be interpreted considering several limitations. First, although we controlled for pre-treatment differences or potential confounders using the propensity score, we cannot rule out that some potential confounders still influence the results (Bartak et al., 2010). Furthermore we used an alternative propensity score enabling to find matching effects. For example we did not control for patient characteristics highly correlated with the concept of *'psychological strengths'* in the propensity score. This concern is somewhat mitigated by the fact that reanalysing de data with or without several correlated strength characteristics in the propensity score did not alter the results. Furthermore, the main effect of destabilization and the lack of matching effects were observed with all variants of the propensity score. Further research is undertaken by our research group to investigate the use of the propensity score in subgroup analyses to optimize the power to find a matching effect, while simultaneously retaining control for confounding effects (Van Eeren et al., 2011). A second limitation is that the treatments available in the destabilizing and stabilizing psychotherapies are a mixture of different settings, theoretical orientations and durations (table 6.1). One could argue that the effects we found can be attributed to the differences in for example the setting, not to the '*level of (de)stabilization*' in the treatments. We considered however that the differences in duration and setting is inherent to the concept of *'(de)stabilization of treatment'.* In other words: the setting and duration are not independent of the level of destabilization. Destabilizing treatments often use a '*high pressure cooker model'* that yield good results in a relatively short time span. Stabilizing treatments use a more supportive and time-consuming trajectory. A third limitation concerns the operationalization and measurement of the concept '*destabilization'.* Although we have indications that the reliability of our operationalization is sufficient (correlation among two ratings was r=.69), the validity of our operationalization might be improved. Further investigations could describe at a detailed level all possible stabilizing and destabilizing therapist interventions. Each treatment could then be scored on the most prominent interventions the therapist uses, for example by rating the videotaped sessions by multiple raters. A fourth limitation is that the operationalization of the psychological strength characteristics remains open for debate. We could not find one variable that captured the whole concept, and others have also outlined this definition problem (Bjorklund, 2000; Lake, 1985). In an attempt to overcome this problem, we decided to use four constructs likely to be

highly associated to the '*latent construct*' of psychological strengths. A fifth limitation is the presence of non-response in our data. This may cause a problem for internal validity if non-response is not at random, but related to systematic bias in effect estimation. However, this bias seems unlikely because responders and non-responders did not differ in psychiatric symptoms at baseline, and therefore it seems that they do not represent two structurally different groups of patients (Bartak et al., 2010).

Conclusion

In conclusion our findings do not encourage clinical practice to routinely match patients low on psychological strengths to supportive or stabilizing variants of psychotherapy, and may encourage to routinely consider predominantly destabilizing psychotherapies as an interesting treatment option in these patients. These findings are in favor of the position of destabilizing psychotherapies in the treatment of PD patients.

CONFLICT OF INTEREST

The authors declare no conflicts of interest of any kind.

ACKNOWLEDGEMENT

We would like to thank the following persons and institutes for their contributions in collecting the data and constructive comments in the early stages of this research. The analysis and conclusions of this research project are the responsibility of the authors only.

- Anke Meerman (Pro Persona, Centre of Psychotherapy, Lunteren)
- Bert Van Rossum (Altrecht, Utrecht)
- Moniek Thunnissen (GGZWNB, Halsteren)
- Piet Rijnierse, Mirjam Soons (Arkin, Amsterdam)
- Uli Ziegler (Zaans Medical Centre, Zaandam)
- Helene Andrea, Hester Van Eeren (Viersprong Institute for Studies on Personality Disorders, Halsteren)

REFERENCES

American Psychiatric Association. (2001). Practice guideline for the treatment of patients with borderline personality disorder. *The American Journal of Psychiatry, 158(Suppl10),* 1–52.

Arrindell, W. A.,&Ettema, J.H.M. (2003). *SCL-90 –Herziene handleiding bij een multidimensionele psychopathologieindicator*. Lisse - The Netherlands: Swets Publishers.

Bartak, A., Andrea, H., Spreeuwenberg, M. D., Ziegler, U. M., Dekker, J., Rossum, B. V., … Emmelkamp, P. M. G. (2011b). Effectiveness of Outpatient, Day Hospital, and Inpatient Psychotherapeutic Treatment for Patients with Cluster B Personality Disorder. *Psychotherapy and Psychosomatics, 80*, 88-99.

Bartak, A., Spreeuwenberg, M. D., Andrea, H., Busschbach, J. J., Croon, M. A., Verheul, R., … 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., Emmelkamp, P. M. G. (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., & Fonagy, P. (2000). Effectiveness of psychotherapeutic treatment of personality disorder. *British Journal of Psychiatry, 177*, 138–143.

Bateman, A., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *American Journal of Psychiatry, 158(1),* 36–42.

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)*, 68.

Bjorklund, P. (2000). Assessing ego strength: Spinning straw into gold. *Perspectives in Psychiatric Care, 36(1),* 14–23.

Blatt, S. J., Zuroff, D. C., Hawley, L. L., & Auerbach, J. S. (2010). Predictors of sustained therapeutic change. *Psychotherapy Research, 20(1),* 37–54.

Bond, M., Gardner, S. T., Christian, J., & Sigal, J. J. (1983). Empirical-study of Self-rated defense styles. *Archives of General Psychiatry, 40(3),* 333–338.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Lawrence Erlbaum.

De Beurs, E., & Zitman, F. (2006). The Brief Symptom Inventory (BSI): Betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. *Maandblad Geestelijke Volksgezondheid, 61(2),* 120–141.

Derogatis, L. R. (1986). *SCL-90-R: Administration, Scoring and Procedure Manual - II for the Revised Version*. Townson, MD: Clinical Psychometric Research.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory report. *Psychological Medicine, 13(3),* 595–605.

Gabbard, G. O. (2005). *Psychodynamic psychiatry in clinical practice (4th ed.)*. Arlington, VA: American Psychiatric Publishing, Inc.

Gabbard, G. O., Coyne, L., Allen, J. G., Spohn, H., Colson, D. B., & Vary, M. (2000). Evaluation of intensive inpatient treatment of patients with severe personality disorders. *Psychiatric Services, 51(7),* 893–898.

Hayes, A. M., & Strauss, J. L. (1998). Dynamic systems theory as a paradigm for the study of change in psychotherapy: An application to cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 66(6),* 939–947.

Horwitz, L., Gabbard, G. O., Allen, J. G., Frieswyk, S. H., Colson, D. B., Newsom, G. E., … Coyne, L. (1996). *Borderline personality disorder: Tailoring the psychotherapy to the patient*. Washington, DC: American Psychiatric Association.

Hutsebaut, J., Bales, D. L., Busschbach, J. J. V., & 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*, 10.

De Jong, C. A. J., Derks, F. C. A., van Oel, C. J., & Rinne, T. (1996). *Structured Interview for DSM-IV personality disorders (SIDP-IV) (Dutch version)* [Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen (SIDP-IV)]. Sint Oedenrode, The Netherlands: Stichting Verslavingszorg Oost Brabant.

Koelen, J. A., Luyten, P., Eurelings-Bontekoe, L. H. M., Diguer, L., Vermote, R., Lowyck, B., Bühring, M.E. (2012). The Impact of Level of Personality Organization on Treatment Response: A Systematic Review. *Psychiatry-Interpersonal and Biological Processes, 75(4)*, 355–374.

Lake, B. (1985). Concept of ego strength in psychotherapy. *British Journal of Psychiatry, 147,* 471–478.

Lambert, M. J., Hansen, L. A., Umpress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., Reisinger, C. W. (1996). *Outcome Questionnaire (OQ-45)*: American Professional Credentialing Services LLC.

Leichsenring, F.,&Leibing, 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.

Lorentzen, S., & Hoglend, P. A. (2008). Moderators of the effects of treatment length in long-term psychodynamic group psychotherapy. *Psychotherapy and Psychosomatics, 77(5),* 321–322.

Luborsky, L., Mintz, J., Auerbach, A., Christoph, P., Bachrach, H., Todd, T., … O'Brien, C. P. (1980). Predicting the outcome of psychotherapy. findings of the Penn Psychotherapy Project. *Archives of General Psychiatry, 37(4),* 471–481.

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.

Piper, W. E., Joyce, A. S., McCallum, M., & Azim, H. F. (1998). Interpretive and supportive forms of psychotherapy and patient personality variables. *Journal of Consulting and Clinical Psychology, 66(3),* 558–567.

Piper, W. E., McCallum, M., Joyce, A. S., Azim, H. F., & Ogrodniczuk, J. S. (1999). Follow-up findings for interpretive and supportive forms of psychotherapy and patiënt personality variables. *Journal of Consulting and Clinical Psychology, 67(2),* 267–273.

Soeteman, D., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. (2008). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry, 69(2)*, 259–265.

Soeteman, D., Verheul, R., & Busschbach, J. J. (2008). The burden of disease in personality disorders: diagnosisspecific quality of life. *Journal of Personality Disorders, 22 (3),* 259–268.

Spreeuwenberg, M. D., Bartak, A., Croon, M. A., Hagenaars, J. A., Busschbach, J. J. V., Andrea, H., … Stijnen, T. (2010). The Multiple Propensity Score as Control for Bias in the Comparison of More Than Two Treatment Arms An Introduction From a Case Study in Mental Health. *Medical Care, 48*(2), 166-174.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge MA: MIT Press.

Thygesen, K. L., Drapeau, M., Trijsburg, R. W., Lecours, S., & de Roten, Y. (2008). Assessing defense styles: Factor structure and psychometric properties of the new Defense Style Questionnaire 60 (DSQ-60). *International Journal of Psychology and Psychological Therapy, 8(2),* 171–181.

Van Beek, N., & Verheul, R. (2008). Motivation for treatment in patients with personality disorders. *Journal of Personality Disorders, 22(1),* 89–100.

Van Eeren, H., Spreeuwenberg, M. D., Van Manen, J. G., De Rooij, M., Stijnen, T., & Busschbach, J. (2011). Finding Treatment Effects within Subgroups when Using the Propensity Score to Control for Selection Bias: a Monte Carlo Simulation Study. *Value in Health, 14(7),* A235–A235.

van Manen, J. G., Andrea, H., van den Eijnden, E., Meerman, A. M., Thunnissen, M. M., Hamers, E. F., … Verheul, R. (2011). Relationship between patient characteristics and treatment allocation for patients with personality disorders. *Journal of Personality Disorders, 25*(5), 656-667.

Van Manen, J. G., Kamphuis, J. H., Goossensen, A., Timman, R., Van Busschbach, J. J., & Verheul, R. (2012). In search of patient characteristics that may guide empirically based treatment selection for personality disorder patients – A concept map approach. *Journal of Personality Disorders, 26,* 481–497.

Van Manen, J. G., Kamphuis, J. H., Visbach, G., Ziegler, U., Gerritsen, A., Rijnierse, P., … Verheul, R. (2008). How do intake clinicians use patient characteristics to select treatment for patients with personality disorders? *Psychotherapy Research, 18(6),* 711–718.

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J., … 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.

Wallerstein, R. S. (1986). *Forty-Two lives in Treatment: A Study of Psychoanalysis and Psychotherapy.* New York: Guilford.

Winston, A., Rosenthal, R. N., & Pinsker, H. (2004). *Introduction to supportive psychotherapy.* Arlington, VA, US: American Psychiatric Publishing, Inc.

CHAPTER 7

General discussion

In the introductory chapter of this thesis, three research questions were formulated and these questions will be discussed below in the light of the results presented in the previous chapters. Subsequently, we will elaborate on the implications of this research for clinical practice and future research.

RESEARCH QUESTIONS OF THIS THESIS

1. *Are the improvements previously observed in patients with a cluster A, B, or C Personality Disorder (PD) or Personality Disorder Not Otherwise Specified (PDNOS) stable over five years of follow-up? Are there differences in the effectiveness of outpatient, day hospital, and inpatient treatments on the long-term outcome in patients with a cluster A, B, or C PD or PDNOS?* Chapter 2 showed that five years after baseline, patients with a cluster A, B or C PD had still maintained a reduced symptom levels and better functioning as compared to before treatment. The differences in outcomes of the various treatment modalities at 1-year follow-up were still present at 5-years follow-up. However, most of these differences were small with the exception of the superiority of short-term inpatient psychotherapy in patients with cluster C PDs. Chapter 3 showed that patients with PDNOS also benefitted from psychotherapy – both at short-term and long-term follow-up. Especially short-term inpatient psychotherapy was superior to the other treatment modalities up to 36-months follow-up. At 60-months follow-up, treatment modalities in patients with PDNOS showed mostly comparable effectiveness. In this regard, the superiority of short-term inpatient treatments in PDNOS is less pronounced than that in cluster C PD. In summary, our findings are consistent with: (1) patients with PDs benefit from psychotherapy; (2) the improvements are long-lasting; (3) the long-term outcome is generally independent of treatment modality, with the exception of (4) short-term inpatient psychotherapy, which continues to be superior to most other treatment modalities in patients with a cluster C PD up to 5-years follow-up.

2. *What are the effectiveness and cost-effectiveness of Short-Term Inpatient Psychotherapy based on Transactional Analysis (STIP-TA) in PD patients?* Chapter 4 showed that in 67 matched pairs of patients, patients in both STIP-TA and Other specialized Psychotherapies (OP) showed large symptomatic and functional improvements up to 3-years follow-up. STIP-TA outperformed OP in terms of improvements in general psychiatric symptomatology and quality of life. The superiority of STIP-TA was most pronounced at 12-months follow-up, but remained intact until 3-years follow-up, when two thirds of the STIP-TA patients were symptomatically recovered versus almost half of the OP patients. Chapter 5 showed that STIP-TA is a cost-effective and dominant treatment

compared to OP: i.e. less costly and more effective. The mean three-year costs were €59,834 for STIP-TA and €69,337 for OP, a difference of €-9,503 (95% CI -32,561 to 15,726). QALYs were 2.29 for STIP-TA and 2.05 for OP, a significant difference of .24 (95% CI .05 to .44).

3. *Is there an interaction between the patient's level of psychological strength and the treatment's level of destabilization with respect to outcome?* Chapter 6 studied the hypothesis that psychological strong patients profit more from destabilizing treatments, while psychological weak patients profit more from stabilizing treatments. This study found that PD patients high on psychological strengths generally improved more than patients low on psychological strengths. There was no interaction with the treatment's level of destabilization: patients profited more from a higher level of destabilizing psychotherapy irrespective of their level of psychological strengths.

Below we highlight some clinical relevant results, together with directions for future research that may surpass the limitations of the present research.

IMPLICATIONS FOR CLINICAL PRACTICE AND FUTURE RESEARCH

Large effects of treatment in PD, but impairment remains

Though we found large and statistically significant effects of PD treatments five years later, patients were still more impaired than the general population. This finding is in line with previous work on treatments in PD patients (Bateman, Gunderson, & Mulder, 2015). Whether this difference is due to research factors (Budge, 2015), the fact that PDs or certain aspects of PD might not be changeable (Ferguson, 2010), or inherent limitations or insufficiencies of the included treatments (Morey & Hopwood, 2013), is impossible to determine on the basis of the results and literature presented in this thesis. The finding that impairment is sustained after treatment, combined with the fact that knowledge about the causes is missing, justifies more research on the intrapsychic and extrapsychic processes underlying the pathogenesis, natural course and change in PD in relation to the mechanisms underlying treatment efficacy.

A neglected group: cluster A PD

The desirability of making treatments for PD patients more effective is especially justifiable for individuals with a cluster A PD, as there is a general paucity of treatment studies in this group. One reason for this is that cluster A PD is relatively rare in psychotherapeutic treatment settings. This might be partially due to aspects of the pathology itself. Their social aversions and relative indifference to their disabilities in relationships make

individuals with a cluster A PD less inclined to seek treatment for their personality problems (Bateman et al., 2015). And if they seek treatment at all, they rather seek treatment for axis 1 disorders (Coid, Yang, Tyrer, Roberts, & Ullrich, 2006). A second reason is that the changeability of cluster A PD is often questioned and – as a result – individuals with this pathology are often excluded from psychotherapeutic treatments (Parnas, Licht, & Bovet, 2005). Recent research on the temporal stability of paranoid and schizotypal PD showed that over two thirds of the stability of cluster A symptoms can be attributed to genetic influences and that these genetic risk factors are highly stable. Furthermore, environmental experiences have less influence on the stability of cluster A symptoms and are found unstable (Kendler et al., 2015). Most treatment and research in these patients is therefore dedicated to the effects of psychopharmacology. Remarkably, in our study we found encouraging effect sizes and high clinical significant change rates for this group five years after baseline, which were comparable to those found in patients with other PDs. This suggests that patients with a cluster A PD can profit from psychotherapeutic treatments and that these positive effects of treatment are long-lasting. Our study in cluster A is one of the largest so far, but includes nevertheless only a modest number of patients. Moreover, a control group not receiving treatment was lacking. Given these limitations, the findings as yet cannot be translated into an evidence-based treatment advice. Larger controlled studies are needed to better understand the natural course and role of treatment in the changeability of cluster A PD. However, this is not to say that patients with a cluster A PD should be excluded from psychotherapeutic treatments by default.

Short-term inpatient treatments: an underestimated treatment

In patients with cluster C PD or PDNOS, STIP-TA has proven to be an effective and a cost-effective treatment. STIP-TA improves quality of life more and faster than other specialized treatments in an outpatient, day hospital or inpatient setting. Furthermore, although this treatment is initially expensive, it is cost saving on the medium and long run in these patients, who are typically seen as less fragile and might thus be responsive to brief and intensive treatments. On the basis of these findings, a short-term inpatient treatment ought to be one of the treatment options, especially in cluster C patients and PDNOS.

Effective ingredients of STIP-TA

We do not know exactly why the effectiveness of STIP-TA is superior, but propose some likely reasons. For one, the amount of confronting and destabilization in this treatment is considerable (van Manen, Horn, Stijnen, Busschbach, & Verheul, 2015), which is why this treatment is also referred to as a "*pressure cooker*". The high amount of destabilization is achieved by several aspects worth mentioning. First, the treatment is limited in time.

This probably promotes therapeutic engagement and motivation and leads to a heightened focus on the individual treatment aims (Town, Abbass, & Hardy, 2011). Second, STIP-TA is a so-called focal treatment and has integrated the redecision model of Goulding and Goulding [13]. This is reflected in the formulation of an individually determined therapeutic contract at the beginning of treatment about the goals for change – and which drives the treatment. In this way the therapeutic focus is established early in treatment and helps to address defences and works against avoidance (Town et al., 2011). The '*focality of treatment*' further enables the therapists to allocate their efforts more effectively, which is often hypothesized to be associated with outcome, although so far no study has found conclusive evidence (Irving et al., 2004). Third, all treatment components are considered to be therapeutic. The pressure is kept high, not only in the treatment sessions but also by the presence and interventions of psychosocial nurses. Furthermore, daily activities are constantly linked to the treatment contract which makes it almost impossible for patients to avoid painful emotions and experiences in the treatment group. The finding that avoidance in treatment is linked to interferences with processes necessary for therapeutic progress reinforces the importance of the high pressure (Meier, 2014). Fourth, the treatment is daily evaluated by the staff and by the patients on the basis of their treatment contracts, which routine helps to keep both patients and the staff focused on the objectives of the therapeutic contract.

This intensive and destabilizing treatment might be possible and effective only in this patient group, since STIP-TA at the same time is also a highly structured treatment and much attention is directed to create and keep a safe and holding environment. This is reflected in the many staff meetings scheduled, on account of which the patients can be carefully monitored in their treatment process and progress and the staff members are able to timely anticipate to events in the treatment. The therapists are actively involved and the transference relationship is used to mirror patients and to obtain more insight in the individual patient. The creation and maintaining of a strong therapeutic alliance on itself might be a strong contributor to effective treatments (Falkenstrom, Granstrom, & Holmqvist, 2013). Furthermore, much attention is directed to the group cohesiveness, which is found to have a large influence on treatment outcome (Burlingame, McClendon, & Alonso, 2011).

Another possible reason is that STIP-TA reinforces patients to be actively involved and responsible for their own treatment progress. After allocation to the treatment, patients spend an orientation day on the ward and are supposed to write a reflection letter afterwards in which they motivate why they have chosen for this treatment. The possibly resultant heightened feeling of autonomy is in itself positively correlated with treatment outcome (Ryan & Deci, 2008). A second aspect which might have a positive influence

on autonomy and motivation is the use of the easily learned theoretical language of TA. This empowers the patient to have a positive influence on his or her own treatment and might enhance compliance to treatment.

STIP-TA is a treatment with half-open groups including '*senior'* patients who have started the treatment six weeks earlier than the *'junior'* patients. In this way, the senior patients can function as role models for the junior patients; they pass on the treatment culture and demonstrate how to explore and deal with problems. This might also lead to heightened feelings of hope and it can foster optimism in junior patients with a beneficial effect on the treatment outcome (Irving et al., 2004). The group in itself may function as a platform for the sharing of corrective experiences, which in combination with building insight and emotional processing appeared to be key factors in psychotherapeutic change (Abbass et al., 2014).

In sum, the combination of a high amount of destabilization and confronting in an empathic manner within a safe and holding environment seems to be the most effective aspect of STIP-TA. Furthermore, the attention to the increase of autonomy and motivation, and the half-open group setting seem to be factors positively related to outcome.

Extension patient group

We found that short-term inpatient treatments were superior in cluster C PD and slightly less so in PDNOS patients. In patients with cluster A PD or cluster B PD, inpatient and day hospital treatments were associated with the largest effect sizes. However, after correction for differences at baseline, the differences between treatment modalities disappeared, too. This difference in results might be due to the broader definition of the treatment modality and the lack of specification of length in cluster A and B PD. While treatment modalities in cluster C and PDNOS were defined in terms of a combination of length and setting of treatment – which resulted in five and six treatment groups, respectively – in the treatment studies of cluster A and B PD, modality was only defined in terms of setting – which resulted in three treatment groups. At least in cluster C we found significant differences between treatment modalities only when the treatment length was added to the modality and the treatments were split up in short-term and long-term treatments. It is possible that smaller effects of long-term inpatient treatments reduced the overall effects of inpatient treatments. Most cluster A and B patients followed long-term treatments however: 81% and 75%, respectively. This implies that clinicians are inclined to send these patients to long-term treatments which might be based on the hypothesis that more vulnerable patients cannot handle the higher pressure in short-term treatments. One can imagine that in BPD patients a destabilizing treatment might impair the mentalizing capabilities and eventually leads to an ineffec-

tive treatment (Fonagy, Luyten, & Strathearn, 2011). Some investigators have pointed to a possible risk for iatrogenic effects in this patient group, such as disturbed therapeutic alliances, suicidality or drop-outs (Fonagy & Bateman, 2006). There is limited evidence for this idea, although the observation by Giesen-Bloo et al. (2006) that BPD patients show significant higher attrition rates in transference focused therapy (TFP) as compared to schema-focused therapy (SFT) is consistent with this idea. This finding might be accounted for by more confronting and therefore destabilizing techniques of TFP, thereby evoking too much anxiety in some BPD patients and leading to dropout. Nonetheless, some patients with a BPD or a different cluster B PD seem to be able to bear the pressure. Though it would be premature to allocate these patients to a short-term inpatient treatment, it is well worth investigating whether some of these patients can profit from a short-term inpatient treatment.

In PDNOS the short-term inpatient treatments in general were most effective up to 36 months of follow-up. Hereafter, differences between modalities disappeared. However, patients in the short-term treatments on average experienced fewer symptoms over the follow-up period than did patients in the other modalities. When combining these results with the fact that almost half of the STIP-TA group consisted of patients with a PDNOS diagnosis, the recommendation that a short-term inpatient treatment should be a treatment option for PDNOS patients is justifiable.

Costs and effects

The recommendation that a short-term inpatient treatment should be a treatment option for patients with a cluster C PD or PDNOS is opposed to the current ideas of deinstitutionalization and the phasing out of inpatient treatments for (PD) patients in the Netherlands and other countries. In the Netherlands, and also in other countries, health care reforms, economic crises, and the resultant emphasis on patient empowerment have led to a reduction of long-term intensive treatments. Within this movement, all inpatient psychotherapies seem to be reduced, both long- and short-term. The present thesis suggests that despite the high initial costs, short-term inpatient treatments are effective as well as cost-effective in patients with cluster C PD or PDNOS. Though the treatment price per patient seems leading in the reimbursement of treatments, intensive treatments in these patient groups pay back their investments as initially inexpensive treatments become more expensive on the long run due to costs outside of the primary treatment. The trend to phase out inpatient treatments will therefore result in the reduction of cost-effectiveness in the treatment of personality disorders.

INTEGRATIVE TREATMENTS: COMBINING THE BEST PARTS OF EQUALS

The current study found that a short-term inpatient treatment was superior in terms of costs and effectiveness compared to other psychotherapies in cluster C PD and PD-NOS. Concerning cluster A and B PDs, the treatment modality, defined as setting and duration, seemed not to influence effectiveness of the treatment. This thesis does not clarify which factors account for a higher or lower effectiveness, whether effectiveness depends on setting, duration, psychotherapeutic orientation, a combination thereof, or even another yet unknown factor. A recent review concluded that evidence-based treatments were more effective than treatment as usual (TAU) and that comparative trials on active treatments have not yet provided conclusive evidence for the superiority of one theoretical orientation over another (Budge et al., 2013). A recent review on potential mechanisms of change in psychotherapeutic treatments for personality disorders found that most evidence points to the importance of '*common or non-specific factors*' such as the therapeutic alliance and the resolution of ruptures in the alliance in the process of change (Forster, Berthollier, & Rawlinson, 2014). Studies confirmed that these common factors have the largest impact on treatment effectiveness and that its impact is considered twice as large compared to the used techniques (Lambert & Barley, 2002). Next to the therapeutic alliance and the resolution of its ruptures, three other factors emerged: 1) therapy factors (e.g. a coherent conceptual structure and a treatment which should provide a rationale for selecting interventions); 2) therapist factors (e.g. the attitude of empathy); and 3) patient factors (e.g. motivation; Livesley, 2012). Instead of further comparing one treatment with another, the next logical step in psychotherapeutic research to improve mental health treatments in PDs are studies on how treatments work: research on the pathways to improvement, on effective ingredients of treatments, and on common elements is warranted.

With more knowledge about effective ingredients and common elements, effective and innovative integrative treatments can be conceptualized for PD patients. Most evidence-based treatments have their own focus, such as emotion regulation in DBT or mentalizing in MBT. However, PDs are '*multidimensional phenomena*' with a large degree of axes I and II comorbidity which require treatments that address several problem areas (Nelson, Beutler, & Castonguay, 2012). Livesley (2012) proposed a framework for an integration of general and PD specific therapeutic approaches and techniques. 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 patients, e.g. focusing on emotion regulation skills or mentalizing (Livesley, 2012). Combining different treatment approaches potentially endangers treatment coherence. The therapist can be inclined to bend with the patient every session. One of the

efficacious elements of treatments in general is the consistent application of a theoretic framework (Bateman & Fonagy, 2000). Therefore, in the integration and evaluation of evidence based treatments, psychotherapy manuals are vital. Instead of providing a session-to-session manual as is mostly done in axis 1 disorders, in PDs this should be rather in the form of guidelines and operationalization of treatment (Westen, 2002). By reducing the individual variability, the quality of treatments is improved by lowering the propensity for errors and treatments can be evaluated on their (cost-)effectiveness. Treatments should furthermore be limited in time and provide for fixed evaluation moments to keep the focus on the target(s) of treatment and to prevent open-ended treatments.

On the other hand, different psychotherapeutic treatments might have more in common than seen at first sight. This line of thought is also reflected in the Common Language for Psychotherapy (CLP) project. This project stresses the fact that the absence of a common language in psychotherapy might lead to confusion in both therapists and patients as different terms seem to be used for identical interventions. By creating an ecumenical lexicon of psychotherapy procedures, a widely accepted, common language for psychotherapy procedure can be developed that will reduce confusion and accelerate the evolution of psychotherapy into a scientific craft (Marks & Fullana).

STRENGTHS AND LIMITATIONS

The strength of the Sceptre study stems from the following factors: a) the high number of patients, b) the long-term follow-up over five years, c) a high follow-up response, d) the wide range of treatments under study (which can be considered highly representative of specialist mental health care in the Netherlands), and e) the naturalistic design of using a sophisticated statistical technique to mimic randomized treatment allocation (propensity score).

The need for such a sophisticated statistical technique is rooted in the principal limitation of this study, the lack of randomization of patients to treatment. This limitation is mitigated that we rigorously controlled for initial patient differences as potential confounders by means of the multiple propensity score. It is important to acknowledge that this score corrects for observed differences only. It is possible that other patient differences, which were either not predetermined or not observable, affected assignment to treatment and confounded the observed differences in treatment effectiveness (Austin, 2008). However, we controlled for a substantial number of social and diagnostic variables, minimizing the likelihood of non-observed confounding factors. Furthermore,

Randomized Controlled Trials (RCTs) are often criticized for their limited external validity as treatments take place under strictly controlled or experimental conditions: only patients who agree to randomization are studied, and exclusion criteria are typically stringent (e.g. Hodgson, Bushe, & Hunter, 2007). Besides, comparisons between markedly different treatment modalities or dosages, as were made in this study, are not feasible in a randomized study. This is reflected in an earlier attempt to compare the short-term inpatient treatment (STIP-TA) to an outpatient treatment in cluster C PD patients in a RCT, and which failed due to patient preferences (PSILO trial, National Academic Research and Collaborations Information System).

A second limitation of this study is the focus on treatment dosage and it did not consider other treatment or patient attributes such as a) the potential impact of theoretical orientation, b) the occurrence of axis I disorders and the use of psychotropic medication or c) treatment fidelity.

We came across three complications to investigate the impact of theoretical orientation (a) in this study design. First, more than 50% of the given treatments were considered integrative treatments (e.g., Schema-Focused Therapy (SFT), Dialective Behavior Therapy (DBT) and Mentalization Based Treatment (MBT)). Second, as theoretical orientation was associated with setting and duration, a different study design would have been necessary to explore the effect of theoretical orientations in addition to treatment modalities. Third, studies for PD typically show that theoretical orientations only account for small differences in effectiveness (Bartak et al., 2007). Other aspects of treatment, such as a coherent theoretical structure or the strength of the therapeutic alliance seem more effective ingredients of treatment than the theoretical orientation (Bateman & Fonagy, 2000).

We do not have information on the occurrence of axis I comorbidity in PD patients (b), though they are often reported to be associated with PDs (Tyrer, Reed, and Crawford, 2015). An earlier study furthermore showed that the severity of symptoms (thereby including the presence of axis I disorders) had an influence on treatment allocation (specifically treatment intensity) (van Manen et al. 2008). The prescription of psychotropic medication is not part of the guideline to the treatment of PDs in general but is advised as a supplementary treatment for comorbid axis I disorders. In our study, almost half of the patients had used psychotropic medication before entering the study. The use of psychotropic medication might have facilitated patients with an axis I disorder to profit from the psychotherapeutic treatment aimed at PD and therefore might have had an effect on treatment effectiveness.

We did not assess the treatment fidelity (c) and therefore do not know whether therapists have administered the treatments as described in the treatment manuals. It is reasonable to assume that treatments delivered in specialized treatment centers by experienced psychotherapists are of relatively high quality, however without evidence of treatment fidelity we do not know whether the treatments described were actually performed.

A third limitation of this study is the absence of a control group (i.e. a group of patients not receiving treatment). Thus, we cannot evaluate which part of the improvement is due to therapy and which part is due to, e.g. natural recovery. For ethical reasons, it would be impossible to compare an active treatment with a placebo or waiting list. The current study can therefore rather be described as a "*comparative effectiveness"* approach, comparing two or more effective treatments in patients that are representative for regular practice. The objective of these studies is to deliver the information decision makers need to know (Sox & Greenfield, 2009; Institute of Medicine, 2009). Nonetheless, previous research has shown that PD patients in treatment have a seven-fold faster recovery than patients without treatment (Perry et al., 1999) and that psychotherapeutic treatment can lead to a remission of symptoms (Zanarini, Frankenburg, Reich, & Fitzmaurice, 2012). However, the changeability of PDs is still an ongoing discussion (Morey & Hopwood, 2013) and in a recent meta-analysis it was concluded that psychotherapy should aim on behaviour change, rather than on global personality change, since expecting changes in fundamental personality structures may be unrealistic (Ferguson, 2010).

Fourth, we do not know whether the underlying personality pathology in our study changed as the outcome variables relied on certain assessment instruments and on self-reporting. However, it seems unlikely that the large effect sizes on the outcome variables found in the current study can occur when the underlying pathology still exists in its original form.

CONCLUSION

The current study showed that psychotherapy in patients with a PD yielded long-lasting treatment results, though functional impairments remained. In patients with cluster A or B PD, differences between modalities or dosages of treatment were small and mostly non-significant in the long run. In patients with cluster C PD or PDNOS, short-term inpatient treatment based on transactional analysis (STIP-TA) proved to be the most effective and cost-effective treatment option, at both short and long term. This specific treatment program deserves more clinical and scientific attention than it currently is receiving.

From a clinical perspective, STIP-TA would be an interesting option to be considered in any treatment allocation process in patients with cluster C PD or PDNOS. From a scientific perspective, it is worthwhile to compare the (cost-)effectiveness of STIP-TA with that of well established, evidence-based treatment options for the target population. More in general, this thesis underlines the need for more studies about the question which treatment works for whom, studies into the effective general ingredients of psychotherapy in PD patients, and the need for the integration of treatment techniques.

> *"The time is ripe for the treatment of PD to move beyond a "competing schools or therapies" approach and adopt a more integrated perspective that combines treatment principles and methods that work regardless of their conceptual origins."* (Livesley, 2012) Dr. John Livesley in 'Integrated Treatment: A Conceptual Framework for an Evidence-Based Approach to the Treatment of Personality Disorder' (Journal of Personality Disorders, 2012. 26(1): p. 17-42)

REFERENCES

Abbass, A. A., Kisely, S. R., Town, J. M., Leichsenring, F., Driessen, E., De Maat, S., … Crowe, E. (2014). Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database Syst Rev, 7*, CD004687.

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.

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.

Bateman, A. W., & Fonagy, P. (2000). Effectiveness of psychotherapeutic treatment of personality disorder. *British Journal of Psychiatry, 177*, 138-143.

Bateman, A. W., Gunderson, J., & Mulder, R. (2015). Treatment of personality disorder. *The Lancet, 385(9969),* 735-743.

Budge, S. L. (2015). The Effectiveness of Psychotherapeutic Treatments for Personality Disorders: A Review and Critique of Current Research Practices. *Canadian Psychology/Psychologie canadienne, 56(2),* 191-196.

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.

Burlingame, G. M., McClendon, D. T., & Alonso, J. (2011). Cohesion in group therapy. *Psychotherapy, 48(1)*, 34-42.

Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry, 188,* 423-431.

Falkenstrom, F., Granstrom, F., & Holmqvist, R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology, 60(3),* 317-328.

Ferguson, C. J. (2010). A Meta-Analysis of Normal and Disordered Personality Across the Life Span. *Journal of Personality and Social Psychology, 98(4),* 659-667.

Fonagy, P., & Bateman, A. (2006). Progress in the treatment of borderline personality disorder. *British Journal of Psychiatry, 188*, 1-3.

Fonagy, P., Luyten, P., & Strathearn, L. (2011). Borderline personality disorder, mentalization, and the neurobiology of attachment. *Infant Mental Health Journal, 32(1),* 47-69.

Forster, C., Berthollier, N., & Rawlinson, D. (2014). A Systematic Review of Potential Mechanisms of Change in Psychotherapeutic Interventions for Personality Disorder. *Journal of Psychology and Psychotherapy, 4(1),* 133.

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., … Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of General Psychiatry, 63(6),* 649-658.

Gunderson, J. G., Stout, R. L. , McGlashan, T. H., Shea, M. T., Morey, L. C., Grilo, C. M. … Skodol, A. (2011). Ten-year course of borderline personality disorder: Psychopathology and function from the collaborative longitudinal personality disorders study. *Archives of General Psychiatry, 68(8):* p. 827-837.

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.

Institute of Medicine (2009). *Initial National Priorities for Comparative Effectiveness Research*. Washington, D.C.: The National Academies Press.

Irving, L. M., Snyder, C. R., Cheavens, J., Gravel, L., Hanke, J., Hilberg, P., Nelson, N. (2004). The Relationships Between Hope and Outcomes at the Pretreatment, Beginning, and Later Phases of Psychotherapy. *Journal of Psychotherapy Integration, 14(4),* 419-443.

Kendler, K. S., Aggen, S. H., Neale, M. C., Knudsen, G. P., Krueger, R. F., Tambs, K., … Reichborn-Kjennerud, T. (2015). A longitudinal twin study of cluster A personality disorders. *Psychol Med, 45(7),* 1531-1538.

Lambert, M. J., & Barley, D. E. (2002). Research summary on the therapeutic relationship and psychotherapy outcome. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 17-32). Oxford, England: Oxford University 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.

Marks, I., & Fullana, M. A. *Common Language for Psychotherapy (CLP) procedures.* Retrieved mey 2nd, 2015, from http://commonlanguagepsychotherapy.org/

Meier, S. T. (2014). Rediscovering the role of avoidance in psychotherapy progress and outcome. *Professional Psychology: Research and Practice, 45(3),* 212-217.

Morey, L. C., & Hopwood, C. J. (2013). Stability and change in personality disorders. *Annual Review of Clinical Psychology, 9,* 499-528.

National Academic Research and Collaborations Information System. *PSILO trial. A randomized comparison of the efficacy and cost-effectiveness of Psychotherapy Short-term Inpatient versus Long-term Outpatient.* Retrieved 06-01-2014, 2014, from http://www.narcis.nl/research/RecordID/OND1324951.

Nelson, D. L., Beutler, L. E., & Castonguay, L. G. (2012). Psychotherapy Integration in the Treatment of Personality Disorders: A Commentary. [doi: 10.1521/pedi.2012.26.1.7]. *Journal of Personality Disorders, 26(1),* 7-16.

Parnas, J., Licht, D., & Bovet, P. (2005). *Cluster A Personality Disorders: A Review Personality Disorders* (pp. 1-124). New York: John Wiley & Sons, Ltd.

Perry, J. C., Banon, E., & Ianni, F. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry, 156(9),* 1312-1321.

Ryan, R. M., & Deci, E. L. (2008). A self-determination theory approach to psychotherapy: The motivational basis for effective change. *Canadian Psychology/Psychologie canadienne, 49(3),* 186-193.

Simon, W. (2009). Follow-up psychotherapy outcome of patients with dependent, avoidant and obsessive-compulsive personality disorders: A meta-analytic review. *International Journal of Psychiatry in Clinical Practice, 13(2),* 153-165.

Sox, H. C. & Greenfield, S. (2009). Comparative effectiveness research: a report from the Institute of Medicine. *Annals of Internal Medicine, 151(3),* 203-205.

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.

Tyrer, P., Reed, G. M., & Crawford, M. J. (2015). Classification, assessment, prevalence, and effect of personality disorder. *The Lancet, 385(9969),* 717-726.

Van Manen, J. G., Kamphuis, J. H., Visbach, G., Ziegler, U., Gerritsen, A., Rijnierse, P., … Verheul, R. (2008). How do intake clinicians use patient characteristics to select treatment for patients with personality disorders? *Psychotherapy Research, 18(6),* 711–718.

Van Manen, J. G., Horn, E. K., Stijnen, T., Timman, R., Busschbach, J. J. V., and Verheul, R. (2015), Tailoring psychotherapy in patients with personality disorders: Matching the level of psychological

strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health, 9*, 133–149.

Westen, D. (2002). Manualizing manual development. *Clinical Psychology-Science and Practice, 9(4),* 416-418.

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2010). The 10-year course of psychosocial functioning among patients with borderline personality disorder and axis II comparison subjects. *Acta Psychiatrica Scandinavica, 122(2)*: p. 103-109.

SUMMARY

The current thesis studied the effectiveness of psychotherapeutic treatments in patients with a personality disorder (PD). Personality disorders are highly invalidating disorders and one of the most frequent disorders treated in outpatient mental health care. In spite of this, there is a paucity of research on effective and cost-effective treatments in this patient group. This thesis uses data collected from a quasi-experimental study on the effectiveness and cost-effectiveness of treatment dosages in patients with a PD, the Study on Cost-Effectiveness of Personality Disorder Treatment (SCEPTRE). Treatment dosages are defined as outpatient, day hospital or inpatient treatment with a short (<six months) or long (≥ six months) duration. Earlier studies based on the SCEPTRE data showed that dosage had some influence on the effectiveness on the short- and mid-term follow-up and some influence on the modelled cost-effectiveness of treatments. The main findings of each chapter are discussed in this summary.

Chapter 1 is the general introduction of this thesis.

Chapters 2 and 3 both report on the long-term effectiveness of different modalities of psychotherapy. As the SCEPTRE study was quasi-experimental we used multiple propensity scores to correct for initial baseline differences. The two research questions were as follows: 1) Are the improvements previously observed in patients with a DSM-IV-TR cluster A, B, C PD, or PDNOS stable over five years of follow-up? 2) Are there differences in the effectiveness of outpatient, day hospital, and inpatient treatments on the long-term outcome in patients with a DSM-IV-TR cluster A, B, C PD, or PDNOS? The effectiveness studies described below (Chapters 2, 3, and 4) all used symptom severity as the primary outcome measure and psychosocial functioning and health related quality of life as secondary outcome measures. Effectiveness was assessed over 60 months after baseline. **Chapter 2** reports on the long-term effectiveness of different treatment modalities in patients with a cluster A, B and/or C PD. Previous research on short- and mid-term follow-up showed some dosage-effect relationships, with superiority of short-term inpatient psychotherapy in patients with a cluster C PD being the most significant finding. We followed 519 patients who were assigned to outpatient, day hospital or inpatient treatments during five years. Of these, 59 patients had a cluster A PD, 210 patients had a cluster B PD, and 416 patients had a cluster C PD. As patients could have more than one PD, the sum of these exceeds the total number. In patients with cluster C PD, short-term treatments (< six months) were distinguished from long-term (≥ six months) treatments. Uncorrected results showed that all patient groups except cluster A outpatients reported significantly less symptom severity, and that initial positive outcomes were sustained over the 5-year follow-up period. Corrected differences showed small and

mostly non-significant differences between the modalities in the three patient groups. When distinguishing between long-term outpatient, short- and long-term day hospital and short- and long-term inpatient treatments for cluster C patients, the previously observed superiority of short-term inpatient psychotherapy compared to other treatment modalities at 12 months follow-up was still present at 5-year follow-up. In **chapter 3** we explored the long-term effectiveness of different modalities in patients with a PDNOS, i.e. short- or long-term outpatient, short- or long-term day hospital, and short- or long-term inpatient treatments. A total of 205 patients with PDNOS were assigned to one of these six treatment modalities. At all time points, all treatment modalities showed positive outcomes, especially in terms of improvements of symptom severity and social role functioning. When we adjusted for initial baseline differences, short-term outpatient psychotherapy and short-term inpatient psychotherapy were associated with the greatest improvement and generally outperformed the other modalities in terms of symptom severity at 12-months follow-up. At 60 months after baseline, observed differences between modalities mostly had diminished. This might be due to a decrease in effects in short-term outpatient and short-term inpatient psychotherapies over time while the effects in long-term outpatient and long-term inpatient psychotherapies were increasing. We concluded that patients with PDNOS benefit from psychotherapy both at the short and long term. And while short-term outpatient psychotherapy and short-term inpatient psychotherapy seem to be superior to the other treatment modalities at 12-months follow-up, the differences in effectiveness disappeared at 60-months follow-up.

Chapters 4 and 5 report on the effectiveness and cost-effectiveness of Short-Term Inpatient Psychotherapy based on Transactional Analysis (STIP-TA) to Other specialized Psychotherapies (OP) in 67 pairs of patients with mainly a cluster C PD or PDNOS. As the SCEPTRE study was quasi-experimental we used the propensity score to match STIP-TA patients with OP patients. The effectiveness and cost-effectiveness were assessed over 36 months after baseline. **Chapter 4** compares the effectiveness of STIP-TA to OP. In 67 pairs of patients, both STIP-TA and OP showed large symptomatic and functional improvements, with STIP-TA patients showing more symptomatic improvement at all time points compared to OP patients. At 36 months, 68% of STIP-TA patients were symptomatically recovered compared to 48% of OP patients. STIP-TA outperformed OP in terms of improvements in general psychiatric symptomatology and quality of life. The superiority of STIP-TA was most pronounced at 12-months follow-up, but remained intact over the course of the 3-year follow-up. In **chapter 5** we compared the cost-effectiveness of STIP-TA to OP. We made use of the societal perspective and measured healthcare costs next to lost production costs. Cost-effectiveness was represented by costs per quality adjusted life years (QALYs) and uncertainty was assessed using bootstrapping. STIP-TA

was less costly and more effective than OP, which makes STIP-TA a dominant treatment. We concluded that STIP-TA is a cost-effective treatment in PD patients.

Chapter 6 reports on the matching hypothesis that patients high on psychological strengths profit more from predominantly destabilizing treatments while patients low on psychological strengths profit more from predominantly stabilizing treatments. We followed 735 PD patients who were assigned to different levels of stabilizing and destabilizing psychotherapies. Outcome measures were symptom severity and psychosocial functioning. At 12-months follow-up, patients who followed destabilizing psychotherapies had slightly better outcomes than patients who followed stabilizing psychotherapies. Patients high on psychological strengths generally improved slightly more than patients low on psychological strengths. We further found that patients profited more from a higher level of destabilizing psychotherapy – irrespective of their psychological strengths. The results therefore imply that destabilizing psychotherapies should be considered a first treatment option for PD patients with either high or low psychological strength.

Chapter 7 is the general discussion where the findings of this thesis are summarized and in which the implications for clinical practice and future research are discussed. We conclude that psychotherapy in patients with a PD yields long-lasting treatment results, but that functional impairments can remain. In patients with cluster A or B PD, differences between modalities or dosages of treatment were small and mostly non-significant in the long run. In patients with cluster C PD or PDNOS, short-term inpatient treatment based on transactional analysis (STIP-TA) proved to be the most effective and cost-effective treatment option, at both short and long term. This specific treatment program deserves more clinical and scientific attention than it currently is receiving. From a clinical perspective, STIP-TA would be an interesting option to be considered in any treatment allocation process in patients with cluster C PD or PDNOS. From a scientific perspective, it is worthwhile to compare the (cost-)effectiveness of STIP-TA with that of well established, evidence-based treatment options for the target population. More in general, this thesis underlines the need for more studies about the question which treatment works for whom, studies into the effective general ingredients of psychotherapy in PD patients, and the need for the integration of treatment techniques.

SUMMARY IN DUTCH

In dit proefschrift staat de effectiviteit van psychotherapeutische behandelingen bij patiënten met een persoonlijkheidsstoornis (PS) centraal. Persoonlijkheidsstoornissen zijn ernstige stoornissen en behoren tot de meest behandelde stoornissen in de ambulante GGZ. Desalniettemin bestaan er slechts een beperkt aantal wetenschappelijke studies naar effectieve en kosteneffectieve behandelingen voor deze stoornis. Het voorliggende proefschrift gebruikt data die zijn verzameld in het kader van een quasi-experimenteel onderzoek naar de effectiviteit en kosteneffectiviteit van psychotherapeutische behandelingen voor patiënten met een PS, de "Study on Cost-Effectiveness of Personality Disorder Treatment" (SCEPTRE). Centraal in de SCEPTRE studie staat de vergelijking van verschillende modaliteiten of doseringen van behandelingen. Deze doseringen zijn gedefinieerd als ambulante, dagklinische en klinische behandelingen met een korte (< zes maanden) of lange (≥ zes maanden) duur. Eerdere onderzoeken die zijn gebaseerd op data van het SCEPTRE onderzoek, wezen uit dat dosering inderdaad invloed heeft op de effectiviteit op de korte en middenlange duur van de follow-up. Met name bij de zogenaamde type C persoonlijkheidsstoornissen bleek een kortdurende klinische psychotherapie de meest effectieve behandeling. Daarnaast bleek dosering van invloed op de kosteneffectiviteit van de behandelingen van type B en C persoonlijkheidsstoornissen. In dit proefschrift wordt gekeken naar de effectiviteit op de lange termijn, naar de effectiviteit van behandeling van "persoonlijkheidsstoornissen niet anderszins omschreven" (PSNAO) en naar de (kosten-) effectiviteit van een bepaalde behandeling, namelijk de kortdurende klinische psychotherapie gebaseerd op de principes van de Transactionele Analyse (Short-term Inpatient Psychotherapy - Transactional Analysis; STIP-TA).

Hoofdstuk 1 vormt de algemene inleiding van dit proefschrift.

In de hoofdstukken 2 en 3 wordt de effectiviteit van verschillende modaliteiten van psychotherapie over vijf jaar follow-up beschreven. Aangezien SCEPTRE als een quasi-experimentele studie is opgezet, hebben we gebruik gemaakt van de multipele propensity score om te corrigeren voor initiële verschillen tussen patiënten vóór begin van de studie. De drie onderzoeksvragen voor onderstaande drie studies zijn als volgt:

1) Zijn de behaalde verbeteringen die eerder zijn gevonden bij patiënten met een cluster A, B of C PS stabiel over de vijf jaar follow-up?
2) Profiteren patiënten met een PSNAO van een psychotherapeutische behandeling en zijn de verbeteringen stabiel over vijf jaar follow-up?
3) Worden er op de lange termijn verschillen gevonden aangaande de effectiviteit tussen ambulante, dagklinische en klinische behandelingen bij patiënten met een cluster A, B of C PS of PSNAO?

De effectiviteitsstudies die hieronder beschreven staan (hoofdstukken 2, 3 en 4), gebruiken allen de '*ernst van de psychische klachten (GSI)*' als de primaire uitkomstmaat. '*Psychosociaal functioneren (OQ-45)*' en '*aan gezondheid gerelateerde kwaliteit van leven (EQ-5D)*' vormen de secundaire uitkomstmaten. **Hoofdstuk 2** beschrijft de effectiviteit van verschillende behandelmodaliteiten bij patiënten met een cluster A, B en/of C PS over vijf jaar follow-up. De studies toonden met name aan dat een kortdurende klinische psychotherapie voor patiënten met een cluster C PS het meest effectief was. In de huidige studie zijn 519 patiënten die toegewezen waren aan ambulante, dagklinische of klinische behandelingen, gedurende vijf jaar gevolgd. Van deze patiënten voldoen 59 patiënten aan een cluster A PS, 210 aan een cluster B PS en 416 patiënten voldoen aan een cluster C PS. Aangezien patiënten kunnen voldoen aan de criteria voor meer dan een PS is de som van deze aantallen patiënten hoger dan het totale aantal patiënten. Bij patiënten met een cluster C PS worden naast de modaliteit ook de kortdurende behandelingen (< zes maanden) van de langdurige behandelingen (≥ zes maanden) onderscheiden. Ongecorrigeerde resultaten wijzen uit dat bijna alle patiëntgroepen na vijf jaar een significant lager niveau van klachten rapporteren en dat de initiële positieve uitkomsten over de vijf jaar van follow-up worden behouden. Alleen binnen de cluster A PS groep die een ambulante behandeling heeft gevolgd, wordt geen significante afname van klachten gevonden. Gecorrigeerde verschillen laten vooral kleine en niet-significante verschillen zien tussen de modaliteiten in de drie patiëntgroepen. Wanneer we naast verschillen tussen modaliteiten ook onderscheid maken tussen de duur van behandeling en zodoende langdurige ambulante behandelingen, kort- en langdurige dagklinische behandelingen en kort- en langdurige klinische behandelingen bij cluster C PS patiënten vergelijken, vinden we na vijf jaar dat kortdurende klinische psychotherapieën nog steeds grotere effecten behalen dan de andere behandelmodaliteiten. In **hoofdstuk 3** wordt het onderzoek naar de effectiviteit van verschillende modaliteiten voor patiënten met een PSNAO over vijf jaar follow-up gepresenteerd. Tweehonderdenvijf patiënten die waren toegewezen aan kort- en langdurige ambulante, kort- en langdurige dagklinische of kort- en langdurige klinische psychotherapie zijn met elkaar vergeleken. Op alle onderzochte tijdstippen wordt een positief effect gevonden van alle behandelmodaliteiten. Het grootste effect wordt gevonden op de primaire uitkomstmaat, de 'ernst van de psychische klachten', en op het '*functioneren in een sociale rol*'. Na een jaar en met correctie voor initiële verschillen tussen patiënten, blijken de kortdurende ambulante en de kortdurende klinische psychotherapie geassocieerd met de grootste verbetering met betrekking tot de '*ernst van de psychische klachten*' en zijn deze verschillen over het algemeen significant in vergelijking met de andere modaliteiten. Vijf jaar na baseline zijn de eerder geobserveerde verschillen tussen de modaliteiten voor het grootste gedeelte verdwenen. Dit is mogelijk te verklaren door de afname van effectiviteit in de kortdurende ambulante en de kortdurende klinische

psychotherapieën over de tijd heen, en een verdere toename van effectiviteit bij de langdurige ambulante en de langdurige klinische psychotherapieën. Op grond hiervan is de conclusie dat patiënten met een PSNAO zowel op de korte als de lange termijn profiteren van psychotherapie. Hoewel kortdurende ambulante en kortdurende klinische psychotherapieën na een jaar superieur lijken boven de andere behandelmodaliteiten, zijn deze verschillen in effectiviteit na vijf jaar verdwenen.

De hoofdstukken 4 en 5 beschrijven de effectiviteit en kosteneffectiviteit van de Kortdurende Klinische Psychotherapie gebaseerd op de Transactionele Analyse (STIP-TA) in vergelijking met andere gespecialiseerde psychotherapieën (Other Psychotherapies; OP). Voor dit onderzoek zijn 67 patiëntparen gevolgd met voornamelijk een cluster C PS en een PSNAO. Aangezien SCEPTRE een quasi-experimentele studie is, is de propensity score gebruikt om STIP-TA patiënten te matchen met OP patiënten. De effectiviteit en kosteneffectiviteit is over 36 maanden na baseline berekend. **Hoofdstuk 4** vergelijkt de effectiviteit van de STIP-TA behandeling met de OP behandeling. Zowel STIP-TA patiënten als ook OP patiënten laten grote symptomatische (‘*ernst van de psychische klachten’)* en functionele verbeteringen (‘*psychosociale functioneren*’ en ‘*aan gezondheid gerelateerde kwaliteit van leven’)* zien, waarbij STIP-TA patiënten echter op alle tijdstippen meer symptomatische verbeteringen laten zien vergeleken met OP patiënten. Na 36 maanden is 68% van de STIP-TA patiënten symptomatisch genezen, vergeleken met 48% van de OP patiënten. STIP-TA patiënten laten meer vooruitgang zien dan OP patiënten met betrekking tot verbeteringen in de ‘*ernst van de psychische klachten*’ en de ‘*aan gezondheid gerelateerde kwaliteit van leven*’. De superioriteit van STIP-TA is het best zichtbaar na 12 maanden van follow-up, maar blijft waarneembaar over de drie jaar van follow-up. In **hoofdstuk 5** is de kosteneffectiviteit van de STIP-TA behandeling vergeleken met de kosteneffectiviteit van de OP behandeling. Het maatschappelijk perspectief is hierin gebruikt. Dat betekent dat naast de kosten in de gezondheidszorg ook de kosten in verband met productieverlies in de berekeningen zijn meegenomen. De kosteneffectiviteit is uitgedrukt in de kosten per extra levensjaar in goede gezondheid (QALYs). De onzekerheid rondom de uitkomsten is gemeten door bootstrapping. STIP-TA patiënten bleken over de drie jaar follow-up minder kosten te generen dan OP patiënten. STIP-TA blijkt aldus goedkoper en effectiever dan OP, waardoor STIP-TA gezien kan worden als een dominante behandeling. We komen tot de conclusie dat STIP-TA een kosteneffectieve behandeling is voor veel patiënten met een cluster C PS en PSNAO.

Hoofdstuk 6 rapporteert over de matching hypothese dat patiënten die over veel psychologische vermogens beschikken meer profiteren van voornamelijk destabiliserende behandelingen, terwijl patiënten die over weinig psychologische vermogens beschikken meer profiteren van voornamelijk stabiliserende behandelingen. We hebben 735

patiënten met een PS gevolgd die zijn toegewezen aan verschillende levels van stabiliserende en destabiliserende psychotherapieën. De gebruikte uitkomstmaten zijn de '*ernst van de psychische klachten*' en het '*psychosociale functioneren*'. Na 12 maanden follow-up behalen patiënten die destabiliserende psychotherapieën hebben gevolgd enigszins betere uitkomsten dan patiënten die stabiliserende psychotherapieën hebben gevolgd. Patiënten die over veel psychologische vermogens beschikken, verbeteren over het algemeen meer dan patiënten die over weinig psychologische vermogens beschikken. Bovendien vonden we dat patiënten onafhankelijk van hun psychologische vermogens meer profiteren van een hogere mate van destabiliserende psychotherapieën. Deze resultaten duiden erop dat destabiliserende psychotherapieën zouden moeten worden overwogen als eerste behandeloptie bij PS patiënten, onafhankelijk van hun psychologische vermogens.

Hoofdstuk 7 bevat de algemene discussie waarin de resultaten van dit proefschrift zijn samengevat en waarin de implicaties voor de klinische praktijk en voor toekomstig onderzoek worden besproken. We concluderen dat psychotherapie bij patiënten met een PS langdurige behandelresultaten teweegbrengt maar dat functionele belemmeringen kunnen blijven bestaan. Bij patiënten met een cluster A of B PS blijken de verschillen tussen de modaliteiten of doseringen klein en over het algemeen niet significant op de lange termijn. Voor patiënten met een cluster C PS of PSNAO blijkt een kortdurende klinische behandeling (STIP-TA) de meest effectieve en kosteneffectieve behandeling te zijn, zowel op de korte als lange termijn. Dit specifieke behandelprogramma verdient dan ook meer klinische en wetenschappelijke aandacht. Vanuit een klinisch perspectief, zou STIP-TA een interessante optie zijn om in elke behandeltoewijzing te overwegen bij patiënten met een cluster C en PSNAO. Vanuit een wetenschappelijk perspectief kan het lonend zijn om de (kosten-) effectiviteit van STIP-TA te vergelijken met erkende en gerenommeerde, evidence-based behandelopties voor deze doelpopulatie. In het algemeen blijkt uit dit proefschrift de behoefte aan meer studies naar de vraag welke behandeling voor wie werkt, studies naar de effectieve bestanddelen van psychotherapie in PS patiënten en studies naar de mogelijke integratie van behandeltechnieken. De Sceptre studie heeft de kennis over zowel de effectiviteit als ook de kosteneffectiviteit van psychotherapie bij patiënten met een persoonlijkheidsstoornis vergroot en kan als voorbeeld dienen voor pragmatisch wetenschappelijk onderzoek.

PHD PORTFOLIO

Name	Eva Karin Horn
Scientific workplace	Viersprong Institute for Studies on Personality Disorders (VISPD)
Erasmus MC department	Department of Psychiatry, section Medical Psychology and Psychotherapy
Promotor	Prof. dr. J.J. van Busschbach
Co-promotor	Dr. R. Verheul

PhD training

2011	'Prognosis Research', Erasmus MC, Rotterdam
2010	'Repeated Measurements in clinical studies', Erasmus Winter Program, Rotterdam
2009	'Courses for the Quantitative Researcher', Erasmus MC, Rotterdam
	'Repeated Measures', Boerhaave, LUmc, Leiden
2007	'Health Technology Assessment', Erasmus MC, Rotterdam
	'Writing English for publication', VUmc, Amsterdam
	'Projectmanagement voor projectleiders', interne training, Trimbos-instituut, Utrecht
2006	'Implementatie van een evidence-based behandeling voor patiënten met een depressieve stoornis', Impact Seattle, Utrecht
	'Doelmatigheidsonderzoek: methoden en principes', VUmc, Amsterdam

Clinical training

2015-2017	Training 'Health Care Psychologist', RINO Zuid, Eindhoven (expected)
2015	'Schemafocused Therapy', Basic course Viersprong Academy, Amsterdam
	'Praktijkdag TPO', Viersprong Academy, Bergen op Zoom
2014	'Attachment theory and the Rorschach (R-PAS). Empirical and clinical implications.' Nederlands Vlaamse Vereniging voor Rorschach en projectieve methoden (NVVR), Den Bosch
2013	'Masterclass: Rorschach Performance Assessment System. Vernieuwingen en verschillen met het Comprehensive System van Exner'. Nederlands Vlaamse Vereniging voor Rorschach en projectieve methoden (NVVR), Utrecht
2012	'TPO bij kinderen, jongeren en hun gezin', Viersprong Academy, Halsteren

2010 'Therapeutic Assessment of an adolescent and her family', Universita Cattolica del Sacro Cuore, Milaan

'Therapeutic use of the Rorschach and other projective tests', Rorschachvereniging, Utrecht

2006 'Problem Solving Treatment', Impact Seattle, Utrecht

2005 'Training SCID I + II-interview', GGNet, Apeldoorn

National presentations

2013 'Cost-effectiveness of Short-Term Inpatient Psychotherapy, based on Transactional Analysis (STIP-TA)', Lowlands Health Economics' Study Group (IolaHESG)

2009 'Effectiveness of STIP-TA', de Viersprong, Halsteren

2007 'De behandeling van depressieve patiënten volgens een collaborative care model met PST en zelfhulp', poster gepresenteerd op de Geestkracht Conferentie (ZonMW).

Teaching activities

2012 'Onbegrepen Pijn', skills training voor medical students, Erasmus MC

Reviewing papers

2015 British Journal of Psychiatry

PUBLICATIONS

Horn EK, Verheul R, Thunnissen M, Delimon J, Goorden M, Hakkaart-van Roijen L, Soons M, Meerman AM, Ziegler UM, Rossum BV, Stijnen T, Emmelkamp PM, Busschbach JJ. Effectiveness of psychotherapy in PD. *Under review.*

Horn EK, Verheul R, Thunnissen M, Delimon J, Goorden M, Hakkaart-van Roijen L, Soons M, Meerman AM, Ziegler UM, Rossum BV, Stijnen T, Emmelkamp PM, Busschbach JJ (2015). Cost-Effectiveness of Short-term Inpatient Psychotherapy based on Transactional Analysis in Patients with Personality Disorders. *Journal of Personality Disorders*, Epub ahead of print.

Van Manen JG, Horn EK, Stijnen T, Timman R, Busschbach JJ, Verheul R (2015). Tailoring psychotherapy in patients with personality disorders: matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health*, 9(2), 133-149.

Horn EK, Verheul R, Thunnissen M, Delimon J, Soons M, Meerman AM, Ziegler UM, Rossum BV, Andrea H, Stijnen T, Emmelkamp PM, Busschbach JJ (2015). Effectiveness of Short-Term Inpatient Psychotherapy based on Transactional Analysis in Patients with Personality Disorders: A Matched Control Study using Propensity Score. *Journal of Personality Disorders* 29(5), 663-683.

van Steenbergen-Weijenburg KM, van der Feltz-Cornelis CM, van Benthem TB, Horn EK, Ploeger R, Brals JW, Leue C, Spijker J, Hakkaart-van Roijen L, Rutten FFH, Beekman ATF (2015). [Collaborative care for co-morbid major depressive disorder in chronically ill outpatients in a general Hospital]. *Tijdschrift voor Psychiatrie* 57(4), 248-257.

Horn EK, Bartak A, Meerman AM, Rossum BV, Ziegler UM, Thunnissen M, Soons M, Andrea H, Hamers EF, Emmelkamp PM, Stijnen T, Busschbach JJ, Verheul R (2015). Effectiveness of Psychotherapy in Personality Disorders Not Otherwise Specified: A Comparison of Different Treatment Modalities. *Clinical Psychology and Psychotherapy*, 22(5), 426-442.

De Saeger H, Kamphuis JH, Finn SE, Smith JD, Verheul R, van Busschbach JJ, Feenstra DJ, Horn EK (2014). Therapeutic assessment promotes treatment readiness but does not affect symptom change in patients with personality disorders: findings from a randomized clinical trial. *Psychological Assessment*, 26(2):474-83.

Birtchnell J, Hammond S, Horn E, De Jong C, Kalaitzaki A (2013). A cross-national comparison of a shorter version of the person's relating to others questionnaire. *Clinical Psychology and Psychotherapy*; 20(1):36-48.

van Steenbergen-Weijenburg KM, van Puffelen AL, Horn EK, Nuyen J, van Dam PS, van Benthem TB, Beekman AT, Rutten FF, Hakkaart-van Roijen L, van der Feltz-Cornelis CM (2011). More comorbid depression in patients with Type 2 diabetes with multiple complications. An observational study at a specialized outpatient clinic. *Diabetic Medicine* 28(1):86-9.

van Steenbergen-Weijenburg KM, van der Feltz-Cornelis CM, Horn EK, van Marwijk HW, Beekman AT, Rutten FF, Hakkaart-van Roijen L (2010). Cost-effectiveness of collaborative care for the treatment of major depressive disorder in primary care. A systematic review. *BMC Health Services Research* 10, 19.

Horn EK, van Benthem TB, Hakkaart-van Roijen L, van Marwijk HW, Beekman AT, Rutten FF, van der Feltz-Cornelis CM (2007). Cost-effectiveness of collaborative care for chronically ill patients with comorbid depressive disorder in the general hospital setting, a randomised controlled trial. BMC Health Services Research 7, 28.

CURRICULUM VITAE

Eva Horn was born on April 21st, 1981 in Saarbrücken, Germany. In 2000, she graduated from high school at the Geschwister Scholl College in Ratingen. In the same year she started her German and English studies at the University of Düsseldorf. In 2001, she started her Psychology study in Nijmegen, the Netherlands. She conducted research on three models measuring interpersonal behaviour in cooperation with the Institute of Psychiatry, London. She furthermore completed a clinical internship at a mental health care centre for personality disorders in Apeldoorn (Spatie, now called GGNet). She obtained her master's degree with merit in 2006, having specialized in clinical psychology. In the same year, Eva started a PhD-project on the cost-effectiveness of collaborative care in medically ill patients with a major depressive disorder at the Trimbos Institute, Utrecht in collaboration with the Onze Lieve Vrouwe Gasthuis (OLVG), Amsterdam. In 2008, Eva switched to start working at the department of medical psychology at the OLVG. In 2009, Eva started working at de Viersprong, a national centre specialized in the assessment and treatment of adolescents and adults with personality problems. She worked at de Viersprong Institute for Studies on Personality Disorders (VISPD) on the PhD project Study on Cost-Effectiveness of Personality Disorders Treatment (SCEPTRE). The results of this project are described in this thesis. During her PhD studies, she also worked as a psychologist at the diagnostics department of de Viersprong. After finishing her PhD project, in 2015 she started a training to become a health care psychologist (Gezondheidszorgpsycholoog).