UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Towards optimal treatment for borderline personality disorder

Comparative effectiveness, treatment selection, and assessment

Wibbelink, C.J.M.

Publication date
2026

Link to publication

Citation for published version (APA):
Wibbelink, C. J. M. (2026). *Towards optimal treatment for borderline personality disorder: Comparative effectiveness, treatment selection, and assessment*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:10 Nov 2026

Chapter 2

Towards optimal treatment selection for borderline personality disorder patients (BOOTS): A study protocol for a multicenter randomized clinical trial comparing Schema Therapy and Dialectical Behavior Therapy

Carlijn J. M. Wibbelink
Arnoud Arntz
Raoul P. P. P. Grasman
Roland Sinnaeve
Michiel Boog
Odile M. C. Bremer
Eliane C. P. Dek
Sevinç G. Alkan
Chrissy James
Annemieke M. Koppeschaar
Linda Kramer
Maria Ploegmakers
Arita Schaling
Faye I. Smits
Jan H. Kamphuis

This chapter is published as:

Wibbelink, C. J. M., Arntz, A., Grasman, R. P. P. P., Sinnaeve, R., Boog, M., Bremer, O. M. C., Dek, E. C. P., Alkan, S. G., James, C., Koppeschaar, A. M., Kramer, L., Ploegmakers, M., Schaling, A., Smits, F. I., & Kamphuis, J. H. (2022). Towards optimal treatment selection for borderline personality disorder patients (BOOTS): a study protocol for a multicenter randomized clinical trial comparing schema therapy and dialectical behavior therapy. *BMC Psychiatry, 22*(1), 89. https://doi.org/10.1186/s12888-021-03670-9

ABSTRACT

Background: Specialized evidence-based treatments have been developed and evaluated for borderline personality disorder (BPD), including Dialectical Behavior Therapy (DBT) and Schema Therapy (ST). Individual differences in treatment response to both ST and DBT have been observed across studies, but the factors driving these differences are largely unknown. Understanding which treatment works best for whom and why remains a central issue in psychotherapy research. The aim of the present study is to improve treatment response of DBT and ST for BPD patients by a) identifying patient characteristics that predict (differential) treatment response (i.e., treatment selection) and b) understanding how both treatments lead to change (i.e., mechanisms of change). Moreover, the clinical effectiveness and cost-effectiveness of DBT and ST will be evaluated.

Methods: The BOOTS trial is a multicenter randomized clinical trial conducted in a routine clinical setting in several outpatient clinics in the Netherlands. We aim to recruit 200 participants who will be randomized to DBT or ST. Patients receive a combined program of individual and group sessions for a maximum duration of 25 months. Data are collected at baseline until 3-year follow-up. Candidate predictors of (differential) treatment response have been selected based on the literature, a patient representative from the Borderline Foundation of the Netherlands, and semi-structured interviews with 18 expert clinicians. In addition, BPD-treatment-specific (ST: beliefs and schema modes; DBT: emotion regulation and skills use), BPD-treatment-generic (therapeutic environment characterized by genuineness, safety, and equality), and non-specific (attachment and therapeutic alliance) mechanisms of change are assessed. The primary outcome measure is change in BPD manifestations. Secondary outcome measures include functioning, additional self-reported symptoms, and well-being.

Discussion: The current study contributes to the optimization of treatments for BPD patients by extending our knowledge of "Which treatment, DBT or ST, works best for which BPD patient, and why?", which is likely to yield important benefits for both BPD patients (e.g., prevention of overtreatment and potential harm of treatments) and society (e.g., increased economic productivity of patients and efficient use of treatments).

INTRODUCTION

Borderline personality disorder (BPD) is a complex and severe mental disorder, characterized by a pervasive pattern of instability in emotion regulation, self-image, interpersonal relationships, and impulse control (American Psychiatric Association, 2013; Skodol et al., 2002). The prevalence in the general population is estimated to be between 1% and 3% (Ellison et al., 2018; ten Have et al., 2016; Tomko et al., 2014), and 10%–25% among psychiatric outpatients and inpatients (Ellison et al., 2018). BPD is associated with severe functional impairment, high rates of comorbid mental disorders, and physical health problems (Dixon-Gordon et al., 2015; El-Gabalawy et al., 2010; Tomko et al., 2014). In addition, BPD is characterized by low quality of life, lower than that observed in other common mental disorders such as depressive disorder, and comparable to that of patients with severe physical conditions, such as Parkinson's disease and stroke (Laurenssen et al., 2016). Moreover, BPD is related to a high risk of suicide (3–6%, or even up to 10%; Paris, 2019; Pompili et al., 2005) and suicide attempts or threats (up to 84%; Black et al., 2004; Soloff & Chiappetta, 2019), and an increased mortality rate (Kjaer et al., 2020). Besides the detrimental effects of BPD on the individual patient, BPD also poses a substantial financial burden on society. BPD patients make extensive use of treatment services, which results in markedly higher healthcare costs for people with BPD compared to people with other mental disorders, such as other personality disorders (Bender et al., 2001) and depressive disorder (Bode et al., 2017). BPD is also associated with high non-healthcare costs, including costs related to productivity losses, informal care, and out-of-pocket costs (van Asselt et al., 2007; Wagner et al., 2014).

Interventions: Dialectical Behavior Therapy and Schema Therapy

BPD has traditionally been viewed as one of the most difficult mental disorders to treat (Sheehan et al., 2016). During recent years, a number of promising treatments have been developed and evaluated, including Dialectical Behavior Therapy (DBT) (Linehan, 1993a, 2015) and Schema Therapy (ST) (Arntz et al., 2009; Young et al., 2003). DBT is a comprehensive cognitive behavioral treatment for BPD, rooted in behaviorism, Zen, and dialectical philosophy (Linehan, 1993a). ST is based on an integrative cognitive therapy, combining cognitive behavior and experiential therapy techniques with concepts derived from developmental theories, including attachment theory, and psychodynamic concepts (Jacob & Arntz, 2013). For detailed information about these treatments, the reader is referred to the Methods/design section.

Several studies have demonstrated the effectiveness and the efficacy of DBT and ST for BPD, although the evidence is mostly based on low-to-moderate-quality evidence and trials focusing on DBT, but especially ST, are limited (Sempertegui et al., 2013; Storebø et al., 2020). In addition, substantial reductions in direct and indirect healthcare costs have been found for both treatments (Meuldijk et al., 2017). However, research on the comparative

effectiveness and cost-effectiveness of the two interventions is lacking. Moreover, research on mediators and moderators of treatment effects is limited. This gap warrants attention, as treatment effectiveness can be optimized by identifying mechanisms within treatments that are associated with improvement and patient characteristics that predict (differential) treatment response (Huibers et al., 2021). Optimizing treatment effectiveness of DBT and ST for BPD is urgently needed since a substantial proportion of patients does not respond fully to either DBT or ST. A systematic review found a mean non-response rate of 46% among BPD patients treated with specialized psychotherapies, including DBT and ST (Woodbridge et al., 2021). In addition, more than one-third of the patients did not achieve a reliable change in BPD symptoms or even showed an increase in BPD severity after DBT or ST (Barnicot, 2020; Giesen-Bloo et al., 2006; McMain et al., 2012). Finally, dropout rates of up to 30% have been found in the first year of DBT and ST (Arntz et al., 2021; Dixon & Linardon, 2020). Individual differences in responses to both ST and DBT have been observed across studies, but the factors driving these differences in treatment response among BPD patients are largely unknown. This state of affairs leaves the principal question "What treatment, by whom, is most effective for this individual with that specific problem, under which set of circumstances?" (Meehl, 1954, p. 111), historically one of the key questions dominating the psychotherapy research agenda, fully open in the treatment of individuals with BPD (DeRubeis et al., 2014; Kazdin, 2014). Identifying factors that specify which patients will benefit most from which treatment (i.e., treatment selection, also referred to as precision medicine or personalized medicine; Cohen & DeRubeis, 2018; Hamburg & Collins, 2010) will lead to fewer mismatches between patients and treatments, and in turn to better outcomes and more efficient use of healthcare resources.

Treatment selection

Several factors predicting treatment response irrespective of type of treatment (i.e., prognostic factors; DeRubeis et al., 2014) among BPD patients have been reported in the literature. The overwhelming list of candidate variables and the general lack of replication hampers research on prognostic factors in BPD (Arntz et al., 2015). Research among BPD patients on prescriptive factors (i.e., factors that predict different outcomes depending on the treatment; moderators) is very scarce indeed. Arntz et al. (2015) examined the effect of several potential predictors of (differential) treatment response across ST and Transference Focused Psychotherapy among BPD patients. The authors failed to find prescriptive factors, but it should be noted that the sample size was inadequate to detect subtle differences between treatments. In addition, Verheul et al. (2003) found that patients with a high frequency of self-mutilating behavior before treatment were more likely to benefit from DBT compared to treatment as usual, whereas for patients with a low frequency of self-mutilating behavior effectiveness did not differ.

Historically, research has focused on single variables to predict treatment response, but often failed to find consistent and clinically meaningful moderators (Bohart & Wade, 2013; Clarkin & Levy, 2004; Wallace et al., 2013; Wampold & Imel, 2015). However, it is highly unlikely that a single variable is responsible for differences in treatment response (Keefe et al., 2020; Wallace et al., 2013; Zilcha-Mano, 2019). In recent decades, novel approaches combining multiple predictors to determine the optimal treatment for a particular patient have been introduced, including the methods of Kraemer (2013; optimal composite moderator) and DeRubeis and colleagues (2014; statistically derived selection algorithm). Several studies have found that a combination of predictors predicted differential treatment response (e.g., Cohen et al., 2020; Huibers et al., 2021; van Bronswijk et al., 2021). For example, by using the method of DeRubeis and colleagues, it was investigated in an effectiveness study among BPD patients which of two different treatments (DBT and General Psychiatric Management; GPM) would have been the optimal treatment option for a particular patient in terms of long-term outcome (Keefe et al., 2020). The authors found that BPD patients with childhood emotional abuse, social adjustment problems, and dependent personality traits were more likely to benefit from DBT compared to GPM, whereas GPM excelled for patients with more severe problems related to impulsivity. The authors also provided an estimate of the advantage that might be gained if patients had been allocated to the optimal treatment option. The average difference in outcomes between the predicted optimal treatment and non-optimal treatment for all patients was small-to-medium (d = 0.36), while the advantage for patients with a relatively stronger prediction increased to a medium-to-large effect (d = 0.61). This suggests that treatment allocation based on a treatment selection procedure may substantially improve outcomes for BPD patients.

Mechanisms of change

Another principal way to improve treatment response is to capitalize on mechanisms underlying change in treatments (Huibers et al., 2021; Kazdin & Nock, 2003; Keefe et al., 2020; Kraemer et al., 2002). Studying mechanisms of change helps to identify core ingredients of interventions and points the way to enhancing crucial elements, while discarding redundant elements. Presumably, this would maximize (cost-)effectiveness and efficiency as well. Since the 1950s, research on change processes has increased exponentially (Orlinsky et al., 2004). However, the majority of the trials on BPD have focused on outcomes, and only a few addressed *how* treatments exerted a positive effect on patient outcomes (Forster et al., 2014; Lynch et al., 2006). Rudge et al. (2020) reviewed studies on mechanisms of change in DBT. They concluded that there is empirical support for behavioral control, emotion regulation, and skills use as mechanisms underlying change in DBT. Recently, Yakın et al. (2020) examined schema modes as mechanisms of change in ST for cluster C, histrionic, paranoid, and narcissistic personality disorders. They found that strengthening a functional schema mode (i.e., healthy adult mode) and weakening four maladaptive schema modes (i.e., vulnerable child mode, impulsive child

mode, avoidant protector mode, and self-aggrandizer mode) predicted improvements in PD symptomatology. However, changes in these schema modes, except for self-aggrandizer mode, also predicted improvements in outcome in treatment-as-usual and clarification-oriented psychotherapy, suggesting that modifying the strength of schema modes might reflect common mechanisms of change. The question of specificity of mechanisms of change is interesting, especially since both DBT and ST have their roots in cognitive behavioral therapy and show similarity in certain treatment parameters, but differ substantially in techniques, explanatory model, and terminology (Fassbinder et al., 2016). Clarifying the treatment-specific and non-specific mechanisms of change may be key to furthering the effectiveness of both DBT and ST, and potentially also for psychotherapy in general.

Current study

BPD-tailored treatments, like DBT and ST, are considered treatments of choice for BPD (Storebø et al., 2020). However, knowledge of the comparative (cost-)effectiveness of DBT and ST is lacking, as is knowledge on mechanisms of change and patient characteristics that predict (differential) treatment response. We will therefore perform a multicenter randomized clinical trial (RCT) comparing DBT and ST for BPD patients to elucidate the question "Which treatment, DBT or ST, works best for which BPD patient, and why?". The main aim of the BOOTS (Borderline Optimal Treatment Selection) study is to improve treatment response of DBT and ST for BPD patients by optimizing treatment selection through the identification of a prediction model based on patient characteristics that predict (differential) treatment response. By doing so, this study is a first step toward the development of a treatment selection procedure for BPD patients. Moreover, the results of this study can serve as a starting point for future studies with the ultimate goal of implementing a treatment selection procedure that can be used in clinical practice to guide BPD patients and clinicians in selecting the optimal treatment. In addition, we aim to elucidate the mechanisms by which DBT and ST lead to change, thus pursuing the other main avenue towards improving BPD treatments.

This study has four primary objectives. The first objective of this study is to develop a treatment selection model based on a combination of patient characteristics that predict (differential) treatment response across DBT and ST. Candidate predictors of (differential) treatment response have been selected based on the literature, suggestions of a patient representative from the Borderline Foundation of the Netherlands, and clinicians' appraisals of BPD patient characteristics that predict (differential) treatment response across DBT and ST. Semi-structured interviews were conducted with 18 expert clinicians to identify patient characteristics they deemed predictive of (differential) treatment response. The extensive investment in the identification of pertinent predictors is a lesson learned from Meehl (1954), who noted that actuarial methods will not outperform clinical judgment when they are based on inadequate knowledge of relevant variables. According to Westen and Weinberger (2004), clinical expertise can serve the important function of identifying relevant variables for use

in research. In addition, the majority of studies examining predictors of treatment response are based on randomized controlled trials with a primary focus on treatment effectiveness (Kessler, 2018), which could result in the exclusion of potentially relevant predictors due to the lack of instruments assessing these constructs (Arntz et al., 2015; Spinhoven et al., 2008). Moreover, findings in the literature may be affected by publication bias, since statistically significant predictors of treatment response are more likely to be published (Zilcha-Mano, 2019). Therefore, candidate predictors of (differential) treatment response are not only based on the literature, but also on clinical expertise and experience-based knowledge. We hypothesize that a combination of multiple patient characteristics will predict and moderate treatment effectiveness of DBT and ST. Hypotheses on the effects of single patient characteristics will not be formulated as research among BPD patients often failed to find consistent prognostic factors, while research on prescriptive factors or a combination of factors is scarce. In addition, there was generally a lack of consensus between the 18 expert clinicians on patient characteristics predicting (differential) treatment response across DBT and ST.

Second, we aim to elucidate how DBT and ST exert their effects by gaining a better understanding of the mechanisms of change of DBT and ST. A first step towards more insight into mechanisms of change is the identification of mediators. Mediators are easily confused with mechanisms of change, despite important differences (Tryon, 2018). A mediator is an intervening variable (partly) accounting for the statistical relationship between the intervention and outcome, and might serve as a statistical proxy for a mechanism of change (Holmes et al., 2018). In this study, we will examine potential BPD-treatment-specific, BPD-treatment-generic, and non-specific mediators. Based on empirical research and the presumed mechanisms of change (e.g., Lynch et al., 2006; Rudge et al., 2020; Yakin et al., 2020), we hypothesize that change in skills use and emotion regulation are the mechanisms underlying change in DBT, and that change in schema modes and beliefs are the mechanisms of change in ST (i.e., BPD-treatment-specific mechanisms of change). In addition, a therapeutic environment characterized by genuineness of therapists and group members, safety, and equality is considered to be especially important for BPD treatment (Farrell & Shaw, 2010; Gunderson, 2008; Linehan, 1997; Swenson & Choi-Kain, 2015) and is therefore assumed to be a BPD-treatment-generic mechanism of change. Finally, attachment and therapeutic alliance are the presumed non-specific mechanisms of change (Barnicot et al., 2012; Taylor et al., 2015).

Third, the comparative effectiveness of DBT and ST will be examined. Accumulating evidence suggests that symptoms and psychosocial functioning are only loosely associated (Gunderson et al., 2018; Skodol, 2018). Patients with BPD are characterized by significant impairments in vocational functioning, relationships, and leisure (Skodol et al., 2002). In addition, social adjustment of BPD patients is considerably lower than social adjustment seen in other mental disorders, such as major depressive disorder and bipolar I disorder (Gunderson et al., 2011). Moreover, although several studies found that even as psychopathology after treatment of BPD decreased, impairments in quality of life and functioning often (partly)

persist (Chakhssi et al., 2019; Zanarini et al., 2010). A more comprehensive view of recovery is therefore needed. This notion is underscored by qualitative research that has shown that patients define recovery in terms of personal well-being, social inclusion, and satisfaction with life (Katsakou et al., 2012; Lariviere et al., 2015). Therefore, the current trial will track outcomes in multiple domains including symptoms, functioning, and well-being.

Finally, the cost-effectiveness of DBT and ST will be compared. Individual ST seems to be a cost-effective treatment (Bamelis et al., 2015; van Asselt et al., 2008). However, although group ST combined with individual ST is widely used in clinical practice, the cost-effectiveness of this combined program is yet unknown. An international RCT evaluating the (cost-)effectiveness of group ST for BPD is currently in progress (Wetzelaer et al., 2014). More economic evaluations of DBT are available and support the cost-effectiveness of DBT. However, the studies vary considerably in design and the number of trials is still somewhat limited (Brettschneider et al., 2014; Krawitz & Miga, 2019; Meuldijk et al., 2017). Therefore, an economic evaluation will be performed and a societal perspective will be applied, including indirect and direct healthcare costs.

In addition to these primary objectives, several secondary investigations will be performed, including, but not limited to: (1) the heterogeneity of BPD, (2) substance use (disorders) among patients with BPD, (3) perspectives of patients and therapists in key areas, including predictors, mechanisms of change, the treatments, and the implementation of the results in clinical practice, and (4) psychometric evaluations of several Dutch questionnaires (e.g., Dialectical Behavior Therapy-Ways of Coping Checklist, Ultrashort BPD Checklist).

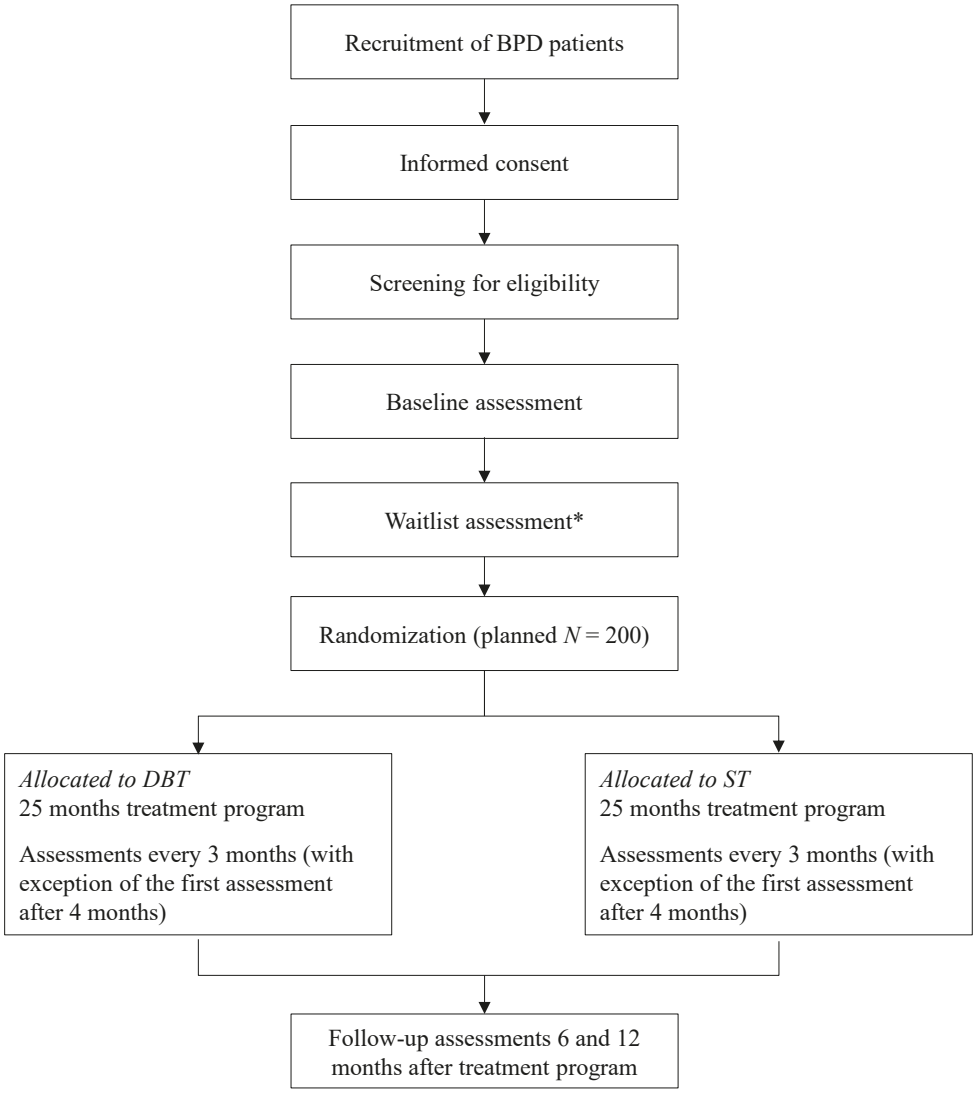
METHODS/DESIGN

Design

The study is a multicenter RCT with two active conditions (DBT or ST). The study is set at various Dutch mental healthcare centers accessible through the public health system, including Antes (Rotterdam), GGZ inGeest (Amsterdam), GGZ NHN (Heerhugowaard), GGZ Rivierduinen (Leiden), NPI (Amsterdam), Pro Persona (Ede and Tiel), PsyQ (Rotterdam-Kralingen), and PsyQ/i-psy (Amsterdam). For an overview of the study design, including the enrollment, randomization, interventions, and assessments, see Figure 1.

The Medical Ethics Committee of the Academic Medical Center (MEC-AMC) Amsterdam approved the study protocol (registration number NL66731.018.18). The study is registered at the Netherlands Trial Register, part of the Dutch Cochrane Center (registration number NL7699), and complies with the World Health Organization Trial Registration Data Set. Modifications to the protocol require a formal amendment to the protocol, which will be examined by the MEC-AMC. The trial adheres to the SPIRIT methodology and guidelines (Chan et al., 2013).

Figure 1. *Flow Chart of the Study Design*



Note. DBT = Dialectical Behavior Therapy; ST = Schema Therapy.

*An additional assessment after the waiting period is included for patients with a waitlist period exceeding 3 months after the baseline assessment.

Patients

Patients are eligible if they (1) are between 18 and 65 years old, (2) have a primary diagnosis of BPD (diagnosed with the Structured Clinical Interview for DSM-5 Personality Disorders; SCID-5-PD), (3) have a BPD severity score > 20 on the Borderline Personality Disorder Severity Index, version 5 (BPDSI-5), (4) have adequate proficiency in the Dutch language, and (5) are

motivated to participate in (group) treatment for a maximum of 25 months and are willing and able to complete the assessments over a period of 3 years. Patients will be excluded if they (1) fulfill the criteria for a psychotic disorder in the past year (diagnosed with the Structured Clinical Interview for DSM-5 Syndrome Disorders; SCID-5-S), (2) have current substance dependence needing clinical detoxification, (3) have been diagnosed with bipolar I disorder with at least one manic episode in the past year, (4) have been diagnosed with antisocial personality disorder (diagnosed with the SCID-5-PD), in combination with a history of physical violence against multiple individuals in the past 2 years, (5) have an IQ below 80, (6) have a travel time to the mental healthcare center longer than 45 minutes (except when the patient lives in the same city), (7) have no fixed address, and (8) have received ST or DBT in the past year.

Sample size

We aim to include 200 participants. Each center intends to recruit at least 18 patients. For the power analysis, we adopted the minimal statistically detectable effect approach (Lakens, 2021). A sample size of 200 will be sufficient to have 80% power to detect moderators of treatment effects that have an effect size of Cohen's f of 0.20 (small to medium effect size), based on a two-tailed significance level of $p < .05$. In addition, the study has 80% power to detect medium effect-sized (i.e., Cohen's $f = 0.25$) moderators of treatment effects, based on a two-tailed significance level of $p < .01$.

Regarding the effectiveness study, with a sample size of $N = 200$ the study is powered at 82% to detect a group difference with a medium effect size of Cohen's $d = 0.50$ at a two-tailed significance level of $p < .05$ and assuming a model with center as random effect and an intraclass correlation value of 0.05 corresponding to the center by treatment interaction (Adams et al., 2004; Eldridge et al., 2004).

Finally, a sample size of $N = 200$ will be sufficient to have 98% power to detect a medium effect size of the mediation effect ($rr = .09$; Kenny & Judd, 2014; Kenny, 2016, 2017), assuming path a (relation between the predictor and mediator) and path b (relation between the mediator and outcome measure) both have a medium effect size ($r = .30$), and based on a simplified trivariate mediation model (Schoemann et al., 2017).

Recruitment

Patients are recruited in the respective participating mental healthcare centers. Patients diagnosed with BPD or for whom this is deemed likely are invited to participate in the screening process. After reading and hearing information about the study and signing an informed consent, patients will start with the screening process. Not only new referrals can be included, but also patients who are already receiving treatment for mental disorders (except patients receiving ST or DBT).

Randomization

A central independent research assistant randomizes the patients per center after a final check of the inclusion and exclusion criteria, and after all baseline measures have been completed. Generally, patients will be randomized using computerized covariate adaptive randomization (Pocock & Simon, 1975; Scott et al., 2002; Taves, 1974), taking into account gender and severity of BPD (BPDSI-5 score ≤ 24; BPDSI-5 score > 24). By using this method, the imbalance of baseline characteristics between the treatments will be minimized. Patients are allocated to the treatment group that results in the least imbalance between the treatments with an allocation probability of 0.8 to preserve unpredictability (Altman & Schulz, 2001). Groups in both treatments are semi-open which implies that new patients can enter the group if treatment slots are available. Therefore, treatment capacity will be taken into account by using unequal ratios if needed (e.g., 2:1 or 1:3).

In exceptional cases, an alternative randomization method will be used if one or more treatment slots are available in only one condition and there is no available treatment slot in the other condition. To prevent long waiting times for treatment and empty places in the groups, the available treatment slot(s) in one condition will be randomized over 2**k* patients whereby *k* stands for the number of available treatment slots, and randomization is done in the subsample of *k* patients who wait the longest. Randomization over 2**k* patients guarantees unpredictable outcomes. For example, if one treatment slot is available in DBT and there is no available treatment slot in ST at that moment, nor within the foreseeable future, the available treatment slot in DBT will be randomized over two patients waiting for treatment. Sensitivity analyses will be performed by excluding patients who have been randomized using the alternative randomization method.

Procedure and assessments

Patients with BPD or suspected of BPD are invited to the screening process by the research assistant or intake staff member. After providing written informed consent, patients are assessed for eligibility to participate in the study based on the inclusion and exclusion criteria. First, to assess DSM-5 syndrome disorders, the SCID-5-S is administered. The SCID-5-PD will also be administered in case the SCID-5-PD is not part of the standard intake procedure of the mental healthcare center. Second, the BPDSI-5 and a screening interview to assess the motivation and availability of the patient are conducted. A simple "yes" answer to the questions posed by the interviewer (e.g., "Are you motivated and available for treatment, including individual and group sessions?") is not sufficient. Patients need to elaborate on their answers and follow-up questions are asked if needed. Patients who are eligible for participation will be invited for the baseline assessment, including interviews and computer-based self-report questionnaires, and intake staff members will fill out a questionnaire (i.e., intake questionnaire; see the Measures section) about these patients. After completing the baseline assessment, patients will be randomized as soon as treatment slots become available.

Patients will be informed that they have been allocated to one of the treatment conditions, but the name of the treatment will not be communicated to the patient until the first treatment session. If patients cannot be randomized within several months after completing the baseline assessment because of unavailability of treatment slots, the BPDSI-5 will be re-assessed after 3 months and the BPDSI-5 and cost interview will be re-assessed after 6 months.

After the treatment phase has started, patients are reassessed every 6 months during the 2 years of treatment. These assessments are a combination of interviews and computer-based self-report questionnaires. In addition, a selection of measures are also assessed every 3 months, by computer-based self-report questionnaires. After the end of treatment, two follow-up assessments (6 and 12 months after the end of treatment) will be administered. An overview of the measures is presented in Table 1. Candidate predictors of (differential) treatment response that are assessed only once at baseline are not included in Table 1. These measures can be found in the Measures section.

Table 1. *Overview of Instruments*

	Before treatment		During treatment								Follow-up	
Timepoint	**S**	**B[a]**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**6FU**	**12FU**
Interventions												
DBT												
Treatment phase			◆	———		◆						
Maintenance phase							◆	———		◆		
ST												
Treatment phase			◆	———		———		◆				
Maintenance phase									◆	———		
Assessments												
Screening												
SCID-5	●											●
Motivation interview	●											
Primary outcome												
BPDSI-5	●			●		●		●		●	●	●
Secondary outcomes												
BSI		●		●		●		●		●	●	●
Cost interview		●		●		●		●		●	●	●
Demographic interview		●		●[b]		●[b]		●[b]		●[b]	●[b]	●[b]
EQ-5D-5L		●		●		●		●		●	●	●
Happiness item		●	●	●	●	●	●	●	●	●	●	●
ISI		●		●		●		●		●	●	●
MHQoL-7D		●		●		●		●		●	●	●

Table 1. *Continued*

	Before treatment		During treatment								Follow-up	
Timepoint	**S**	**B[a]**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**6FU**	**12FU**
NFQ		●		●		●		●		●	●	●
ORS		●	●	●	●	●	●	●	●	●	●	●
Ultrashort BPD Checklist		●	●	●	●	●	●	●	●	●	●	●
WHODAS 2.0		●		●		●		●		●	●	●
Mediators												
Beliefs[c]		●	●	●	●	●	●	●	●	●	●	●
DBT-WCCL		●	●	●	●	●	●	●	●	●	●	●
DERS-18[d]		●		●		●		●		●	●	●
DERS-SF[e]		●		●		●		●		●	●	●
ECR-RS[f]		●[g]		●		●		●		●	●	●
SMI		●	●[h]	●	●[h]	●	●[h]	●	●[h]	●	●	●
Therapeutic environment[i]			●	●	●	●	●	●	●	●	●	●
WAI-S		●[j]		●		●		●		●	●	●

Note. S = screening; B = baseline assessment; 3/6/9/12/15/18/21/24 = 3/6/9/12/15/18/21/24 months after start of the treatment phase; 6 FU = follow-up at 6 months after end of treatment; 12 FU = follow-up at 12 months after end of treatment. BPDSI-5 = Borderline Personality Severity Interview, fifth version; BPD = borderline personality disorder; BSI = Brief Symptom Inventory; DBT-WCCL = Dialectical Behavior Therapy-Ways of Coping Checklist; DERS-18 = Difficulties in Emotion Regulation Scale 18; DERS-SF = Difficulties in Emotion Regulation Scale Short Form; ECR-RS = Experiences in Close Relationships-Relationship Structures questionnaire; EQ-5D-5L = 5-level EuroQol 5D version; ISI = Insomnia Severity Index; MHQoL-7D = Mental Health Quality of Life seven-dimensional Questionnaire; NFQ = Nightmare Frequency Questionnaire; ORS = Outcome Rating Scale; SCID-5 = Structured Clinical Interview for the DSM-5 Disorders; SMI = Schema Mode Inventory; WAI-S = Working Alliance Inventory-Short; WHODAS 2.0 = World Health Organization Disability Assessment Schedule 2.0.

[a] The potential predictors of (differential) treatment response that are assessed only once at baseline are not included in this table.

[b] A shortened version.

[c] Including three to five idiosyncratic dysfunctional beliefs and one functional belief.

[d] DERS-18 subscale 'Awareness'.

[e] Excluding the subscale 'Awareness'.

[f] Three versions of the ECR-RS will be assessed, measuring general attachment style and attachment styles with respect to two targets (i.e., most important therapist and group members).

[g] Two versions, measuring attachment styles with respect to two targets (i.e., most important therapist and group members), are assessed after the third group session.

[h] SMI subscales Vulnerable Child, Angry Child, Impulsive Child, Detached Protector, Punitive Parent, and Healthy Adult.

[i] Key characteristics of a promoting therapeutic environment (i.e., genuineness, safety, and equality) are assessed by 13 items formulated by ST experts.

[j] Assessed after the third group session.

All assessments are performed by trained local research assistants blind to the patients' treatment condition, with the exception of the SCID-5 interviews, demographic interview, and cost interview. The SCID-5 interviews can be administered by trained research assistants as well as trained intake staff members, both blind to treatment condition. The demographic interview and cost interview contain questions on healthcare utilization and

are therefore performed by non-blinded local research assistants. Due to the nature of the interventions, blinding of therapists and patients is not possible. All interviews, except for the SCID interviews, are audio-recorded. Participants receive financial compensation for their involvement in the study. Patients who discontinue their treatment or deviate from the treatment protocol will be encouraged to continue the assessments.

Treatments

Format

For patients of both DBT and ST, treatment has a maximum duration of 25 months and starts with a pretreatment phase of approximately 4 weeks consisting of several (ST: ± three; DBT: ± five) individual sessions in which patients are prepared for the group sessions and become accustomed to their therapists and the treatment model. After the pretreatment phase, patients receive a combined program of individual sessions and group sessions (i.e., treatment phase). Group sessions of both treatments are offered in a semi-open format. If treatment slots are available, new patients can enter the ST group every 10 weeks and for DBT groups at the start of a mindfulness skills module. In DBT, the treatment phase has a maximum duration of 12 months and consists of weekly group sessions (i.e., skills training groups; 150 minutes), weekly individual psychotherapy sessions (50 minutes), and between-session consultation. The between-session consultation, often called telephone consultation although all kinds of technology can be used (van Leeuwen et al., 2021), is offered to the patient within limitations set by the individual therapist, varying from access to between-session support within working hours to 24/7 access to between-session support, which is officially the standard in DBT. In ST, the treatment phase has a maximum duration of 18 months consisting of weekly group (90 minutes) and individual (45 minutes) psychotherapy sessions for a period of 12 months, followed by weekly group psychotherapy sessions and biweekly individual psychotherapy sessions for a period of 6 months. Following the treatment phase, patients continue their treatment with a maintenance phase. The maintenance phase of DBT is a recently developed blended aftercare program with a maximum duration of 12 months. The blended aftercare program was developed based on results of previous studies (e.g., McMain et al., 2012; van den Bosch et al., 2005) and recommendations by several authors (e.g., Royal College of Psychiatrists, 2019; Safer et al., 2009; van den Bosch et al., 2005) to extend the duration of DBT to sustain or even enhance treatment effects. The DBT aftercare program consists of monthly individual psychotherapy sessions, 3-monthly group sessions, and an eHealth intervention in which patients have online access to DBT handouts and worksheets (van Leeuwen et al., 2017). The maintenance phase of ST consists of biweekly individual psychotherapy sessions for a period of 3 months, followed by 3 months of one individual session each month. Excluding the time spent on telephone consultation, homework assignments, and eHealth, and based on 48 working weeks per year, patients will receive about 167 hours of treatment if they follow the treatment protocol. Patients

who have completed treatment successfully before they reach the maximum number of treatment sessions are allowed to complete treatment earlier, although the assessments will be conducted at the originally planned assessment points. Early termination of treatment requires substantial improvements in the primary and secondary outcomes and is decided in a joint decision by the patient and therapist. The treatments are covered by the public health insurance. See Table 2 for an overview of the treatment formats.

Table 2. *Treatment Formats*

ST	Duration	DBT	Duration
Pretreatment phase	4 weeks	Pretreatment phase	4 weeks
Individual sessions		Individual sessions	
Treatment phase	18 months	Treatment phase	12 months
Weekly individual sessions	12 months	Weekly individual sessions	
Weekly group sessions		Weekly skills training groups	
Biweekly individual sessions	6 months	Telephone consultation	
Weekly group sessions			
Maintenance phase	6 months	Maintenance phase	12 months
Biweekly individual sessions	3 months	eHealth intervention	
Monthly individual sessions	3 months	Monthly individual sessions	
		3-monthly group sessions	

Note. Early termination of treatment during the treatment phase or maintenance phase is permitted in case of successful recovery.

Schema Therapy (ST)

ST, developed by Jeffrey Young (Young & Klosko, 1993; Young et al., 2003), is based on an integrative cognitive model, combining cognitive behavioral therapy and experiential techniques with insights from developmental theories, including attachment theory, and psychodynamic concepts (Jacob & Arntz, 2013). Central concepts are early maladaptive schemas and schema modes. Early maladaptive schemas can be defined as broad, pervasive patterns of thoughts, emotions, memories, and cognitions regarding oneself and relationships with others, developed during childhood (Young et al., 2003). ST assumes that the frustration of core needs and early traumatic experiences lead to the development of early maladaptive schemas. A schema mode refers to an activated set of schemas and the associated coping response (i.e., overcompensation, avoidance, and surrender), and describes the momentary emotional, cognitive, and behavioral state of the patient. The following schema modes are characteristic of BPD (Arntz et al., 2005): (1) vulnerable child mode, associated with a fear of abandonment and strong emotions, such as loneliness, sadness, and helplessness, (2) angry and impulsive child mode, characterized by anger, frustration, hostility, and impulsivity,

(3) punitive parent mode, representing the internalized voice of very punitive and critical attachment figures and associated with self-criticism, self-hatred, guilt, and self-denial, (4) detached protector mode, characterized by attempts to cut off the self from needs and feelings, resulting in symptoms of detachment, substance misuse, social withdrawal, and self-harm, and (5) healthy modes, reflecting functional thoughts, cognitions, and behavior (i.e., healthy adult mode) and the feeling that core needs are being fulfilled (i.e., happy child mode). The first four modes are maladaptive schema modes and central to BPD. The last two modes are functional and often only weakly present at the beginning of the treatment (Kellogg & Young, 2006). Idiosyncratic schema mode models usually cover additional modes, depending on the specific problems and comorbidity of the patient.

ST aims to enable patients to fulfill their needs, reduce maladaptive schema modes, and strengthen adaptive schema modes. In this study, ST is offered in a combined group-individual format developed by Farrell and Shaw (2012). The group acts as an analogue of a family with the other patients as "siblings" and the two therapists as "parents" (Farrell & Shaw, 2012). The group may speed up and amplify the effect of treatment by offering corrective emotional experiences, peer support, opportunities for in vivo practice, and a sense of understanding (Farrell et al., 2009). The individual ST follows the protocol as described by Arntz and van Genderen (2020).

Dialectical Behavior Therapy (DBT)

DBT is a comprehensive cognitive behaviorally based treatment for BPD, integrating strategies from cognitive and behavioral treatments, Zen-based acceptance strategies, and dialectical strategies (Linehan, 1993a, 1993b). Linehan (1993a, 2015) proposed a skills deficit model in which emotion regulation is central. More specifically, the model holds that the problematic behaviors associated with BPD (e.g., suicide attempts, self-injury, substance use) are in fact best understood as dysfunctional attempts to regulate emotions. Emotion dysregulation results from the complex transaction between dispositional emotional vulnerabilities and an adverse invalidating environment. Therefore, the treatment involves balancing problem solving strategies with loads of validation. DBT aims to help patients develop new skills, enhance motivation, ensure generalization of skills use, and change their environment if needed. In addition, DBT aims to enhance therapists' motivation to deliver effective treatment (Linehan, 2015).

DBT involves skills training groups, individual therapy, between-session consultation, and therapist consultation team meetings. DBT skills training groups teach patients behavioral skills in four different, yet inter-related, areas: mindfulness, interpersonal effectiveness, emotion regulation, and distress tolerance / radical acceptance. Individual therapy focuses on motivational issues and the acquisition and use of skills in daily life. A predetermined ordering of treatment targets is used in individual sessions and part of different stages of the treatment. Stage 1 focuses on stabilizing the patient and behavior control. Targets in this stage

of the treatment include: life-threatening behavior, therapy-interfering behavior, quality-of-life-interfering behavior, and behavior skills. Stage 2 focuses on reducing posttraumatic stress and requires exposure to trauma-related cues (Linehan, 1993a). Finally, Stages 3 and 4 target self-respect and the sense of incompleteness. However, due to time constraints, some patients might not enter all stages and most studies have focused on Stage 1 DBT (Wagner & Linehan, 2006). Individual therapists provide between-session (telephone) consultation if needed. According to the guidelines of DBT, access to between-session consultation outside of office hours, preferably by the individual therapist, is part of DBT (Linehan, 1993a). In this trial, between-session consultation by the individual therapist will be within limitations set by the therapist, which can vary from support provided within working hours to 24/7 access to telephone consultation. As access to between-session (telephone) consultation will vary between centers and individual therapists, the effect of therapists' availability for between-session support will be examined. Finally, DBT therapists meet weekly in a DBT consultation team to motivate and support each other.

Therapists, training, and supervision

The therapists in this study will be licensed psychologists, psychotherapists, psychiatrists, or psychiatric nurses. Individual and group schema therapists must have completed a basic training in individual ST. Group schema therapists must have also completed a four-day training in the group schema therapy model of Farrell and Shaw (2012). All schema therapists receive a one-day training in experiential techniques by a certified ST trainer. DBT therapists are required to complete a three-day training in DBT and at least one member of the DBT team must have completed the 10-day intensive DBT training. In addition, DBT therapists receive a two-day kick-off training by certified DBT trainers to expand their knowledge of DBT. Moreover, DBT therapists were given the opportunity to participate in a one-day training in imaginal exposure. According to Linehan (1993a), reducing behaviors and stress response patterns related to traumatic life events is a primary DBT target. Reducing posttraumatic stress is mostly part of Stage 2 of DBT and involves exposure to trauma-associated cues (Dimeff et al., 2020; Linehan, 1993a). However, some of the DBT therapists expressed concerns about their ability to apply the principles and procedures of exposure to treat traumatic memories in BPD patients. Therefore, the opportunity to participate in a one-day exposure training was offered to the therapists.

Before the start of the study, schema therapists should have received at least 10 individual supervision sessions by a licensed supervisor. There is no requirement for the minimum number of DBT supervision sessions. During the study, therapists receive supervision over a period of 2 years by certified supervisors. ST supervision is provided through teleconferencing biweekly in the beginning, then (2-)monthly after 6 to 12 months, depending on the experience of the therapists. DBT therapists receive supervision at the location every 3 months. Moreover, there will be weekly DBT team meetings (i.e., DBT consultation team

meetings) and biweekly ST team meetings. All individual ST sessions will be audiotaped, while individual DBT sessions and ST and DBT group sessions will be videotaped. These recordings are used for supervision and treatment adherence ratings. Treatment adherence, a component of treatment integrity (i.e., the extent to which a treatment is implemented as intended; Perepletchikova et al., 2007), refers to the extent to which the therapist utilizes prescribed techniques and procedures and avoids the use of proscribed techniques and procedures (Waltz et al., 1993). Adherence will be assessed in a random selection of session recordings by trained raters (master's students in psychology) blind to treatment condition. Master's students in psychology will be trained by ST and DBT experts by using session recordings not used in the final adherence rating to practice with the instruments. Individual ST sessions will be rated on an adapted version of the Therapy Adherence and Competence scale for ST for BPD (Young et al., 2006) and group ST sessions will be rated on the Group Schema Therapy Rating Scale – Revised (Zarbock et al., 2014). Individual DBT sessions will be rated on the Dutch translation of the observer-rated version of the DBT Adherence Checklist for Individual Therapy (Harned et al., 2021). An observer-rated instrument will be developed to assess the skills training groups.

Other treatment

During the treatment, patients are not allowed to engage in any other form of psychological treatment. However, in case of acute crisis, the crisis procedures of the treatments will be followed (e.g., telephone consultation by the therapist, contact with a crisis line, visit the emergency room, hospitalization, individual crisis management sessions). Any additional treatment will be recorded and included in the analyses. Patients will only be withdrawn from the study at their request.

Coronavirus disease (COVID-19) pandemic

This study is conducted during the COVID-19 pandemic. The pandemic is expected to have adverse effects on patients with mental health disorders (Hao et al., 2020). In addition, in case face-to-face treatment is restricted in mental healthcare centers because of government and healthcare center policy, the treatment will be delivered via videoconferencing. Consequently, differences between patients will arise regarding the amount of treatment sessions delivered during the pandemic and/or via videoconferencing. We will control for a potential influence of the COVID-19 pandemic by, for example, adding dynamic regression parameters that include the impact of time in treatment during the pandemic. The definition of the indicator variable indicating the COVID-19 pandemic will be decided before start of the data analyses (e.g., dummy variable indicating pandemic/no pandemic or continuous variable indicating the amount of time in treatment during the pandemic), given the unpredictability of the current situation. Moreover, exploratory analyses may be conducted to investigate the potential influence of the deviating treatment format (i.e., online vs. face-to-face individual sessions and/or group sessions) on the treatment effectiveness.

In addition, the assessments will be conducted via videoconferencing or phone, and the computer-based questionnaires will be completed by participants at home, if face-to-face assessments are not allowed. Before receiving the treatment and/or assessments via videoconferencing, patients will sign an additional informed consent form.

Data management, storage, monitoring, and dissemination

Data are collected with a unique identifier for each patient (i.e., pseudonym) using the online survey software program Qualtrics (2019) and the web tool Lotus, which has been especially developed for longitudinal research by the University of Amsterdam. The list of pseudonyms and personal information of patients within a particular mental healthcare center is securely stored at the center and only accessible for the research assistant and coordinator of this center. A different set of pseudonyms is used for data collected by clinicians (i.e., intake questionnaire and recordings). The list with the combination of both pseudonyms of patients is only accessible for the research assistant and coordinator of the center and the authorized researchers. The data are stored on a secure storage server of the University of Amsterdam, accessible only to authorized researchers.

All (serious) adverse events reported by the patient or observed by clinicians or researchers will be recorded. There is no data monitoring committee and the study will not be audited. The results of the study will be disseminated in scientific journals and at (inter)national scientific conferences.

Measures

The instruments include screening measures, measures to assess potential predictors and mediators of treatment response, and outcome measures. The instruments that were not available in Dutch were translated (i.e., Brief Experiential Avoidance Questionnaire, Dialectical Behavior Therapy-Ways of Coping Checklist, Gordon Test of Visual Imagery Control, Positive Mental Health scale, and social problems) by bilingual experts. The translations were checked for consistency with the original version. Items, questionnaires, and interviews that have been developed or modified by the authors are available upon request from the first author.

Screening

Mental disorders. The SCID-5 is a semi-structured interview used to diagnose DSM-5 disorders. Personality disorders are assessed with the SCID-5-PD (First et al., 2015) and syndrome disorders are assessed using the SCID-5-S (First et al., 2018), which is an extended version of the SCID-5 Clinician Version (SCID-5-CV; First et al., 2016). Supplement 1 offers an overview of all syndrome disorders that are assessed by the SCID-5-S. Based on a first psychometric evaluation in a psychiatric patient sample, Somma et al. (2017) found an adequate interrater reliability of the SCID-5-PD. In addition, the SCID-5-CV has demonstrated good psychometric properties (Mohammadkhani et al., 2020; Osório et al., 2019; Shankman et al., 2018).

Before administering the SCID-5-S and/or SCID-5-PD, self-report screening questionnaires (SCID-5-SPQ; First et al., 2016, and SCID-5-SV; Arntz et al., 2018) may be administered. In accordance with the instructions for administering the SCID, disorders and criteria of disorders not affirmed by the screening questionnaires and not considered as false negatives by the clinician will be assumed to be absent. The SCID-5 will be assessed during the screening phase and 12 months after the end of treatment.

Motivation and availability. A 13-item semi-structured motivation interview is used to assess several exclusion criteria (e.g., no fixed address, have received ST or DBT in the past year) and patients' motivation and availability.

Predictors

As mentioned, candidate predictor variables of (differential) treatment response have been selected using a multi-method approach (i.e., literature, suggestions of a patient representative of the Borderline Foundation of the Netherlands, and semi-structured interviews with 18 expert clinicians). Supplement 2, Table S1 presents an overview of the predictors that have emerged during the semi-structured interviews with clinicians. Supplement 2, Table S2 provides an overview of the predictors based on the literature and suggestions of a patient representative. The candidate predictors of (differential) treatment response are assessed at baseline. Only the measures that are not part of the screening, mediator or outcome measures will be briefly described in this paragraph.

Autistic traits. Autistic traits are assessed by the abbreviated version of the Autism Spectrum Quotient, the AQ-10 (Allison et al., 2012). The AQ-10 consists of 10 items rated on a four-point Likert scale. The AQ-10 has demonstrated acceptable psychometric properties in an adult general population sample (Lundin et al., 2019).

Commitment. Patient commitment to treatment is measured with a selection of items of the subscale Motivation to Engage of the Treatment Motivation Scales for forensic outpatient treatment (TMS-F; Drieschner & Boomsma, 2008). The four items can be rated on a seven-point Likert scale.

Experiential avoidance. The Brief Experiential Avoidance Questionnaire (BEAQ; Gámez et al., 2014) is a 15-item scale assessing experiential avoidance across six domains (i.e., behavioral avoidance, distress aversion, suppression, procrastination, repression/denial, and distress endurance). The items can be rated on a six-point Likert scale. The BEAQ has shown good psychometric properties among psychiatric outpatients (Gámez et al., 2014).

Frustration intolerance. Frustration intolerance is assessed by the Frustration Tolerance subscale of the Severity Indices of Personality Problems (SIPP-118; Verheul et al., 2008). This subscale consists of eight 4-point Likert scale items measuring the capacity to cope with setbacks and disappointments. In previous research among Dutch patients with a personality disorder, the subscale demonstrated moderate to good reliability (Verheul et al., 2008).

Insight. A modified version of the Self-Reflection and Insight Scale (SRIS; Roberts & Stark, 2008; van Dulmen et al., 2014) is used to assess self-reflection and insight. The SRIS contains 20 five-point Likert scale items. The SRIS has shown good reliability and validity in student samples (Grant et al., 2002; Roberts & Stark, 2008).

Internal locus of control. Internal locus of control, defined as the extent to which a person experiences an outcome as the result of their own behavior or personal characteristics rather than external circumstances, is assessed by the Locus of Control scale (IE; Sanderman, 1988). The IE contains 10 five-point Likert scale items. Previous research has demonstrated adequate psychometric properties (Kamphuis & Emmelkamp, 1998; Sanderman, 1988).

Level of personality functioning. The Level of Personality Functioning Scale-Brief Form 2.0 (LPFS-BF 2.0; Weekers et al., 2019) assesses impairment in personality functioning according to the DSM-5 alternative model for personality disorders. The LPFS-BF 2.0 contains 12 four-point Likert-scale items. Based on a first psychometric evaluation among Dutch patients referred to a specialized mental healthcare center for personality disorders, the LPFS-BF 2.0 has demonstrated satisfactory psychometric properties (Weekers et al., 2019).

Mental imagery capacity. Mental imagery capacity is assessed with the 12-item Gordon Test of Visual Imagery Control (TVIC; Richardson, 1969). The TVIC assesses the ability to visualize and manipulate a given scenario in response to a set of cues. Participants can respond on a three-point Likert scale. In addition to the 12 Likert scale items, we measure the time it takes the participant to visualize the scenarios. Finally, we have added two 100 mm visual analog scale (VAS) items measuring how well participants see the scenarios that were described and how difficult it was for the participant to visualize the different scenarios. The TVIC has demonstrated fair to satisfactory internal consistency and validity among community samples and undergraduates (Lequerica et al., 2002; Odou & Vella-Brodrick, 2013; Pérez-Fabello & Campos, 2004; White et al., 1977).

Mentalizing capacity. Mentalizing capacity is measured using an eight-item version of the Reflective Functioning Questionnaire (RFQ-8; Fonagy et al., 2016). The RFQ-8 comprises two dimensions: uncertainty about mental states, reflecting hypomentalizing, and certainty about mental states, indicating hypermentalizing. The RFQ-8 uses a seven-point Likert scale. In previous research among BPD patients, the questionnaire has demonstrated satisfactory psychometric properties (de Meulemeester et al., 2018; Fonagy et al., 2016; Morandotti et al., 2018).

Perfectionism. The eight-item Frost Multidimensional Perfectionism Scale-Brief (F-MPS-Brief; Burgess et al., 2016) assesses perfectionism across two dimensions (evaluative concerns and striving). Items are rated on a five-point Likert scale. Psychometric properties of the F-MPS-Brief were found to be good in clinical and community samples (Burgess et al., 2016).

Personality traits. Personality traits are measured, among others, with the Ten-Item Personality Inventory (TIPI; Gosling et al., 2003; Hofmans et al., 2008), which is a brief

measure of the Big-Five personality dimensions. The 10 items can be rated on a seven-point Likert scale. The TIPI has shown low to moderate internal consistency and adequate validity among students (Gosling et al., 2003; Hofmans et al., 2008).

Positive mental health. Positive mental health, often referred to as mental well-being, is assessed using the nine-item Positive Mental Health scale (PMH-scale; Lukat et al., 2016). The items can be rated on a nine-point Likert scale. Based on a previous study on the psychometric properties of the PMH-scale in student, patient and general samples, the PMH-scale was found to be a reliable and valid instrument (Lukat et al., 2016).

Psychopathology and maladaptive personality traits. The Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008) measures a wide range of psychopathology symptoms, personality characteristics, and behavioral proclivities. The MMPI-2-RF consists of 338 true-false items aggregating onto 51 individual scales. The psychometric properties of the MMPI-2-RF varied from inadequate to good among normative, outpatient and inpatient samples, as documented in detail in the Technical Manual (Tellegen & Ben-Porath, 2008).

Readiness to change. Readiness to change is assessed by two subscales (contemplation and action) of the 24-item version of the University of Rhode Island Change Assessment (URICA; DiClemente & Hughes, 1990; DiClemente et al., 1999; McConnaughy et al., 1983). Both subscales are measured by six 5-point Likert scale items and have demonstrated good reliability across a diversity of studies (e.g., Carbonari et al., 1994; Carbonari & DiClemente, 2000; Dozois et al., 2004).

Rigidity. Rigidity is measured by the rigidity subscale of the Computerized Adaptive Test of Personality Disorder-Static Form (CAT-PD-SF; Simms et al., 2011). The Rigidity subscale contains 10 five-point Likert scale items reflecting an unwillingness to consider alternative perspectives and inflexibility in values and beliefs. The subscale has demonstrated good reliability among community adults with current or a history of mental health treatment (Simms et al., 2011).

Social problems. By using the social problems list, derived from the Improving Access to Psychological Therapies (IAPT) program (Clark, 2011), social problems (e.g., financial problems, housing problems, and unemployment) are assessed in direct discussion with the patient.

Social support. The Multidimensional Scale of Perceived Social Support (MSPSS; Zimet et al., 1988) is assessed to investigate perceived support from three sources: significant others, family, and friends. The MSPSS contains 12 items that can be rated on a seven-point Likert scale. Psychometric properties of the MSPSS are satisfactory among psychiatric outpatients and BPD patients (Cecil et al., 1995; O'Toole et al., 2012). In addition to the MSPSS, the research assistant rates the patient's social network taking into account the size of the network and potential pathogenic influences.

Stigma of immutability. BPD has been associated with stigma of immutability (Aviram et al., 2006). We have developed five 7-point Likert scale items assessing the extent to which participants believe that BPD is resistant to treatment.

Trauma. The Traumatic Experience Checklist (TEC; Nijenhuis et al., 2002) is used to assess traumatic experiences, including emotional abuse, emotional neglect, sexual abuse, sexual harassment, physical abuse, and threat to life/ bizarre punishment/ intense pain. The TEC includes 30 descriptions of various traumatic experiences. The TEC has demonstrated favorable psychometric properties in Dutch psychiatric patients (Nijenhuis et al., 2002).

Verbal intelligence. The Dutch version of the National Adult Reading Test (DART; Schmand et al., 1992) is used as a proxy for verbal intelligence. The DART is a reading test including 50 irregularly spelled words. Based on previous research, the DART yields an adequate estimation of verbal intelligence and has shown adequate psychometric properties across a variety of populations (Schmand et al., 1991).

Other patient characteristics, collected using a self-report questionnaire. In addition to the questionnaires, participants fill out several questions developed by the authors about the willingness and ability to engage in a therapeutic relationship, perceived suitability of DBT and ST (treatment preference), and the absence or presence of an attachment figure in the past.

Other patient characteristics, collected using a questionnaire filled out by clinicians (intake questionnaire). Clinicians responsible for the intake assessment will fill out the nine-item intake questionnaire for each participant, including questions about the willingness and ability to engage in a therapeutic relationship, the willingness and ability to examine the link between childhood history and present problems, high vs. low level borderline personality organization (Lenzenweger et al., 2008), the request for help, the degree to which a syndrome disorder might interfere with treatment response, and perceived suitability of DBT and ST. These questions have been formulated by the authors.

Mediators

Both treatments include non-specific (attachment and therapeutic alliance), BPD-treatment-generic (therapeutic environment characterized by genuineness, safety, and equality), and BPD-treatment-specific (ST: beliefs and schema modes; DBT: emotion regulation and skills use) mechanisms of change. The proposed mediators are repeatedly measured: at baseline, except for measures requiring information about the therapy (i.e., therapeutic environment, therapeutic alliance, and attachment styles with respect to the most important therapist and group members), and every 6 months after the start of the treatment phase. In addition, a selection of the proposed mediators (i.e., selection of schema modes, skills use, beliefs, and therapeutic environment) are also collected every 3 months after the start of the treatment phase, during the first 2 years.

Attachment. The Experiences in Close Relationships-Relationship Structures Questionnaire (ECR-RS; Fraley et al., 2011) is a brief version of the Experiences in Close Relationships-Revised (ECR-R; Fraley et al., 2000). The ECR-RS measures attachment patterns in different relational domains, such as relationships with parents and friends. The ECR-RS can also be adapted to measure a person's general attachment style. In this study, three versions of the ECR-RS are used, measuring general attachment style and attachment styles with respect to two targets (i.e., most important therapist and group members). The ECR-RS contains nine items, assessing two attachment dimensions: attachment-related anxiety and avoidance. The items can be rated on a seven-point Likert scale. The ECR-RS has shown adequate psychometric properties in a large web-based sample ($N > 21.000$), comparable to the ECR-R (Fraley et al., 2011). As experience with the treatment is required in order to be able to complete the questions about the most important therapist and group members, these questions will be filled out 3 weeks after the start of the treatment phase.

Beliefs. Idiosyncratic dysfunctional beliefs were elicited with a semi-structured interview at baseline. Three to five idiosyncratic dysfunctional beliefs related to the self (e.g., "I am worthless"), others (e.g., "People always reject me"), and emotions (e.g., "Expressing emotions is a sign of weakness") are formulated. Participants rate the degree to which they believe in each statement on a 100 mm VAS at baseline and at every subsequent assessment. This procedure has been used in previous research (e.g., Videler et al., 2018; Vuijk & Arntz, 2017). The VAS has been found to be useful for assessing variations in the intensity of beliefs in patients with a personality disorder (Vuijk & Arntz, 2017). In addition to the idiosyncratic dysfunctional beliefs, participants rate the credibility of one functional belief ("I consider myself a good person") on a 100 mm VAS.

Emotion regulation. Emotion regulation is assessed by the Difficulties in Emotion Regulation Scale Short Form (DERS-SF; Kaufman et al., 2016), a brief version of the widely used DERS (Gratz & Roemer, 2004). The DERS-SF measures non-acceptance of emotional responses, difficulties engaging in goal-directed behavior, impulse control difficulties, limited access to emotion regulation strategies, lack of emotional clarity, and lack of emotional awareness. The Awareness subscale is excluded based on recommendations of among others Hallion et al. (2018) and Bardeen et al. (2012). Lack of emotional awareness is assessed by the Awareness subscale of the Difficulties in Emotion Regulation Scale 18 (DERS-18; Victor & Klonsky, 2016). The DERS-SF, without the awareness subscale, consists of 15 items. The Awareness subscale of the DERS-18 is measured by three items. All items can be rated on a five-point Likert scale. Both questionnaires have demonstrated good psychometric properties among outpatients (Hallion et al., 2018).

Schema mode ratings. The Schema Mode Inventory (SMI; Young et al., 2007) measures the extent to which 16 different (dysfunctional as well as functional) schema modes are endorsed. The SMI consists of 143 items that are scored on a six-point Likert scale. Previous research

using a sample of non-patients and patients with a syndrome disorder and/or personality disorder has demonstrated acceptable psychometric properties (Lobbestael et al., 2010). The five maladaptive schema modes that are central to BPD (i.e., vulnerable child, angry child, impulsive child, detached protector, and punitive parent; Arntz et al., 2005) and one functional schema mode (i.e., healthy adult) are assessed every 3 months during the first 2 years.

Skills use. The 59-item Dialectical Behavior Therapy-Ways of Coping Checklist (DBT-WCCL; Neacsiu et al., 2010) is an adaptation of the Revised Ways of Coping Checklist (RWCCL; Vitaliano et al., 1985). The DBT-WCCL measures DBT skills use and maladaptive coping skills use over the previous month. All items are assessed using a four-point Likert scale. The DBT-WCCL has shown adequate to excellent reliability and validity among BPD patients (Neacsiu et al., 2010).

Therapeutic alliance. The therapeutic alliance is measured with the Working Alliance Inventory-Short (WAI-S; Stinckens et al., 2009; Tracey & Kokotovic, 1989). The WAI-S consists of three subscales (agreement on goals, agreement on tasks, and bond between patient and therapist), each consisting of four items which can be scored on a five-point Likert scale. Observed psychometric properties of the WAI-S were satisfactory in a patient sample (Smits et al., 2015; Stinckens et al., 2009). Since experience with the treatment is required in order to be able to complete the WAI-S, the WAI-S will be filled out 3 weeks after the start of the treatment phase.

Therapeutic environment. Key characteristics of a promoting therapeutic environment (i.e., genuineness, safety, and equality) are assessed by 13 items formulated by ST experts (A. Arntz and O. Brand-de Wilde) and rated on a 100 mm VAS. The items measure the extent to which the participant feels a) the individual therapist, group therapists, and group members are genuine with him/her; b) he or she can tell the individual therapist and group therapists everything; c) safe in the individual and group therapy; d) safe to show vulnerability and express negative feelings in the individual and group therapy; e) the individual and group therapists take personal responsibility for their mistakes; and f) the individual and group therapists see him/her as equal. Since experience with the treatment is required in order to be able to complete this questionnaire, this questionnaire will not be assessed at baseline.

Primary outcome

BPD severity. The primary outcome measure is the change in severity and frequency of the DSM-5 BPD manifestations from baseline to 3-year follow-up, assessed with the total score of the Borderline Personality Disorder Severity Index version 5 (BPDSI-5; Arntz et al., 2003; Giesen-Bloo et al., 2010). The BPDSI-5 is a semi-structured interview consisting of 70 items rating the nine DSM-5 BPD criteria over the prior 3 months. All items are rated on an 11-point Likert scale (0 = *never* to 10 = *daily*), except for the subscale Identity Disturbance which is rated on a 5-point Likert scale (0 = *absent* to 4 = *dominant, clear, and well-defined*) and multiplied by 2.5. The total score is the sum of the nine criteria scores and ranges from 0 to 90. The scores

on the BPDSI-5 subscales provide information on the severity of each of the nine criteria. The BPDSI-5 is a modified version of the BPDSI-IV (Arntz et al., 2003; Giesen-Bloo et al., 2010) in which a few questions have been slightly reworded and exact frequency scores have been added in addition to the Likert scale. The BPDSI-IV has proven to be a reliable and valid measure among non-patients and (BPD) patients (Arntz et al., 2003; Giesen-Bloo et al., 2010). Previous research has shown that a cutoff score of 15 differentiates between BPD patients and controls (Giesen-Bloo et al., 2010). In addition, a score of 20 distinguishes BPD patients from non-BPD patients (Dickhaut & Arntz, 2014; Giesen-Bloo et al., 2010; Nadort et al., 2009).

Secondary outcome measures

As accumulating evidence suggests that BPD severity and level of functioning are only loosely associated, attention will be paid to outcomes in different areas, including symptoms, functioning, and well-being. The outcome measures are administered at baseline and every 6 months after the start of the treatment phase. In addition, patients' ratings of experienced burden due to BPD manifestations and well-being are collected every 3 months after the start of the treatment phase, during the first 2 years.

Costs. Costs, including healthcare costs, patient and family costs, and costs outside the healthcare sector, are measured using a retrospective cost interview especially designed for BPD patients (Wetzelaer et al., 2014). Healthcare costs include visits to general practitioners, hospitals, crisis centers, psychologists and psychiatrists, use of medication, social work, paramedical care, and alternative treatments. Patient and family costs include informal care (i.e., care provided by the patient's family, friends, or neighbors) and out-of-pocket costs (e.g., drugs, alcohol, excessive spending). Costs in other sectors include productivity losses from unpaid work (study and voluntary work) and paid work. Since it is difficult to distinguish between BPD-related costs and costs due to other psychological disorders (Wagner et al., 2014), only a distinction will be made between costs due to psychological disorders and costs due to somatic diseases. The cost interview will be conducted by trained research assistants using a recall period of 6 months (baseline assessment), the number of weeks since randomization (assessment 6 months after the start of the treatment phase), or the number of weeks since the previous assessment (assessments 12, 18, and 24 months after the start of the treatment phase and both follow-up assessments).

Dutch guidelines (Hakkaart-van Roijen et al., 2015; Kanters et al., 2017) will be used to determine total costs. Healthcare costs will be calculated as volumes of resource use multiplied by their corresponding unit costs, derived from Hakkaart-van Roijen et al. (2015). Prescribed medication costs will be determined based on national reference prices. Informal care costs will be computed by multiplying the number of hours the patient receives informal care by shadow prices (Hakkaart-van Roijen et al., 2015). Shadow prices will also be used to value lost productivity in study and voluntary work. Productivity losses from paid work will be valued according to the Human Capital Approach (Hodgson & Meiners, 1982). Out-of-

pocket costs, such as alcohol and excessive spending, will be directly retrieved from the cost interview or, in case of over-the-counter medication, from the Dutch Pharmacotherapeutic Compass (Hakkaart-van Roijen et al., 2015).

Demographics. General patient characteristics (e.g., age, ethnicity, marital status, educational level, employment status) will be collected using a semi-structured demographic interview. During this interview, additional patient characteristics such as treatment history, request for help, medication use, substance use, and duration of BPD manifestations will be recorded. For an overview of all characteristics, see Supplement 2.

Experienced burden due to BPD. Patients' self-reported experienced burden of BPD manifestations is measured using the Ultrashort BPD Checklist, a shortened version of the validated BPD Checklist (Bloo et al., 2017). The Ultrashort BPD Checklist consists of nine to 11 five-point Likert scale items (the number of items will be based on the upcoming validation study), each related to a specific DSM-5 BPD criterion. Based on an initial psychometric evaluation, the Ultrashort BPD Checklist showed good to excellent psychometric properties in a sample with BPD and cluster C patients, patients with a syndrome disorder, and non-patients, similar to the BPD Checklist (Bloo et al., 2017).

General psychopathological symptoms. The Brief Symptom Inventory (BSI; de Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983) is a self-report instrument measuring general psychiatric symptoms at the time of assessment. The BSI is a short version of the Symptom-Check-List (SCL-90-R) and contains 53 items assessing nine symptom dimensions: somatization, obsession-compulsion, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. All items are assessed using a five-point Likert scale. Previous research in Dutch community and patient samples has demonstrated good reliability and validity (de Beurs & Zitman, 2006; de Beurs, 2011).

Global functioning and impairment. Global functioning and impairment is assessed by the 36-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) interview version (Üstün et al., 2010). The WHODAS 2.0 is a general measure to assess disability in six major life domains (cognition, mobility, self-care, getting along, life activities, and participation). For each item, participants have to report how much difficulty they experienced in the last 30 days. The six domain scores and overall functioning score have shown good psychometric properties in a general population sample as well as a patient sample (Üstün et al., 2010).

Sleep. Insomnia complaints are assessed by the Insomnia Severity Index (ISI; Bastien et al., 2001). The ISI contains seven items that are scored on a five-point Likert scale. The ISI has been shown to be a valid measure in community and insomnia patient samples (Bastien et al., 2001), although the reliability was questionable in some studies (e.g., Lancee et al., 2016; Thorndike et al., 2011). In addition to insomnia, the number of nights with nightmares

and the total number of nightmares in the week prior to the assessment are measured using the Nightmare Frequency Questionnaire (NFQ; Krakow et al., 2000). Based on previous research among posttraumatic stress disorder (PTSD) patients, the NFQ appears reliable for measuring nightmare frequency (Krakow et al., 2000).

Quality of life. Generic quality of life is assessed using the 5-level EuroQol 5D version (EQ-5D-5L; Herdman et al., 2011). The questionnaire measures five health state dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). Each dimension is divided into five severity levels: no problem, slight problems, moderate problems, severe problems, and extreme problems. The profiles from the five health state dimensions are assigned a value based on the Dutch social tariffs to generate health utilities (Versteegh et al., 2016). These utilities will be used to calculate Quality Adjusted Life Years (QALYs) by multiplying the change in utility values between assessments by the length of the period between assessments. In addition to the five health state dimensions, the EQ-5D-5L contains a VAS item which records the patient's self-reported health status ranging from 0 (worst health you can imagine) to 100 (best health you can imagine). The EQ-5D-5L has been shown to be a reliable and valid measure among different patient groups in different countries (Janssen et al., 2013).

As a complement to the EQ-5D-5L, the Mental Health Quality of Life seven-dimensional Questionnaire (MHQoL-7D; van Krugten et al., 2019) will be administered. The MHQoL-7D is a recently developed instrument to assess quality of life specifically in people with mental health problems. The MHQoL-7D consists of seven quality of life domains (self-image, independence, mood, relationships, daily activities, physical health, and hope) and a VAS item which records the patient's self-reported psychological well-being. A study into the psychometric properties of the MHQoL-7D is currently running. The MHQoL-7D will only be included in the analysis if it is demonstrated to be a psychometrically sound instrument and Dutch social tariffs are available.

Well-being. Well-being is measured using a single item measuring happiness (Abdel-Khalek, 2006) and the Outcome Rating Scale (ORS; Miller et al., 2003). The single item measures general happiness in the months prior to the assessment on a seven-point Likert scale. Reliability and validity were good among undergraduates (Abdel-Khalek, 2006), and sensitivity to change was excellent in a BPD sample (Dickhaut & Arntz, 2014). The ORS consists of four VAS items assessing four areas of functioning: individual (personal well-being), interpersonal (family and close relationships), social (work and/or school functioning), and overall (general sense of well-being). We slightly adapted the third dimension of the ORS by excluding friendships, because of its overlap with the second dimension (interpersonal functioning). Hafkenscheid et al. (2010) reported adequate psychometric properties of the ORS in a Dutch outpatient sample.

Statistical analyses

The statistical analyses for the (cost-)effectiveness, mechanisms of change, and treatment selection studies are under development. For example, according to Cohen et al. (2020), the treatment selection field is still in its developmental stage and statistical methods are constantly evolving. Recently, great efforts have been made by several authors (e.g., Buckman et al., 2020; Webb et al., 2020) to select the optimal prediction model by comparing different variable selection techniques. Considering the ongoing advances in methodological approaches, the statistical analyses described below should be considered as examples of appropriate analytic methods. We will determine the optimal methods at the time of the analyses. An update of the protocol will be published, including the selected statistical methods, before the start of the data analyses. The statistical analyses will be performed according to the intention-to-treat (ITT) principle (i.e., including all patients who have been randomized and received at least one treatment session). In addition to the primary analysis based on the ITT principle, a completers analysis will be conducted by excluding patients who dropped out prematurely (i.e., termination of the treatment before planned end, without patient and therapist agreeing that enough improvement has been reached to justify the termination) or deviated from the protocol (e.g., sought other psychological treatment in addition to the study treatment). No interim analyses are planned.

Treatment selection

A two-step approach will be applied to determine the optimal treatment for a particular patient by identifying patient characteristics that predict (differential) treatment response. First, we will examine which of the candidate predictors (see Supplement 2 for an overview) predict (differential) treatment response. Many different variable selection approaches can be used to identify which of the candidate predictors contribute to the prediction of treatment outcome, for example elastic net regularization (Friedman et al., 2010), Bayesian additive regression trees (Kapelner & Bleich, 2016), or a combination of different variable selection procedures (Cohen et al., 2020). Second, individual treatment recommendations are generated based on a prediction model including the variables that predict (differential) treatment response. For each patient, the most beneficial treatment will be identified by using the prediction model to estimate the predicted outcomes for both treatments including the difference in predicted outcomes.

Our primary analysis will focus on individual treatment recommendations based on change in BPD manifestations and will therefore reveal the advantage in symptom relief that may be gained if patients are allocated to their predicted optimal treatment compared to their predicted non-optimal treatment. To investigate the advantage that may be gained in other outcomes, such as functioning and cost-effectiveness, generalization analyses will be performed by testing the performance of the prediction model for these outcomes.

Mechanisms of change

It is hypothesized that the treatments exert a remedial effect on the frequency and severity of BPD manifestations by their impact on the BPD-treatment-specific (ST: beliefs and schema modes; DBT: emotion regulation and skills use), BPD-treatment-generic (therapeutic environment characterized by genuineness, safety, and equality), and non-specific (attachment and therapeutic alliance) mechanisms of change. Since potential mediators (i.e., a variable that (partially) explains the relationship between the treatment and outcome) and outcome will be assessed multiple times, temporal patterns of change can be studied by performing mediation analysis for longitudinal data, for example multilevel autoregressive mediation analysis (Zhang et al., 2018) or multilevel structural equation modeling (Preacher et al., 2010). By using advanced statistical models, the hierarchical structure of the data (repeated measures nested within patients, who in turn are nested within centers) can be taken into account and possible concurrent and temporal relations between mediators and outcome can be investigated.

Clinical effectiveness

Change in the outcome measures and the relative effectiveness of the two treatment conditions will be analyzed using mixed regression so that all available data are used, while taking into account the dependencies among observations nested within individuals nested within centers. Potential center effects are modeled by including a random effect which enables generalization of results outside the trial and maximizes statistical power (Feaster et al., 2011). Since group sessions in both treatments are offered in a semi-open format, patients will start with group treatment at different time points. One can imagine that patients starting treatment at the same time point are more interdependent compared to patients starting treatment at different time points. Therefore, we will take into account, if needed, the interdependency of patients. The underlying distributions of the mixed regression models will be determined based on the variable type (i.e., scale or nominal) and the distribution of residuals (e.g., normal, gamma, negative binomial).

Cost-effectiveness

The cost-effectiveness evaluation will be performed from a societal perspective and includes a cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). The primary clinical outcome for the CEA will be the severity of the BPD manifestations and for the CUA utility scores will be derived from the quality of life instrument(s), both with a time horizon of 12 months after the end of treatment. The net benefit will be used to express cost-effectiveness. For each patient, the net benefit will be calculated by subtracting the costs incurred by the patient from the amount that society is willing to pay for the health benefit (Hoch et al., 2002). The development of the net benefit over time and differences between the treatments will be modeled using multilevel modeling in which the hierarchical structure of the data and potential missing values are taken into account. The best fitting model to describe

the development over time and the appropriate distribution of the net benefit data (e.g., gamma distribution, log-normal distribution) will be based on the data. Cost-effectiveness acceptability curves (CEACs) will be drawn showing the probability that one treatment is more cost-effective compared to the other treatment, given the observed data, for a range of willingness-to-pay values. Sensitivity analyses will be performed to address the uncertainties in methodology and assumptions and to test for the robustness of findings.

Additional substudies

Several secondary studies will be conducted, including, but not limited to, the investigation of the heterogeneity of BPD and substance abuse among BPD patients, a qualitative study into the perspectives of patients and therapists, and psychometric evaluations. First, BPD is characterized by considerable heterogeneity (Leichsenring et al., 2011; Lenzenweger et al., 2008). Over the past years, researchers have attempted to identify BPD subtypes based on different indicator variables (e.g., DSM-5 criteria, interpersonal characteristics, temperament) and different statistical strategies (e.g., exploratory factor analysis, Q-factor analysis, finite mixture modeling) (Hallquist & Pilkonis, 2012). The BPD subtypes that emerged differed substantially between studies. According to Hallquist and Pilkonis (2012), advances in classifying BPD subtypes can be made by using a theoretical model as guidance, for example Kernberg's theory (Kernberg & Caligor, 2005). Therefore, a substudy into the heterogeneity of BPD will be performed based on theoretically justified indicators and state-of-the-art statistical methods.

A second substudy will focus on the co-occurrence of substance abuse and BPD. Research suggests that patients with BPD and substance abuse have more severe problems, including higher rates of suicide attempts, more treatment noncompliance, and increased risk of violence, compared to BPD patients without substance abuse (e.g., Heath et al., 2018; Links et al., 1995; Wilson et al., 2006). However, few trials have assessed the effectiveness of treatments for BPD patients with substance abuse. In addition, research into the effect of BPD treatment on substance abuse is also limited (Trull et al., 2018). Third, qualitative research will be conducted to explore the perspectives of patients and therapists in key areas, including predictors, mechanisms of change, the treatments, and the implementation of the results in clinical practice. Finally, psychometric evaluations of several Dutch questionnaires (e.g., Dialectical Behavior Therapy-Ways of Coping Checklist, Ultrashort BPD Checklist) will be performed.

DISCUSSION

This article described the study protocol of a multicenter RCT focusing on the (differential) treatment effectiveness of DBT and ST for patients with BPD. The primary aim of the study is to improve the treatment outcomes of DBT and ST for BPD patients by optimizing treatment selection by identifying patient characteristics that specify which patients will benefit most from which treatment. In addition, we aim to elucidate the change mechanisms of DBT and ST, which is crucial for improving treatments and, in turn, treatment response (Huibers, 2015; Kazdin & Nock, 2003; Kraemer et al., 2002). Finally, the comparative effectiveness and cost-effectiveness of DBT and ST will be evaluated.

This trial provides a unique opportunity to gain more insight into one of the main questions dominating the psychotherapy research agenda: "What works for whom and why?". Although DBT and ST share some important characteristics, different interventions related to different assumed core deficits in BPD are provided (Fassbinder et al., 2016). As each treatment provides a different therapeutic milieu and focuses on different goals and tasks, a particular treatment may be a better fit with some patients compared to others (Keefe et al., 2020). In this study, patient characteristics of (differential) treatment response will be identified and individual treatment recommendations (DBT or ST) will be generated. In addition, for each patient, an estimate will be provided of the potential advantage in symptom relief that might be gained in case the patient was allocated to their indicated treatment. Moreover, the potential advantage in other outcomes, for example functioning and cost-effectiveness, will also be estimated. Knowing which treatment is most cost-effective for whom may lead to more efficient allocation of healthcare resources, which is important, as the current healthcare system is characterized by constraints in resources (e.g., people, time, budget; Drummond et al., 2015). However, before a treatment selection procedure can be implemented in clinical practice, replication and external validation of the prediction model are needed. Subsequently, a prospective study in which the patient and clinician collaborate in selecting the optimal treatment (i.e., shared decision making; Hamann et al., 2003), guided by treatment recommendations based on the prediction model, should be conducted to evaluate the advantage of a treatment selection procedure. By using a state-of-the-art approach, the current study can serve as the starting point for future studies into personalized medicine among BPD patients, and is therefore of great importance.

In addition, this trial provides insight into the comparative (cost-)effectiveness of DBT and ST. Although the effectiveness of both treatments has been established, DBT and ST have not been directly compared. Therefore, and because outcome measures differ substantially between studies on the effectiveness of DBT or ST, hypotheses concerning the differential effectiveness can hardly be formulated. According to the "Dodo Bird effect" (Rosenzweig, 1936; Wampold et al., 1997), all evidence-based psychotherapies are equally effective, suggesting that DBT and ST will produce equivalent outcomes. However, a meta-analysis

into the comparative effectiveness of evidence-based treatments for personality disorders demonstrated that some treatments may be more effective than others (Budge et al., 2013). In addition, Fassbinder et al. (2018) hypothesized that ST may be more effective than DBT in reducing psychiatric comorbidity and improving quality of life, while DBT may lead to a better and faster reduction in self-harming and suicidal behaviors. Moreover, although not assessed in direct comparison with ST, the meta-analysis of Storebø et al. (2020) into psychological treatments for BPD indicated that DBT may be especially effective for BPD severity, self-harm, and psychosocial functioning. They also pointed out that more research into the effects of BPD-tailored treatments, including head-to-head comparisons, is needed. By focusing on an array of outcomes, this study will extend our knowledge of the potential differential effects of DBT and ST.

This study has several strengths. First, this RCT is quite inclusive in terms of patient characteristics, and as such designed to reflect clinical practice to enhance ecological validity. Second, this trial is conducted by a research group including researchers with balanced allegiance to either ST or DBT and an independent researcher (i.e., C.J.M. Wibbelink), to prevent the potential effect of research allegiance on treatment outcomes (Leykin & DeRubeis, 2009). Third, we adopt a broad view on treatment response by including outcome measures reflecting different areas of recovery (e.g., BPD symptoms, functioning, well-being). Focusing on outcomes beyond symptom reduction is in line with patients' views on recovery (Katsakou et al., 2012; Lariviere et al., 2015). In addition, the study follows a multi-method assessment approach, as the outcome measures include both self-report questionnaires and semi-structured interviews. Fourth, we include a large number and broad range of patient characteristics potentially predictive of (differential) treatment response across DBT and ST. Finally, the presumed mediators and outcomes will be frequently measured on multiple time points during the treatments and mediation analyses will be performed by using state-of-the-art statistical analysis methods (Lemmens et al., 2016). This allows us to establish concurrent as well as temporal relationships between the mediators and outcomes (Lemmens et al., 2016). However, according to Lemmens et al. (2017), understanding psychotherapeutic change may be too challenging, even in optimal research designs. Psychotherapy consists of a complex interplay of multiple mechanisms on different levels. Finding that a construct (e.g., therapeutic alliance) mediates treatment outcome does not explain *how* changes in this construct lead to changes in the outcome as it could involve several processes (e.g., cognitions, behaviors, emotions, neural systems) (Holmes et al., 2018). It is therefore highly questionable if these complex processes can be assessed by relatively simple mediational models. As such, this is one of the potential limitations of the current study.

This study has several other limitations that should be considered when evaluating the results. First, as power is conventionally set at 80% (Cohen, 1988, 1992; Di Stephano, 2003; Lakens, 2021), we used a minimum criterion of 80% power for the power analyses. However, this means that we accept a 20% chance of a false negative result. Since DBT and ST are both

evidence-based treatments for BPD, differential effects in treatment outcome may be small or non-existing. To demonstrate equivalence or small effects between treatments, a very large sample size is needed. The sample size of the current study is not large enough (i.e., does not have ≥ 80% power) to reliably detect a small differential treatment effect. However, the comparison of treatments is not the main aim of the study. In addition, according to Luedtke et al. (2019), a sample size of at least 300 patients per condition is required to have sufficient power for applying multivariable prediction models. Nonetheless, they also noticed that a smaller sample size might be justified if studies are designed to develop prediction models that can be tested in future studies. Moreover, the results of this study can contribute to building a database including trials on BPD that can be analyzed with meta-analytic techniques.

Second, this study does not include a no-treatment control group, which might affect internal validity. When improvements are found in both treatments, but no significant differences between the treatments, the absence of a control group implies that it cannot be ruled out that non-specific factors such as attention or time (maturation) caused the improvements. However, including a control group receiving no treatment would clearly be unethical (e.g., patients are at risk of suicide). For similar reasons, it is not possible to standardize medication use and crisis management sessions. Any additional treatment or medication use will be monitored and included in the analyses.

Third, one of the treatment elements of DBT is out-of-office hours between-session (telephone) consultation by the individual therapist. The targets of telephone consultation include, among others, reducing self-harm and suicidal behavior and teaching patients how to apply learned skills in everyday life in order to encourage skills generalization (Linehan, 1993a). In the current study, one center will provide 24/7 access to telephone consultation by the patient's individual therapist, while the other centers will provide telephone consultation within the limitations of the individual therapists, or within working hours. In case of emergency, the standard emergency procedures of each center will be followed. Although outside of office hours availability is considered to be an essential element of DBT by some authors (van den Bosch et al., 2002), the link between telephone consultation and outcome in DBT has not been evaluated (Feigenbaum et al., 2012; van Leeuwen et al., 2021). There is some preliminary support for the importance of telephone consultation (Koons, 2011). However, studies into the effectiveness of DBT that did not apply 24-hour telephone consultation by the individual therapist have found positive outcomes (e.g., Comtois et al., 2007; Feigenbaum et al., 2012). Van den Bosch and Sinnaeve (2015) studied treatment programs of 25 DBT teams in the Netherlands. They found that only 36% of the DBT teams applied telephone consultation according to the guidelines of DBT. It can therefore be concluded that the current study is a good reflection of clinical practice, which enhances generalizability of our findings. Notwithstanding, we will monitor between-session (telephone) consultation within centers and examine potential effects.

Fourth, in this study, a component of treatment integrity (treatment adherence) will be assessed, which is, surprisingly, not standard procedure in trials investigating BPD treatments (Storebø et al., 2020). However, treatment integrity also consists of treatment differentiation and therapist competence (Waltz et al., 1993). Treatment adherence and treatment differentiation are closely related, in contrast to treatment adherence and therapist competence (Miller & Binder, 2002). Treatment adherence represents a quantitative aspect of treatment integrity (i.e., how frequently a therapist utilizes prescribed techniques and procedures and avoids proscribed techniques and procedures), while competence represents a qualitative aspect (i.e., how well prescribed techniques and procedures are implemented) (Perepletchikova et al., 2007). Adherence does not necessarily presuppose competence; even with adequate adherence, therapists may deliver the treatment in an incompetent manner. The absence of competence ratings may threaten the validity of our results (Perepletchikova et al., 2007). Moreover, treatment adherence will be assessed by trained master's students in psychology, whereas for DBT, adherence ratings by reliably trained therapists are considered the gold standard (Harned et al., 2021; Lynch et al., 2007). However, students will receive training from experienced therapists.

Fifth, it is a subject of some debate whether the EQ-5D is a valid instrument to measure quality of life in BPD patients, which can affect the economic evaluation (Brettschneider et al., 2014). According to Brazier (2010), the EQ-5D might not measure what matters to patients with psychiatric disorders. In addition, in a cost-effectiveness study among BPD patients, van Asselt et al. (2008) found contradictory results on the incremental risk ratios when recovery was based on the EQ-5D compared to the BPDSI-IV. In contrast, adequate responsiveness of the EQ-5D has been found in a BPD sample (van Asselt et al., 2009). In addition, Soeteman et al. (2010) concluded that the EQ-5D is sensitive to changes in the health status of patients with cluster B personality disorders. In addition to the EQ-5D, quality of life will be assessed by a recently developed instrument specifically developed for patients with mental health problems (MHQoL-7D; van Krugten et al., 2019). The validation of this instrument is currently in progress, but preliminary results are promising (van Krugten et al., 2019). Another point for consideration is that conclusions with regard to the most cost-effective treatment choice can be affected by the amount society is willing to pay for an additional unit of effectiveness (i.e., willingness-to-pay threshold). Soeteman et al. (2010) concluded that outpatient psychotherapy for cluster B personality disorder patients is the optimal treatment choice in case society is not willing to pay more than €12,274; otherwise, day hospital psychotherapy was the optimal treatment choice. To date, there is no consensus about reasonable willingness-to-pay thresholds, although guidelines have been proposed by the Dutch healthcare authority (Zwaap et al., 2015). We will therefore calculate the probability of each treatment being cost-effective for different willingness-to-pay values. As a result, the optimal treatment choice can be different for different willingness-to-pay values.

Finally, this study is conducted during the COVID-19 pandemic. The COVID-19 pandemic has had significant disruptive effects on society and is related to increased burden of mental health among individuals with mental disorders (Inchausti et al., 2020; Ventura Wurman et al., 2020). Moreover, some authors suggest that patients with severe psychopathology, including BPD, may be especially at risk for symptom deterioration (Alvaro et al., 2020; Ventura Wurman et al., 2020). In addition, some patients will temporarily receive treatment via videoconferencing if face-to-face treatment is restricted in mental healthcare centers. Research on the effectiveness of online individual psychotherapy has found positive effects for several mental health disorders, including PTSD (Liu et al., 2020), anxiety disorders (Berryhill, Halli-Tierney et al., 2019), and depression (Berryhill, Culmer et al., 2019). However, research on the effectiveness of online group psychotherapy is scarce (van Leeuwen et al., 2021; Weinberg, 2020). Consequently, we will control for a potential effect of the COVID-19 pandemic in the analyses.

Conclusion

Specialized evidence-based treatments have been developed and evaluated for BPD, including DBT and ST. However, BPD patients vary widely in their response to treatment, and poor response to one treatment does not imply poor response to another treatment. The selection of the optimal treatment for a particular patient is a daily task of the clinician, but very scant evidence is available to guide these decisions. This study will extend our knowledge of one of the main issues in psychotherapy research; understanding for whom a treatment works and how. As such, this study helps pave the way for an evidence-based personalized medicine for patients with BPD.

REFERENCES

Abdel-Khalek, A. M. (2006). Measuring happiness with a single-item scale. *Social Behavior and Personality: An International Journal, 34*(2), 139–150. https://doi.org/10.2224/sbp.2006.34.2.139

Adams, G., Gulliford, M. C., Ukoumunne, O. C., Eldridge, S., Chinn, S., & Campbell, M. J. (2004). Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of Clinical Epidemiology, 57*(8), 785–794. https://doi.org/10.1016/j.jclinepi.2003.12.013

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(2), 202–212. e7. https://doi.org/10.1016/j.jaac.2011.11.003

Altman, D. G., & Schulz, K. F. (2001). Statistics notes - Concealing treatment allocation in randomised trials. *British Medical Journal, 323*(7310), 446–447. https://doi.org/10.1136/bmj.323.7310.446

Alvaro, F., Navarro, S., Palma, C., Farriols, N., Aliaga, F., Solves, L., Hernandez, M., Anton, M., & Riera, A. (2020). Clinical course and predictors in patients with borderline personality disorder during the COVID-19 outbreak: A 2.5-month naturalistic exploratory study in Spain. *Psychiatry Research, 292*, 113306. https://doi.org/10.1016/j.psychres.2020.113306

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)* (5th ed.). American Psychiatric Publishing.

Arntz, A., Klokman, J., & Sieswerda, S. (2005). An experimental test of the schema mode model of borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 36*(3), 226–239. https://doi.org/10.1016/j.jbtep.2005.05.005

Arntz, A., Mensink, K., Cox, W., Verhoef, R., van Emmerik, A. A. P., & Grasman, R. P. P. P. (2021). Dropout from psychological treatment for borderline personality disorder: a multilevel survival meta-analysis. *Manuscript submitted for publication.*

Arntz, A., van den Hoorn, M., Cornelis, J., Verheul, R., van den Bosch, W., & de Bie, A. (2003). Reliability and validity of the borderline personality disorder severity index. *Journal of Personality Disorders, 17*(1), 45–59. https://doi.org/10.1521/pedi.17.1.45.24053

Arntz, A., Kamphuis, J. H., & Derks, J. L. (2018). *Gestructureerd klinisch interview voor DSM-5 Syndroomstoornissen Vragenlijst (SCID-5-SV)*. Boom.

Arntz, A., Stupar-Rutenfrans, S., Bloo, J., van Dyck, R., & Spinhoven, P. (2015). Prediction of treatment discontinuation and recovery from Borderline Personality Disorder: Results from an RCT comparing Schema Therapy and Transference Focused Psychotherapy. *Behaviour Research and Therapy, 74*, 60–71. https://doi.org/10.1016/j.brat.2015.09.002

Arntz, A., & van Genderen, H. (2020). *Schema Therapy for Borderline Personality Disorder*. Wiley-Blackwell.

Aviram, R. B., Brodsky, B. S., & Stanley, B. (2006). Borderline personality disorder, stigma, and treatment implications. *Harvard Review of Psychiatry, 14*(5), 249–256. https://doi.org/10.1080/10673220600975121

Bamelis, L. L. M., Arntz, A., Wetzelaer, P., Verdoorn, R., & Evers, S. M. A. A. (2015). Economic Evaluation of Schema Therapy and Clarification-Oriented Psychotherapy for Personality Disorders: A Multicenter, Randomized Controlled Trial. *Journal of Clinical Psychiatry, 76*(11), E1432–+. https://doi.org/10.4088/JCP.14m09412

Bardeen, J. R., Fergus, T. A., & Orcutt, H. K. (2012). An examination of the latent structure of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 34*(3), 382–392. https://doi.org/10.1007/s10862-012-9280-y

Barnicot, K. (2020). Evidence-based psychological interventions for borderline personality disorder in the United Kingdom—who falls through the gaps? *Journal of Psychological Therapies, 5*(2), 148–167.

Barnicot, K., Katsakou, C., Bhatti, N., Savill, M., Fearns, N., & Priebe, S. (2012). Factors predicting the outcome of psychotherapy for borderline personality disorder: A systematic review. *Clinical Psychology Review, 32*(5), 400–412. https://doi.org/10.1016/j.cpr.2012.04.004

Bastien, C. H., Vallieres, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine, 2*(4), 297–307. https://doi.org/10.1016/S1389-9457(00)00065-4

Bender, D. S., Dolan, R. T., Skodol, A. E., Sanislow, C. A., Dyck, I. R., McGlashan, T. H., Shea, M. T., Zanarini, M. C., Oldham, J. M., & Gunderson, J. G. (2001). Treatment utilization by patients with personality disorders. *American Journal of Psychiatry, 158*(2), 295–302. https://doi.org/10.1176/appi.ajp.158.2.295

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF. Manual for Administration, Scoring, and Interpretation*. University of Minnesota Press.

Berryhill, M. B., Halli-Tierney, A., Culmer, N., Williams, N., Betancourt, A., King, M., & Ruggles, H. (2019). Videoconferencing psychological therapy and anxiety: a systematic review. *Family Practice, 36*(1), 53–63. https://doi.org/10.1093/fampra/cmy072

Berryhill, M. B., Culmer, N., Williams, N., Halli-Tierney, A., Betancourt, A., Roberts, H., & King, M. (2019). Videoconferencing Psychotherapy and Depression: A Systematic Review. *Telemedicine and E-Health, 25*(6), 435–446. https://doi.org/10.1089/tmj.2018.0058

Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: Prevalence, risk factors, prediction, and prevention. *Journal of Personality Disorders, 18*(3), 226–239. https://doi.org/10.1521/pedi.18.3.226.35445

Bloo, J., Arntz, A., & Schouten, E. (2017). The Borderline Personality Disorder Checklist: Psychometric evaluation and factorial structure in clinical and nonclinical samples. *Roczniki Psychologiczne, 20*(2), 311–336. https://doi.org/10.18290/rpsych.2017.20.2-3en

Bode, K., Vogel, R., Walker, J., & Kroeger, C. (2017). Health care costs of borderline personality disorder and matched controls with major depressive disorder: a comparative study based on anonymized claims data. *European Journal of Health Economics, 18*(9), 1125–1135. https://doi.org/10.1007/s10198-016-0858-2

Bohart, A. C., & Wade, A. G. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 219–257). Wiley.

Brazier, J. (2010). Is the EQ–5D fit for purpose in mental health? *The British Journal of Psychiatry, 197*(5), 348–349. https://doi.org/10.1192/bjp.bp.110.082453

Brettschneider, C., Riedel-Heller, S., & Koenig, H. (2014). A Systematic Review of Economic Evaluations of Treatments for Borderline Personality Disorder. *Plos One, 9*(9), e107748. https://doi.org/10.1371/journal.pone.0107748

Buckman, J. E. J., Cohen, Z. D., O'Driscoll, C., Fried, E. I., Saunders, R., Ambler, G., DeRubeis, R. J., Gilbody, S., Hollon, S., & Kendrick, T. (2020). Predicting prognosis for adults with depression using individual symptom data: a comparison of modelling approaches, 1–11. https://doi.org/10.31219/osf.io/xkwdc

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. https://doi.org/10.1016/j.cpr.2013.08.003

Burgess, A. M., Frost, R. O., & DiBartolo, P. M. (2016). Development and Validation of the Frost Multidimensional Perfectionism Scale-Brief. *Journal of Psychoeducational Assessment, 34*(7), 620–633. https://doi.org/10.1177/0734282916651359

Carbonari, J. P., DiClemente, C. C., & Zweben, A. (1994). A readiness to change scale: Its development, validation, and usefulness. Paper presented at the *Annual Meeting of the Association for Advancement of Behavior Therapy.*

Carbonari, J. P., & DiClemente, C. C. (2000). Using transtheoretical model profiles to differentiate levels of alcohol abstinence success. *Journal of Consulting and Clinical Psychology, 68*(5), 810. https://doi.org/10.1037/0022-006X.68.5.810

Cecil, H., Stanley, M. A., Carrion, P. G., & Swann, A. (1995). Psychometric Properties of the MSPSS and NOS in Psychiatric Outpatients. *Journal of Clinical Psychology, 51*(5), 593–602. https://doi.org/10.1002/1097-4679(199509)51:5%3C593::aid-jclp2270510503%3E3.0.co;2-w

Chakhssi, F., Zoet, J. M., Oostendorp, J. M., Noordzij, M. L., & Sommers-Spijkerman, M. (2019). Effect of psychotherapy for borderline personality disorder on quality of life: a systematic review and meta-analysis. *Journal of Personality Disorders,* 1–15. https://doi.org/10.1521/pedi_2019_33_439

Chan, A., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gotzsche, P. C., Krleza-Jeric, K., Hrobjartsson, A., Mann, H., Dickersin, K., Berlin, J. A., Dore, C. J., Parulekar, W. R., Summerskill, W. S. M., Groves, T., Schulz, K. F., Sox, H. C., Rockhold, F. W., Rennie, D., & Moher, D. (2013). SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine, 158*(3), 200–207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *International Review of Psychiatry, 23*(4), 318–327. https://doi.org/10.3109/09540261.2011.606803

Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 194–226). John Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment Selection in Depression. *Annual Review of Clinical Psychology, Vol 14, 14*, 209–236. https://doi.org/10.1146/annurev-clinpsy-050817084746

Cohen, Z. D., Kim, T. T., Van, H. L., Dekker, J. J. M., & Driessen, E. (2020). A demonstration of a multi-method variable selection approach for treatment selection: Recommending cognitive-behavioral versus psychodynamic therapy for mild to moderate adult depression. *Psychotherapy Research, 30*(2), 137–150. https://doi.org/10.1080/10503307.2018.1563312

Comtois, K. A., Elwood, L., Holdcraft, L. C., Smith, W. R., & Simpson, T. L. (2007). Effectiveness of dialectical behavior therapy in a community mental health center. *Cognitive and Behavioral Practice, 14*(4), 406–414. https://doi.org/10.1016/j.cbpra.2006.04.023

de Beurs, E., & Zitman, F. (2006). De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief [Reliability and validity of a practical alternative to the SCL 90]. *Maandblad Geestelijke Volksgezondheid, 61*, 120–141.

de Beurs, E. (2011). *Brief Symptom Inventory (BSI) en BSI 18. Handleiding 2011 [Brief Symptom Inventory and BSI 18. Manual 2011]*. PITS.

de Meulemeester, C., Vansteelandt, K., Luyten, P., & Lowyck, B. (2018). Mentalizing as a Mechanism of Change in the Treatment of Patients With Borderline Personality Disorder: A Parallel Process Growth Modeling Approach. *Personality Disorders-Theory Research and Treatment, 9*(1), 22–29. https://doi.org/10.1037/per0000256

Derogatis, L. R., & Melisaratos, N. (1983). The brief symptom inventory: an introductory report. *Psychological Medicine, 13*(3), 595–605. https://doi.org/10.1017/S0033291700048017

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: Translating Research on Prediction into Individualized Treatment Recommendations. A Demonstration. *Plos One, 9*(1), e83875. https://doi.org/10.1371/journal.pone.0083875

Di Stephano, J. (2003). How much power is enough? Against the development of an arbitrary convention for statistical power calculations. *Functional Ecology, 17*(5), 707–709. https://doi.org/10.1046/j.1365-2435.2003.00782.x

Dickhaut, V., & Arntz, A. (2014). Combined group and individual schema therapy for borderline personality disorder: A pilot study. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(2), 242–251. https://doi.org/10.1016/j.jbtep.2013.11.004

DiClemente, C. C., Bellino, L. E., & Neavins, T. M. (1999). Motivation for change and alcoholism treatment. *Alcohol Research & Health, 23*(2), 86–92.

DiClemente, C. C., & Hughes, S. O. (1990). Stages of change profiles in outpatient alcoholism treatment. *Journal of Substance Abuse, 2*(2), 217–235. https://doi.org/10.1016/S0899-3289(05)80057-4

Dimeff, L. A., Rizvi, S. L., & Koerner, K. (2020). *Dialectical Behavior Therapy in Clinical Practice. Applications across Disorders and Settings* (2nd ed.). Guilford Press.

Dixon, L. J., & Linardon, J. (2020). A systematic review and meta-analysis of dropout rates from dialectical behaviour therapy in randomized controlled trials. *Cognitive Behaviour Therapy, 49*(3), 181–196. https://doi.org/10.1080/16506073.2019.1620324

Dixon-Gordon, K. L., Whalen, D. J., Layden, B. K., & Chapman, A. L. (2015). A Systematic Review of Personality Disorders and Health Outcomes. *Canadian Psychology-Psychologie Canadienne, 56*(2), 168–190. https://doi.org/10.1037/cap0000024

Dozois, D. J., Westra, H. A., Collins, K. A., Fung, T. S., & Garry, J. K. (2004). Stages of change in anxiety: Psychometric properties of the University of Rhode Island Change Assessment (URICA) scale. *Behaviour Research and Therapy, 42*(6), 711–729. https://doi.org/10.1016/S0005-7967(03)00193-1

Drieschner, K. H., & Boomsma, A. (2008). The treatment motivation scales for forensic outpatient treatment (TMS-F) construction and psychometric evaluation. *Assessment, 15*(2), 224–241. https://doi.org/10.1177/1073191107311650

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford university press.

Eldridge, S. M., Ashby, D., Feder, G. S., Rudnicka, A. R., & Ukoumunne, O. C. (2004). Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clinical Trials, 1*(1), 80–90. https://doi.org/10.1191/1740774504cn006rr

El-Gabalawy, R., Katz, L. Y., & Sareen, J. (2010). Comorbidity and Associated Severity of Borderline Personality Disorder and Physical Health Conditions in a Nationally Representative Sample. *Psychosomatic Medicine, 72*(7), 641–647. https://doi.org/10.1097/PSY.0b013e3181e10c7b

Ellison, W. D., Rosenstein, L. K., Morgan, T. A., & Zimmerman, M. (2018). Community and Clinical Epidemiology of Borderline Personality Disorder. *Psychiatric Clinics of North America, 41*(4), 561–573. https://doi.org/10.1016/j.psc.2018.07.008

Farrell, J. M., & Shaw, I. A. (2010). Schema therapy groups for borderline personality disorder patients: The best of both worlds of group psychotherapy. In E. Roediger, & G. Jacobs (Eds.), *Fortschritte der Schematherapie. (Advances in Schema Therapy).* Hogrefe.

Farrell, J. M., Shaw, I. A., & Webber, M. A. (2009). A schema-focused approach to group psychotherapy for outpatients with borderline personality disorder: A randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry, 40*(2), 317–328. https://doi.org/10.1016/j.jbtep.2009.01.002

Farrell, J. M., & Shaw, I. A. (2012). *Group schema therapy for borderline personality disorder: A step-by-step treatment manual with patient workbook*. John Wiley & Sons.

Fassbinder, E., Assmann, N., Schaich, A., Heinecke, K., Wagner, T., Sipos, V., Jauch-Chara, K., Hüppe, M., Arntz, A., & Schweiger, U. (2018). PRO* BPD: effectiveness of outpatient treatment programs for borderline personality disorder: a comparison of Schema therapy and dialectical behavior therapy: study protocol for a randomized trial. *BMC Psychiatry, 18*(1), 341. https://doi.org/10.1186/s12888-018-1905-6

Fassbinder, E., Schweiger, U., Martius, D., Brand-de Wilde, O., & Arntz, A. (2016). Emotion Regulation in Schema Therapy and Dialectical Behavior Therapy. *Frontiers in Psychology, 7*, 1373. https://doi.org/10.3389/fpsyg.2016.01373

Feaster, D. J., Mikulich-Gilbertson, S., & Brincks, A. M. (2011). Modeling site effects in the design and analysis of multi-site trials. *The American Journal of Drug and Alcohol Abuse, 37*(5), 383–391. https://doi.org/10.3109/00952990.2011.600386

Feigenbaum, J. D., Fonagy, P., Pilling, S., Jones, A., Wildgoose, A., & Bebbington, P. E. (2012). A real-world study of the effectiveness of DBT in the UK National Health Service. *British Journal of Clinical Psychology, 51*, 121–141. https://doi.org/10.1111/j.2044-8260.2011.02017.x

First, M. B., Williams, J., Benjamin, L. S., & Spitzer, R. L. (2015). *User's Guide for the SCID-5-PD (Structured Clinical Interview for DSM-5 Personality Disorder)*. American Psychiatric Association.

First, M. B., Williams, J. B. W., Benjamin, L. S., & Spitzer, R. L. (2016). *Structured Clinical Interview for DSM-5 Screening Personality Questionnaire (SCID-5-SPQ)*. American Psychiatric Association.

First, M. B., Williams, J., Karg, R. S., & Spitzer, R. L. (2018). *Gestructureerd klinisch interview voor DSM-5 Syndroomstoornissen (SCID-5-S).* Nederlandse vertaling van Structured Clinical Interview for DSM-5® Disorders-Clinician Version (SCID-5-CV) en User's Guide to Structured Clinical Interview for DSM-5® Disorders-Clinician Version en delen van Structured Clinical Interview for DSM-5® Disorders-Research Version (SCID-5-RV), first edition. Boom.

First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2016). *Structured Clinical Interview for DSM-5 Disorders, Clinician Version (SCID-5-CV)*. American Psychiatric Association.

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y., Warren, F., Howard, S., Ghinai, R., Fearon, P., & Lowyck, B. (2016). Development and Validation of a Self-Report Measure of Mentalizing: The Reflective Functioning Questionnaire. *Plos One, 11*(7), e0158678. https://doi.org/10.1371/journal.pone.0158678

Forster, C., Berthollier, N., & Rawlinson, D. (2014). A systematic review of potential mechanisms of change in psychotherapeutic interventions for personality disorder. *Journal of Psychology & Psychotherapy, 4*(133). https://doi.org/10.4172/2161-0487.1000133

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350–365. https://doi.org/10.1037/0022-3514.78.2.350

Fraley, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The Experiences in Close Relationships-Relationship Structures Questionnaire: A Method for Assessing Attachment Orientations Across Relationships. *Psychological Assessment, 23*(3), 615–625. https://doi.org/10.1037/a0022898

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software, 33*(1), 1–22. https://doi.org/10.18637/jss.v033.i01

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., Suzuki, N., & Watson, D. (2014). The brief experiential avoidance questionnaire: development and initial validation. *Psychological Assessment, 26*(1), 35–45. https://doi.org/10.1037/a0034473

Giesen-Bloo, J. H., Wachters, L. M., Schouten, E., & Arntz, A. (2010). The Borderline Personality Disorder Severity Index-IV: Psychometric evaluation and dimensional structure. *Personality and Individual Differences, 49*(2), 136–141. https://doi.org/10.1016/j.paid.2010.03.023

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., Kremers, I., Nadort, M., & 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. https://doi.org/10.1001/archpsyc.63.6.649

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Grant, A. M., Franklin, J., & Langford, P. (2002). The self-reflection and insight scale: A new measure of private self-consciousness. *Social Behavior and Personality, 30*(8), 821–835. https://doi.org/10.2224/sbp.2002.30.8.821

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41–54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gunderson, J. (2008). Borderline personality disorder: An overview. *Social Work in Mental Health, 6*(1-2), 5–12. https://doi.org/10.1300/J200v06n01_02

Gunderson, J. G., Herpertz, S. C., Skodol, A. E., Torgersen, S., & Zanarini, M. C. (2018). Borderline personality disorder. *Nature Reviews Disease Primers, 4*, 18029. https://doi.org/10.1038/nrdp.2018.29

Gunderson, J. G., Stout, R. L., McGlashan, T. H., Shea, T., Morey, L. C., Grilo, C. M., Zanarini, M. C., Yen, S., Markowitz, J. C., Sanislow, C., Ansell, E., Pinto, A., & Skodol, A. E. (2011). Ten-Year Course of Borderline Personality Disorder Psychopathology and Function From the Collaborative Longitudinal Personality Disorders Study. *Archives of General Psychiatry, 68*(8), 827–837. https://doi.org/10.1001/archgenpsychiatry.2011.37

Hafkenscheid, A., Duncan, B. L., & Miller, S. D. (2010). The outcome and session rating scales. A cross-cultural examination of the psychometric properties of the Dutch translation. *Journal of Brief Therapy, 7*(1), 1–12.

Hakkaart-van Roijen, L., Van der Linden, N., Bouwmans, C., Kanters, T., & Tan, S. S. (2015). *Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg [Costing manual: methodology of costing research and refeence prices for economic evaluations in healthcare].* Diemen, the Netherlands: Zorginstituut Nederland.

Hallion, L. S., Steinman, S. A., Tolin, D. F., & Diefenbach, G. J. (2018). Psychometric properties of the Difficulties in Emotion Regulation Scale (DERS) and its short forms in adults with emotional disorders. *Frontiers in Psychology, 9*, 539. https://doi.org/10.3389/fpsyg.2018.00539

Hallquist, M. N., & Pilkonis, P. A. (2012). Refining the Phenotype of Borderline Personality Disorder: Diagnostic Criteria and Beyond. *Personality Disorders-Theory Research and Treatment, 3*(3), 228–246. https://doi.org/10.1037/a0027953

Hamann, J., Leucht, S., & Kissling, W. (2003). Shared decision making in psychiatry. *Acta Psychiatrica Scandinavica, 107*(6), 403–409. https://doi.org/10.1034/j.1600-0447.2003.00130.x

Hamburg, M. A., & Collins, F. S. (2010). The Path to Personalized Medicine. *New England Journal of Medicine, 363*(4), 301–304. https://doi.org/10.1056/NEJMp1006304

Hao, F., Tan, W., Jiang, L., Zhang, L., Zhao, X., Zou, Y., Hu, Y., Luo, X., Jiang, X., & McIntyre, R. S. (2020). Do psychiatric patients experience more psychiatric symptoms during COVID-19 pandemic and lockdown? A case-control study with service and research implications for immunopsychiatry. *Brain, Behavior, and Immunity, 87*, 100–106. https://doi.org/10.1016/j.bbi.2020.04.069

Harned, M. S., Schmidt, S. C. & Korslund, K. E. (2021). *The Dialectical Behavior Therapy Adherence Checklist for Individual Therapy (DBT AC-I).* https://www.dbtadherence.com/.

Harned, M. S., Korslund, K. E., Schmidt, S. C., & Gallop, R. J. (2021). The Dialectical Behavior Therapy Adherence Coding Scale (DBT ACS): Psychometric properties. *Psychological Assessment, 33(6),* 552–561. https://doi.org/10.1037/pas0000999

Heath, L. M., Laporte, L., Paris, J., Hamdullahpur, K., & Gill, K. J. (2018). Substance Misuse is Associated with Increased Psychiatric Severity among Treatment-Seeking Individuals with Borderline Personality Disorder. *Journal of Personality Disorders, 32*(5), 694–708. https://doi.org/10.1521/pedi_2017_31_307

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

Hoch, J. S., Briggs, A. H., & Willan, A. R. (2002). Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Economics, 11*(5), 415–430. https://doi.org/10.1002/hec.678

Hodgson, T. A., & Meiners, M. R. (1982). Cost-of-illness methodology: a guide to current practices and procedures. *The Milbank Memorial Fund Quarterly. Health and Society, 60*, 429–462. https://doi.org/10.2307/3349801

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences, 45*(8), 750–755. https://doi.org/10.1016/j.paid.2008.08.004

Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., Roiser, J. P., Bockting, C. L. H., O'Connor, R. C., Shafran, R., Moulds, M. L., & Craske, M. G. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *Lancet Psychiatry, 5*(3), 237–286.

Huibers, M. (2015). Voorbij het oordeel van de dodo. *Tijdschrift Voor Psychotherapie, 41*(3), 174–186. https://doi.org/10.1007/s12485-015-0027-6

Huibers, M. J. H., Lorenzo-Luaces, L., Cuijpers, P., & Kazantzis, N. (2021). On the Road to Personalized Psychotherapy: A Research Agenda Based on Cognitive Behavior Therapy for Depression. *Frontiers in Psychiatry, 11*, 607508. https://doi.org/10.3389/fpsyt.2020.607508

Huibers, M. J., Cohen, Z. D., Lemmens, L. H., Arntz, A., Peeters, F. P., Cuijpers, P., & DeRubeis, R. J. (2015). Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the personalized advantage index approach. *PloS One, 10*(11), e0140771. https://doi.org/10.1371/journal.pone.0140771

Inchausti, F., MacBeth, A., Hasson-Ohayon, I., & Dimaggio, G. (2020). Psychological Intervention and COVID-19: What We Know So Far and What We Can Do. *Journal of Contemporary Psychotherapy, 50*(4), 243–250. https://doi.org/10.1007/s10879-020-09460-w

Jacob, G. A., & Arntz, A. (2013). Schema Therapy for Personality Disorders-A Review. *International Journal of Cognitive Therapy, 6*(2), 171–185. https://doi.org/10.1521/ijct.2013.6.2.171

Janssen, M. F., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P., & Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research, 22*(7), 1717–1727. https://doi.org/10.1007/s11136-012-0322-4

Kamphuis, J. H., & Emmelkamp, P. (1998). Crime-related trauma: Psychological distress in victims of bankrobbery. *Journal of Anxiety Disorders, 12*(3), 199–208. https://doi.org/10.1016/S0887-6185(98)00009-7

Kanters, T. A., Bouwmans, C. A. M., van der Linden, N., Tan, S. S., & Hakkaart-van Roijen, L. (2017). Update of the Dutch manual for costing studies in health care. *Plos One, 12*(11), e0187477. https://doi.org/10.1371/journal.pone.0187477

Kapelner, A., & Bleich, J. (2016). bartMachine: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software, 70*(4), 1–40. https://doi.org/10.18637/jss.v070.i04

Katsakou, C., Marougka, S., Barnicot, K., Savill, M., White, H., Lockwood, K., & Priebe, S. (2012). Recovery in Borderline Personality Disorder (BPD): A Qualitative Study of Service Users' Perspectives. *Plos One, 7*(5), e36517. https://doi.org/10.1371/journal.pone.0036517

Kaufman, E. A., Xia, M., Fosco, G., Yaptangco, M., Skidmore, C. R., & Crowell, S. E. (2016). The difficulties in emotion regulation scale short form (DERS-SF): validation and replication in adolescent and adult samples. *Journal of Psychopathology and Behavioral Assessment, 38*(3), 443–455. https://doi.org/10.1007/s10862-015-9529-3

Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry, 44*(8), 1116–1129. https://doi.org/10.1111/1469-7610.00195

Kazdin, A. E. (2014). Evidence-based psychotherapies I: qualifiers and limitations in what we know. *South African Journal of Psychology, 44*(4), 381–403. https://doi.org/10.1177/0081246314533750

Keefe, J. R., Kim, T. T., DeRubeis, R. J., Streiner, D. L., Links, P. S., & McMain, S. F. (2020). Treatment selection in borderline personality disorder between dialectical behavior therapy and psychodynamic psychiatric management. *Psychological Medicine, 24*, 1–9. https://doi.org/10.1017/S0033291720000550

Kellogg, S. H., & Young, J. E. (2006). Schema therapy for borderline personality disorder. *Journal of Clinical Psychology, 62*(4), 445–458. https://doi.org/10.1002/jclp.20240

Kenny, D. A. (2016). *Mediation.* http://davidakenny.net/cm/mediate.htm.

Kenny, D. A. (2017). MedPower: An interactive tool for the estimation of power in tests of mediation [computer software]. https://davidakenny.shinyapps.io/MedPower/.

Kenny, D. A., & Judd, C. M. (2014). Power Anomalies in Testing Mediation. *Psychological Science, 25*(2), 334–339. https://doi.org/10.1177/0956797613502676

Kernberg, O. F., & Caligor, E. (2005). A psychoanalytic theory of personality disorders. In M. F. Lenzenweger, & J. F. Clarkin (Eds.), *Major theories of personality disorder* (pp. 114–156). Guilford Press.

Kessler, R. C. (2018). The potential of predictive analytics to provide clinical decision support in depression treatment planning. *Current Opinion in Psychiatry, 31*(1), 32–39. https://doi.org/10.1097/YCO.0000000000000377

Kjaer, J. N., Biskin, R., Vestergaard, C., & Munk-Jorgensen, P. (2020). All-Cause Mortality of Hospital-Treated Borderline Personality Disorder: a Nationwide Cohort Study. *Journal of Personality Disorders, 34*(6), 723–735. https://doi.org/10.1521/pedi_2018_32_403

Koons, C. R. (2011). The Role of the Team in Managing Telephone Consultation in Dialectical Behavior Therapy: Three Case Examples. *Cognitive and Behavioral Practice, 18*(2), 168–177. https://doi.org/10.1016/j.cbpra.2009.10.008

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877–883. https://doi.org/10.1001/archpsyc.59.10.877

Kraemer, H. C. (2013). Discovering, comparing, and combining moderators of treatment on outcome after randomized clinical trials: a parametric approach. *Statistics in Medicine, 32*(11), 1964–1973. https://doi.org/10.1002/sim.5734

Krakow, B., Hollifield, M., Schrader, R., Koss, M., Tandberg, D., Lauriello, J., McBride, L., Warner, T. D., Cheng, D., & Edmond, T. (2000). A controlled study of imagery rehearsal for chronic nightmares in sexual assault survivors with PTSD: a preliminary report. *Journal of Traumatic Stress, 13*(4), 589–609. https://doi.org/10.1023/A:1007854015481

Krawitz, R., & Miga, E. M. (2019). Cost-effectiveness of dialectical behaviour therapy for borderline personality disorder. In M. A. Swales (Ed.), *The Oxford handbook of dialectical behaviour therapy*. Oxford University Press.

Lakens, D. (2021). Sample size justification. https://doi.org/10.31234/osf.io/9d3yf

Lancee, J., van Straten, A., Morina, N., Kaldo, V., & Kamphuis, J. H. (2016). Guided online or face-to-face cognitive behavioral treatment for insomnia: a randomized wait-list controlled trial. *Sleep, 39*(1), 183–191. https://doi.org/10.5665/sleep.5344

Lariviere, N., Couture, E., Blackburn, C., Carbonneau, M., Lacombe, C., Schinck, S., David, P., & St-Cyr-Tribble, D. (2015). Recovery, as Experienced by Women with Borderline Personality Disorder. *Psychiatric Quarterly, 86*(4), 555–568. https://doi.org/10.1007/s11126-015-9350-x

Laurenssen, E. M. P., Eeren, H. V., Kikkert, M. J., Peen, J., Westra, D., Dekker, J. J. M., & Busschbach, J. J. V. (2016). The burden of disease in patients eligible for mentalization-based treatment (MBT): quality of life and costs. *Health and Quality of Life Outcomes, 14*, 145. https://doi.org/10.1186/s12955-016-0538-z

Leichsenring, F., Leibing, E., Kruse, J., New, A. S., & Leweke, F. (2011). Borderline personality disorder. *The Lancet, 377*(9759), 74–84. https://doi.org/10.1016/S0140-6736(10)61422-5

Lemmens, L. H. J. M., Mueller, V. N. L. S., Arntz, A., & Huibers, M. J. H. (2016). Mechanisms of change in psychotherapy for depression: An empirical update and evaluation of research aimed at identifying psychological mediators. *Clinical Psychology Review, 50*, 95–107. https://doi.org/10.1016/j.cpr.2016.09.004

Lemmens, L. H. J. M., Galindo-Garre, F., Arntz, A., Peeters, F., Hollon, S. D., DeRubeis, R. J., & Huibers, M. J. H. (2017). Exploring mechanisms of change in cognitive therapy and interpersonal psychotherapy for adult depression. *Behaviour Research and Therapy, 94*, 81–92. https://doi.org/10.1016/j.brat.2017.05.005

Lenzenweger, M. F., Clarkin, J. F., Yeomans, F. E., Kernberg, O. F., & Levy, K. N. (2008). Refining the borderline personality disorder phenotype through finite mixture modeling: Implications for classification. *Journal of Personality Disorders, 22*(4), 313–331. https://doi.org/10.1521/pedi.2008.22.4.313

Lequerica, A., Rapport, L., Axelrod, B. N., Telmet, K., & Whitman, R. D. (2002). Subjective and objective assessment methods of mental imagery control: Construct validation of self-report measures. *Journal of Clinical and Experimental Neuropsychology, 24*(8), 1103–1116. https://doi.org/10.1076/jcen.24.8.1103.8370

Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in Psychotherapy Outcome Research: Separating Association From Bias. *Clinical Psychology-Science and Practice, 16*(1), 54–65. https://doi.org/10.1111/j.1468-2850.2009.01143.x

Linehan, M. M. (1993a). *Cognitive-behavioral therapy of borderline personality disorder*. Guilford press.

Linehan, M. M. (1993b). *Skills training manual for treating borderline personality disorder.* Guilford press.

Linehan, M. M. (1997). Validation and psychotherapy. In A. Bohart, & L. Greenberg (Eds.), *Empathy reconsidered: New directions in psychotherapy.* (pp. 353–392). American Psychological Association.

Linehan, M. M. (2015). *DBT Skills Training Manual* (2nd ed.). Guilford Publications.

Links, P. S., Heslegrave, R. J., Mitton, J. E., Vanreekum, R., & Patrick, J. (1995). Borderline Personality-Disorder and Substance-Abuse - Consequences of Comorbidity. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 40*(1), 9–14. https://doi.org/10.1177/070674379504000105

Liu, L., Thorp, S. R., Moreno, L., Wells, S. Y., Glassman, L. H., Busch, A. C., Zamora, T., Rodgers, C. S., Allard, C. B., Morland, L. A., & Agha, Z. (2020). Videoconferencing psychotherapy for veterans with PTSD: Results from a randomized controlled non-inferiority trial. *Journal of Telemedicine and Telecare, 26*(9), 507–519. https://doi.org/10.1177/1357633X19853947

Lobbestael, J., van Vreeswijk, M., Spinhoven, P., Schouten, E., & Arntz, A. (2010). Reliability and validity of the short Schema Mode Inventory (SMI). *Behavioural and Cognitive Psychotherapy, 38*(4), 437–458. https://doi.org/10.1017/S1352465810000226

Luedtke, A., Sadikova, E., & Kessler, R. C. (2019). Sample Size Requirements for Multivariate Models to Predict Between-Patient Differences in Best Treatments of Major Depressive Disorder. *Clinical Psychological Science, 7*(3), 445–461. https://doi.org/10.1177/2167702618815466

Lukat, J., Margraf, J., Lutz, R., van der Veld, W. M., & Becker, E. S. (2016). Psychometric properties of the positive mental health scale (PMH-scale). *BMC Psychology, 4*(1), 1–14. https://doi.org/10.1186/s40359-016-0111-x

Lundin, A., Kosidou, K., & Dalman, C. (2019). Measuring Autism Traits in the Adult General Population with the Brief Autism-Spectrum Quotient, AQ-10: Findings from the Stockholm Public Health Cohort. *Journal of Autism and Developmental Disorders, 49*(2), 773–780. https://doi.org/10.1007/s10803-018-3749-9

Lynch, T. R., Chapman, A. L., Rosenthal, M. Z., Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology, 62*(4), 459–480. https://doi.org/10.1002/jclp.20243

Lynch, T. R., Trost, W. T., Salsman, N., & Linehan, M. M. (2007). Dialectical behavior therapy for borderline personality disorder. *Annual Review of Clinical Psychology, 3*, 181–205. https://doi.org/10.1146/annurev.clinpsy.2.022305.095229

McConnaughy, E. A., Prochaska, J. O., & Velicer, W. F. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research & Practice, 20*(3), 368–375. https://doi.org/10.1037/h0090198

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. https://doi.org/10.1176/appi.ajp.2012.11091416

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* University of Minnesota Press.

Meuldijk, D., McCarthy, A., Bourke, M. E., & Grenyer, B. F. S. (2017). The value of psychological treatment for borderline personality disorder: Systematic review and cost offset analysis of economic evaluations. *Plos One, 12*(3), e0171592. https://doi.org/10.1371/journal.pone.0171592

Miller, S. J., & Binder, J. L. (2002). The effects of manual-based training on treatment fidelity and outcome: A review of the literature on adult individual psychotherapy. *Psychotherapy, 39*(2), 184–198. https://doi.org/10.1037/0033-3204.39.2.184

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*(2), 91–100.

Mohammadkhani, P., Forouzan, A. S., Hooshyari, Z., & Abasi, I. (2020). Psychometric Properties of Persian Version of Structured Clinical Interview for DSM-5-Research Version (SCID-5-RV): A Diagnostic Accuracy Study. *Iranian Journal of Psychiatry and Behavioral Sciences, 14*(2). https://doi.org/10.5812/ijpbs.100930

Morandotti, N., Brondino, N., Merelli, A., Boldrini, A., De Vidovich, G. Z., Ricciardo, S., Abbiati, V., Ambrosi, P., Caverzasi, E., Fonagy, P., & Luyten, P. (2018). The Italian version of the Reflective Functioning Questionnaire: Validity data for adults and its association with severity of borderline personality disorder. *Plos One, 13*(11), e0206433. https://doi.org/10.1371/journal.pone.0206433

Nadort, M., Arntz, A., Smit, J. H., Giesen-Bloo, J., Eikelenboom, M., Spinhoven, P., van Asselt, T., Wensing, M., & van Dyck, R. (2009). Implementation of outpatient schema therapy for borderline personality disorder: study design. *BMC Psychiatry, 9*, 64. https://doi.org/10.1186/1471-244X-9-64

Neacsiu, A. D., Rizvi, S. L., Vitaliano, P. P., Lynch, T. R., & Linehan, M. M. (2010). The Dialectical Behavior Therapy Ways of Coping Checklist: Development and Psychometric Properties. *Journal of Clinical Psychology, 66*(6), 563–582. https://doi.org/10.1002/jclp.20685

Nijenhuis, E., Van der Hart, O., & Kruger, K. (2002). The psychometric characteristics of the Traumatic Experiences Checklist (TEC): First findings among psychiatric outpatients. *Clinical Psychology & Psychotherapy, 9*(3), 200–210. https://doi.org/10.1002/cpp.332

Odou, N., & Vella-Brodrick, D. A. (2013). The Efficacy of Positive Psychology Interventions to Increase Well-Being and the Role of Mental Imagery Ability. *Social Indicators Research, 110*(1), 111–129. https://doi.org/10.1007/s11205-011-9919-1

Orlinsky, D. E., Rönnestad, M. H., & Willutzki, U. (2004). Fifty years of process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 307–389). Wiley.

Osório, F. L., Loureiro, S. R., Hallak, J. E. C., Machado-de-Sousa, J. P., Ushirohira, J. M., Baes, C. V., Apolinario, T. D., Donadon, M. F., Bolsoni, L. M., & Guimarães, T. (2019). Clinical validity and intrarater and test–retest reliability of the Structured Clinical Interview for DSM-5–Clinician Version (SCID-5-CV). *Psychiatry and Clinical Neurosciences, 73*(12), 754–760. https://doi.org/10.1111/pcn.12931

O'Toole, S. K., Diddy, E., & Kent, M. (2012). Mindfulness and Emotional Well-being in Women with Borderline Personality Disorder. *Mindfulness, 3*(2), 117–123. https://doi.org/10.1007/s12671-011-0085-y

Paris, J. (2019). Suicidality in Borderline Personality Disorder. *Medicina-Lithuania, 55*(6). https://doi.org/10.3390/medicina55060223

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*(6), 829–841. https://doi.org/10.1037/0022-006X.75.6.829

Pérez-Fabello, M. J., & Campos, A. (2004). Factor structure and internal consistency of the Spanish version of the Gordon Test of Visual Imagery Control. *Psychological Reports, 94*(3), 761–766. https://doi.org/10.2466/pr0.94.3.761-766

Pocock, S. J., & Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics, 31*(1), 103–115.

Pompili, M., Girardi, P., Ruberto, A., & Tatarelli, R. (2005). Suicide in borderline personality disorder: A meta-analysis. *Nordic Journal of Psychiatry, 59*(5), 319–324. https://doi.org/10.1080/08039480500320025

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A General Multilevel SEM Framework for Assessing Multilevel Mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

Qualtrics. (2019). Qualtrics [computer software]. Provo, UT: Qualtrics.

Richardson, A. (1969). *Mental imagery*. Springer.

Roberts, C., & Stark, P. (2008). Readiness for self-directed change in professional behaviours: factorial validation of the Self-reflection and Insight Scale. *Medical Education, 42*(11), 1054–1063. https://doi.org/10.1111/j.1365-2923.2008.03156.x

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry, 6*(3), 412–415. https://doi.org/10.1111/j.1939-0025.1936.tb05248.x

Royal College of Psychiatrists. (2019). *PS06/19 Services for People Diagnosed with Personality Disorder.* https://www.rcpsych.ac.uk/docs/default-source/improvingcare/better-mh-policy/position-statements/ps01_20.pdf?sfvrsn=85af7fbc_2.

Rudge, S., Feigenbaum, J. D., & Fonagy, P. (2020). Mechanisms of change in dialectical behaviour therapy and cognitive behaviour therapy for borderline personality disorder: a critical review of the literature. *Journal of Mental Health, 29*(1), 92–102. https://doi.org/10.1080/09638237.2017.1322185

Safer, D. L., Telch, C. F., & Chen, E. Y. (2009). *Dialectical behavior therapy for binge eating and bulimia*. Guilford Press.

Sanderman, R. (1988). *Life events, mediating variables and psychological distress: a longitudinal study*. University of Groningen.

Schmand, B., Lindeboom, J., & Van Harskamp, F. (1992). *De nederlandse leestest voor volwassenen [the dutch adult reading test]*. Swets & Zeitlinger.

Schmand, B. A., Bakker, D., Saan, R. J., & Louman, J. (1991). De Nederlandse Leestest voor Volwassenen: een maat voor het premorbide intelligentieniveau [The Dutch Adult Reading Test: A measure of premorbid intelligence]. *Tijdschrift Voor Gerontologie En Geriatrie, 22*, 15–19.

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining Power and Sample Size for Simple and Complex Mediation Models. *Social Psychological and Personality Science, 8*(4), 379–386. https://doi.org/10.1177/1948550617715068

Scott, N. W., McPherson, G. C., Ramsay, C. R., & Campbell, M. K. (2002). The method of minimization for allocation to clinical trials: a review. *Controlled Clinical Trials, 23*(6), 662–674. https://doi.org/10.1016/S0197-2456(02)00242-8

Sempertegui, G. A., Karreman, A., Arntz, A., & Bekker, M. H. J. (2013). Schema therapy for borderline personality disorder: A comprehensive review of its empirical foundations, effectiveness and implementation possibilities. *Clinical Psychology Review, 33*(3), 426–447. https://doi.org/10.1016/j.cpr.2012.11.006

Shankman, S. A., Funkhouser, C. J., Klein, D. N., Davila, J., Lerner, D., & Hee, D. (2018). Reliability and validity of severity dimensions of psychopathology assessed using the Structured Clinical Interview for DSM-5 (SCID). *International Journal of Methods in Psychiatric Research, 27*(1), e1590. https://doi.org/10.1002/mpr.1590

Sheehan, L., Nieweglowski, K., & Corrigan, P. (2016). The Stigma of Personality Disorders. *Current Psychiatry Reports, 18*(1), 11. https://doi.org/10.1007/s11920-015-0654-1

Simms, L. J., Goldberg, L. R., Roberts, J. E., Watson, D., Welte, J., & Rotterman, J. H. (2011). Computerized Adaptive Assessment of Personality Disorder: Introducing the CAT-PD Project. *Journal of Personality Assessment, 93*(4), 380–389. https://doi.org/10.1080/00223891.2011.577475

Skodol, A. E., Gunderson, J. G., Pfohl, B., Widiger, T. A., Livesley, W. J., & Siever, L. J. (2002). The borderline diagnosis I: Psychopathology comorbidity, and personality structure. *Biological Psychiatry, 51*(12), 936–950. https://doi.org/10.1016/S0006-3223(02)01324-0

Skodol, A. E. (2018). Impact of personality pathology on psychosocial functioning. *Current Opinion in Psychology, 21*, 33–38. https://doi.org/10.1016/j.copsyc.2017.09.006

Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., Grilo, C. M., Shea, M. T., Zanarini, M. C., & Morey, L. C. (2002). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry, 159*(2), 276–283.

Smits, D., Luyckx, K., Smits, D., Stinckens, N., & Claes, L. (2015). Structural Characteristics and External Correlates of the Working Alliance Inventory-Short Form. *Psychological Assessment, 27*(2), 545–551. https://doi.org/10.1037/pas0000066

Soeteman, D. I., Verheul, R., Delimon, J., Meerman, A. M. M. A., van den Eijnden, E., Rossum, B. V., Ziegler, U., Thunnissen, M., Busschbach, J. J. V., & Kim, J. J. (2010). Cost-effectiveness of psychotherapy for cluster B personality disorders. *British Journal of Psychiatry, 196*(5), 396–403. https://doi.org/10.1192/bjp.bp.109.070482

Soloff, P. H., & Chiappetta, L. (2019). 10-Year Outcome of Suicidal Behavior in Borderline Personality Disorder. *Journal of Personality Disorders, 33*(1), 82–100. https://doi.org/10.1521/pedi_2018_32_332

Somma, A., Borroni, S., Maffei, C., Besson, E., Garbini, A., Granozio, S., Limuti, B., Perego, G., Pietrobon, A., & Rugi, C. (2017). Inter-rater reliability of the Italian translation of the structured clinical interview for DSM-5 personality disorders (SCID-5-PD): a study on consecutively admitted clinical adult participants. *J Psychopathol, 23*, 105–111.

Spinhoven, P., Giesen-Bloo, J., van Dyck, R., & Arntz, A. (2008). Can assessors and therapists predict the outcome of long-term psychotherapy in borderline personality disorder? *Journal of Clinical Psychology, 64*(6), 667–686. https://doi.org/10.1002/jclp.20466

Stinckens, N., Ulburghs, A., & Claes, L. (2009). De werkalliantie als sleutelelement in het therapiegebeuren. *Meting Met Behulp Van De WAV-12: De Nederlandse Vertaling Van De Working Alliance Inventory. Tijdschr Klin Psychol, 39*, 44–60.

Storebø, O. J., Stoffers-Winterling, J. M., Völlm, B. A., Kongerslev, M. T., Mattivi, J. T., Jørgensen, M. S., Faltinsen, E., Todorovac, A., Sales, C. P., Callesen, H. E., Lieb, K., & Simonsen, E. (2020). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews,* (5), CD012955. https://doi.org/10.1002/14651858.CD012955.pub2

Swenson, C. R., & Choi-Kain, L. W. (2015). Mentalization and dialectical behavior therapy. *American Journal of Psychotherapy, 69*(2), 199–217.

Taves, D. R. (1974). Minimization: a new method of assigning patients to treatment and control groups. *Clinical Pharmacology & Therapeutics, 15*(5), 443–453. https://doi.org/10.1002/cpt1974155443

Taylor, P. J., Rietzschel, J., Danquah, A., & Berry, K. (2015). The role of attachment style, attachment to therapist, and working alliance in response to psychological therapy. *Psychology and Psychotherapy-Theory Research and Practice, 88*(3), 240–253. https://doi.org/10.1111/papt.12045

Tellegen, A., & Ben-Porath, Y. S. (2008). *MMPI-2-RF. Technical Manual*. University of Minnesota Press.

ten Have, M., Verheul, R., Kaasenbrood, A., van Dorsselaer, S., Tuithof, M., Kleinjan, M., & de Graaf, R. (2016). Prevalence rates of borderline personality disorder symptoms: a study based on the Netherlands Mental Health Survey and Incidence Study-2. *Bmc Psychiatry, 16*, 249. https://doi.org/10.1186/s12888-016-0939-x

Thorndike, F. P., Ritterband, L. M., Saylor, D. K., Magee, J. C., Gonder-Frederick, L. A., & Morin, C. M. (2011). Validation of the Insomnia Severity Index as a Web-Based Measure. *Behavioral Sleep Medicine, 9*(4), 216–223. https://doi.org/10.1080/15402002.2011.606766

Tomko, R. L., Trull, T. J., Wood, P. K., & Sher, K. J. (2014). Characteristics of Borderline Personality Disorder in a Community Sample: Comorbidity, Treatment Utilization, and General Functioning. *Journal of Personality Disorders, 28*(5), 734–750. https://doi.org/10.1521/pedi_2012_26_093

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the working alliance inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(3), 207. https://doi.org/10.1037/1040-3590.1.3.207

Trull, T. J., Freeman, L. K., Vebares, T. J., Choate, A. M., Helle, A. C., & Wycoff, A. M. (2018). Borderline personality disorder and substance use disorders: an updated review. *Borderline Personality Disorder and Emotion Dysregulation, 5*(1), 1–12. https://doi.org/10.1186/s40479-018-0093-9

Tryon, W. W. (2018). Mediators and Mechanisms. *Clinical Psychological Science, 6*(5), 619–628. https://doi.org/10.1177/2167702618765791

Üstün, T. B., Kostanjsek, N., Chatterji, S., & Rehm, J. (2010). *Measuring health and disability: Manual for WHO disability assessment schedule WHODAS 2.0*. World Health Organization.

van Asselt, A. D. I., Dirksen, C. D., Arntz, A., & Severens, J. L. (2007). The cost of borderline personality disorder: Societal cost of illness in BPD-patients. *European Psychiatry, 22*(6), 354–361. https://doi.org/10.1016/j.eurpsy.2007.04.001

van Asselt, A., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., & Severens, J. L. (2009). The EQ-5D: A useful quality of life measure in borderline personality disorder? *European Psychiatry, 24*(2), 79–85. https://doi.org/10.1016/j.eurpsy.2008.11.001

van Asselt, A. D. I., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., van Dyck, R., Spinhoven, P., van Tilburg, W., Kremers, I. P., Nadort, M., & 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. https://doi.org/10.1192/bjp.bp.106.033597

van Bronswijk, S. C., DeRubeis, R. J., Lemmens, L. H., Peeters, F. P., Keefe, J. R., Cohen, Z. D., & Huibers, M. J. (2021). Precision medicine for long-term depression outcomes using the Personalized Advantage Index approach: cognitive therapy or interpersonal psychotherapy? *Psychological Medicine, 51*(2), 279–289. https://doi.org/10.1017/S0033291719003192

van den Bosch, L. M. C., & Sinnaeve, R. (2015). Dialectische gedragstherapie in Nederland: implementatie en consolidatie. *Tijdschrift Voor Psychiatrie, 57*(10), 719–727.

van den Bosch, L. M. C., Verheul, R., Schippers, G. M., & van den Brink, W. (2002). Dialectical Behavior Therapy of borderline patients with and without substance use problems - Implementation and long-term effects. *Addictive Behaviors, 27*(6), 911–923. https://doi.org/10.1016/S0306-4603(02)00293-9

van den Bosch, L. M., Koeter, M. W., Stijnen, T., Verheul, R., & van den Brink, W. (2005). Sustained efficacy of dialectical behaviour therapy for borderline personality disorder. *Behaviour Research and Therapy, 43*(9), 1231–1241. https://doi.org/10.1016/j.brat.2004.09.008

van Dulmen, S. A., Maas, M., Staal, J. B., Rutten, G., Kiers, H., Nijhuis-van der Sanden, M., & van der Wees, P. (2014). Effectiveness of Peer Assessment for Implementing a Dutch Physical Therapy Low Back Pain Guideline: Cluster Randomized Controlled Trial. *Physical Therapy, 94*(10), 1396–1409. https://doi.org/10.2522/ptj.20130286

van Krugten, F., van Busschbach, J. J., Versteegh, M. M., Hakkaart-van Roijen, L., & Brouwer, W. (2019). *The Mental Health Quality of Life Seven-Dimensional Questionnaire (MHQoL-7D): Development and First Psychometric Evaluation of a New Measure to Assess Quality of Life in People with Mental Health Problems.* Paper presented at the 2019 World Congress on Health Economics.

van Leeuwen, H., Sinnaeve, R., Witteveen, U., Van Daele, T., Ossewaarde, L., Egger, J. I., & van den Bosch, L. (2021). Reviewing the availability, efficacy and clinical utility of Telepsychology in dialectical behavior therapy (Tele-DBT). *Borderline Personality Disorder and Emotion Dysregulation, 8*(1), 1–15. https://doi.org/10.1186/s40479-021-00165-7

van Leeuwen, J., van den Bosch, L., Ossewaarde, L., Sinnaeve, R., & Egger, J. (2017). *Introducing a pre-post study: Efficacy of a blended Dialectical Behavior Therapy consolidation group following one year of standard.* Paper presented at the Strategic Planning Meeting for Dialectical Behavior.

Ventura Wurman, T., Lee, T., Bateman, A., Fonagy, P., & Nolte, T. (2020). Clinical management of common presentations of patients diagnosed with BPD during the COVID-19 pandemic: the contribution of the MBT framework. *Counselling Psychology Quarterly,* 1–27. https://doi.org/10.1080/09515070.2020.1814694

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J., Bateman, A. W., & Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): Development, factor structure, reliability, and validity. *Psychological Assessment, 20*(1), 23. https://doi.org/10.1037/1040-3590.20.1.23

Verheul, R., Van Den Bosch, L. M., Koeter, M. W., De Ridder, M. A., Stijnen, T., & Van Den Brink, W. (2003). Dialectical behaviour therapy for women with borderline personality disorder: 12-month, randomised clinical trial in The Netherlands. *The British Journal of Psychiatry, 182*(2), 135–140.

Versteegh, M. M., Vermeulen, K. M., Evers, S. M. A. A., de Wit, G. A., Prenger, R., & Stolk, E. A. (2016). Dutch Tariff for the Five-Level Version of EQ-5D. *Value in Health, 19*(4), 343–352. https://doi.org/10.1016/j.jval.2016.01.003

Victor, S. E., & Klonsky, E. D. (2016). Validation of a brief version of the difficulties in emotion regulation scale (DERS-18) in five samples. *Journal of Psychopathology and Behavioral Assessment, 38*(4), 582–589. https://doi.org/10.1007/s10862-016-9547-9

Videler, A. C., van Alphen, S. P. J., van Royen, R. J. J., van der Feltz-Cornelis, C. M., Rossi, G., & Arntz, A. (2018). Schema therapy for personality disorders in older adults: a multiple-baseline study. *Aging & Mental Health, 22*(6), 738–747. https://doi.org/10.1080/13607863.2017.1318260

Vitaliano, P. P., Russo, J., Carr, J. E., Maiuro, R. D., & Becker, J. (1985). The ways of coping checklist: Revision and psychometric properties. *Multivariate Behavioral Research, 20*(1), 3–26. https://doi.org/10.1207/s15327906mbr2001_1

Vuijk, R., & Arntz, A. (2017). Schema therapy as treatment for adults with autism spectrum disorder and comorbid personality disorder: Protocol of a multiple-baseline case series study testing cognitive-behavioral and experiential interventions. *Contemporary Clinical Trials Communications, 5*, 80–85. https://doi.org/10.1016/j.conctc.2017.01.001

Wagner, A. W., & Linehan, M. M. (2006). Applications of dialectical behavior therapy to posttraumatic stress disorder and related problems. In V. M. Follette, & J. I. Ruzek (Eds.), *Cognitive Behavioral Therapies for Trauma* (2nd ed., pp. 117–145). The Guilford Press.

Wagner, T., Fydrich, T., Stiglmayr, C., Marschall, P., Salize, H., Renneberg, B., Flessa, S., & Roepke, S. (2014). Societal cost-of-illness in patients with borderline personality disorder one year before, during and after dialectical behavior therapy in routine outpatient care. *Behaviour Research and Therapy, 61*, 12–22. https://doi.org/10.1016/j.brat.2014.07.004

Wallace, M. L., Frank, E., & Kraemer, H. C. (2013). A novel approach for developing and interpreting treatment moderator profiles in randomized clinical trials. *JAMA Psychiatry, 70*(11), 1241–1247. https://doi.org/10.1001/jamapsychiatry.2013.1960

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the Integrity of a Psychotherapy Protocol - Assessment of Adherence and Competence. *Journal of Consulting and Clinical Psychology, 61*(4), 620–630. https://doi.org/10.1037/0022-006X.61.4.620

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: Empirically, ‘’all must have prizes’’. *Psychological Bulletin, 122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

Webb, C. A., Cohen, Z. D., Beard, C., Forgeard, M., Peckham, A. D., & Bjorgvinsson, T. (2020). Personalized Prognostic Prediction of Treatment Outcome for Depressed Patients in a Naturalistic Psychiatric Hospital Setting: A Comparison of Machine Learning Approaches. *Journal of Consulting and Clinical Psychology, 88*(1), 25–38. https://doi.org/10.1037/ccp0000451

Weekers, L. C., Hutsebaut, J., & Kamphuis, J. H. (2019). The Level of Personality Functioning Scale-Brief Form 2.0: Update of a brief instrument for assessing level of personality functioning. *Personality and Mental Health, 13*(1), 3–14. https://doi.org/10.1002/pmh.1434

Weinberg, H. (2020). Online Group Psychotherapy: Challenges and Possibilities During COVID-19-A Practice Review. *Group Dynamics-Theory Research and Practice, 24*(3), 201–211. https://doi.org/10.1037/gdn0000140

Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist, 59*(7), 595–613. https://doi.org/10.1037/0003-066X.59.7.595

Wetzelaer, P., Farrell, J., Evers, S. M. A. A., Jacob, G. A., Lee, C. W., Brand, O., van Breukelen, G., Fassbinder, E., Fretwell, H., Harper, R. P., Lavender, A., Lockwood, G., Malogiannis, I. A., Schweiger, U., Startup, H., Stevenson, T., Zarbock, G., & Arntz, A. (2014). Design of an international multicentre RCT on group schema therapy for borderline personality disorder. *Bmc Psychiatry, 14*, 319. https://doi.org/10.1186/s12888-014-0319-3

White, K., Sheehan, P. W., & Ashton, R. (1977). Imagery assessment: A survey of self-report measures. *Journal of Mental Imagery, 1*(1), 145–169.

Wilson, S. T, Fertuck, E. A., Kwitel, A., Stanley, M. C., & Stanley, B. (2006). Impulsivity, suicidality and alcohol use disorders in adolescents and young adults with borderline personality disorder. *International Journal of Adolescent Medicine and Health, 18*(1), 189–196. https://doi.org/10.1515/IJAMH.2006.18.1.189

Woodbridge, J., Townsend, M., Reis, S., Singh, S., & Grenyer, B. F. (2021). Non-response to psychotherapy for borderline personality disorder: A systematic review. *Australian & New Zealand Journal of Psychiatry*, 1–17. https://doi.org/10.1177/00048674211046893

Yakin, D., Grasman, R., & Arntz, A. (2020). Schema modes as a common mechanism of change in personality pathology and functioning: Results from a randomized controlled trial. *Behaviour Research and Therapy, 126*, 103553. https://doi.org/10.1016/j.brat.2020.103553

Young, J. E., Arntz, A., Atkinson, T., Lobbestael, J., Weishaar, M. E., Van Vreeswijk, M. F., & Klokman, J. (2007). *The schema mode inventory*. Schema Therapy Institute.

Young, J. E., Arntz, A., & Giesen-Bloo, J. (2006). *Therapy adherence and competence scale for Schema focused therapy for borderline personality disorder-adapted version.* Maastricht: University of Maastricht.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy*. Guilford.

Young, J. E., & Klosko, J. S. (1993). *Reinventing your life: how to break free from negative life patterns*. Dutton.

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. https://doi.org/10.1111/j.1600-0447.2010.01543.x

Zarbock, G., Farrell, J. M., Schikowski, A., Heimann, A., Shaw, I., Reiss, N., Verhagen, S., Bot, S., Lee, C. W. & Arntz, A. (2014). *Group Schema Therapy Rating Scale–Revised (GSTRS-R).* http://www.schematherapysociety.org/Resources/Documents/GSTRS-R-fin-2014-07-07.pdf.

Zhang, Q., Wang, L., & Bergeman, C. S. (2018). Multilevel Autoregressive Mediation Models: Specification, Estimation, and Applications. *Psychological Methods, 23*(2), 278–297. https://doi.org/10.1037/met0000161

Zilcha-Mano, S. (2019). Major developments in methods addressing for whom psychotherapy may work and why. *Psychotherapy Research, 29*(6), 693–708. https://doi.org/10.1080/10503307.2018.1429691

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment, 52*(1), 30–41. https://doi.org/10.1207/s15327752jpa5201_2

Zwaap, J., Knies, S., Van der Meijden, C., Staal, P., & Van der Heiden, L. (2015). *Kosteneffectiviteit in de praktijk.* Diemen: Zorginstituut Nederland.

SUPPLEMENTARY MATERIALS

Supplement 1

Table S1. *Overview of Syndrome Disorders Assessed with the SCID-5-S*

Module	Content	Time period
Module A: Mood Episodes and Persistent Depressive Disorder	Major Depressive Episode	Current (past month) and past
	Manic Episode	Current (past month) and past
	Hypomanic Episode	Current (past month) and past
	Persistent Depressive Disorder	Current (past 2 years)
Module B: Psychotic and Associated Symptoms	Delusions	Lifetime
	Hallucinations	Lifetime
	Disorganized Speech and Behavior	Lifetime
	Catatonic Behavior	Lifetime
	Negative Symptoms	Lifetime
Module C: Differential Diagnosis of Psychotic Disorders	Schizophrenia	Current (past month) and past
	Schizophreniform Disorder	Current (past month) and past
	Schizoaffective Disorder	Current (past month) and past
	Delusional Disorder	Current (past month) and past
	Brief Psychotic Disorder	Current (past month) and past
	Other Specified Psychotic Disorder	Current (past month) and past (remission)
	Psychotic Disorder Due to AMC	Lifetime
	Substance/Medication-Induced Psychotic Disorder	Lifetime
Module D: Differential Diagnosis of Mood Disorders	Bipolar I Disorder	Current (past month) and past (remission)
	Bipolar II Disorder	Current (past month) and past (remission)
	Other Specified/Unspecified Bipolar Disorder	Current (past month) and past (remission)
	Bipolar Disorder Due to AMC	Lifetime
	Substance/Medication-Induced Bipolar Disorder	Lifetime
	Major Depressive Disorder	Current (past month) and past (remission)
	Other Specified/Unspecified Depressive Disorder	Current (past month) and past (remission)
	Depressive Disorder Due to AMC	Lifetime
	Substance/Medication-Induced Depressive Disorder	Lifetime

Table S1. *Continued*

Module	Content	Time period
Module E: Substance Use Disorders	Alcohol Use Disorder	Current (past 12 months)
	Sedative, Hypnotic, or Anxiolytic Use Disorder	Current (past 12 months)
	Cannabis Use Disorder	Current (past 12 months)
	Stimulant Use Disorder	Current (past 12 months)
	Opioid Use Disorder	Current (past 12 months)
	Phencyclidine Use Disorder	Current (past 12 months)
	Other Hallucinogen Use Disorder	Current (past 12 months)
	Inhalant Use Disorder	Current (past 12 months)
	Other or Unknown Substance Use Disorder	Current (past 12 months)
	Gambling Disorder	Current (past 12 months)
Module F: Anxiety Disorders	Panic Disorder	Current (past month) and past
	Agoraphobia	Current (past 6 months)
	Social Anxiety Disorder	Current (past 6 months)
	Generalized Anxiety Disorder	Current (past 6 months)
	Anxiety Disorder Due to AMC	Lifetime
	Substance/Medication-Induced Anxiety Disorder	Lifetime
	Specific Phobia	Current (past 6 months)
Module G: Obsessive-Compulsive and Related Disorders	Obsessive-Compulsive Disorder	Current (past month)
	OC and Related Disorder Due to AMC	Lifetime
	Substance/Medication-Induced Obsessive-Compulsive and Related Disorder	Lifetime
	Body Dysmorphic Disorder	Current (past month)
Module H: Posttraumatic Stress Disorder	Posttraumatic Stress Disorder	Current (past month) and past
Module I: Externalizing Disorders	Attention-Deficit/Hyperactivity Disorder	Current (past 6 months)
	Intermittent Explosive Disorder	Current (past 12 months)
Module J: Feeding and Eating Disorders	Anorexia Nervosa	Current (past 3 months) and past
	Bulimia Nervosa	Current (past 3 months)
	Binge-Eating Disorder	Current (past 3 months)
	Avoidant/Restrictive Food Intake Disorder	Current (past month)
	Other Specified Feeding or Eating Disorder	Current
Module K: Somatic Symptom and Related Disorders	Somatic Symptom Disorder	Current (past 6 months)
	Illness Anxiety Disorder	Current (past 6 months)
Module L: Sleep Disorders	Insomnia Disorder	Current (past 3 months)
	Hypersomnolence Disorder	Current (past 3 months)
	Nightmare Disorder	Current (past month)
Module M: Screening for other disorders	Separation Anxiety Disorder	n/a
	Hoarding Disorder	n/a
	Trichotillomania	n/a
	Excoriation Disorder	n/a
Module N: Adjustment Disorder	Adjustment Disorder	Current (past 6 months)

Note. n/a = not available

Supplement 2

Table S1. *Candidate Predictors Based on Clinicians' Appraisals*

Domains and (sub)categories
Personality and psychopathology
Comorbidity
Syndrome disorders
Primary syndrome disorder
Substance use disorder
Depressive disorder
Anxiety disorder
Eating disorder
Anorexia nervosa
Bulimia nervosa
Posttraumatic stress disorder
Re-experiencing symptoms
Developmental disorder
Autism
Personality disorders
Cluster A
Paranoid personality disorder
Schizoid personality disorder
Schizotypal personality disorder
Cluster B
Antisocial personality disorder
Narcissistic personality disorder
Cluster C
Avoidant personality disorder
Obsessive-compulsive personality disorder
Dependent personality disorder
Somatic symptom disorder
Physical diseases
Borderline personality organization (low level vs. high level)
Duration of BPD manifestations
Internalizing BPD manifestations
Feelings of emptiness
Separation anxiety
Identity disturbance
(Self-)destructive behavior
Self-harm
Suicidal behavior
Acting out/impulse control problems
Personality traits
Antisocial
Narcissistic
Avoidant
Obsessive-compulsive
Psychotic symptoms
Dissociation
Flexibility
Pretend mode / detached protector mode
Self-criticism
Perfectionism
Worrying / ruminating
Distrust
Emotion regulation problems
Hostility
Crisis susceptibility
Frustration tolerance
Attachment

Table S1. *Continued*

Domains and (sub)categories
Locus of control
Alexithymia
Experiential avoidance / emotional avoidance
Mental imagery capacity
Untrustworthy
Curious / eager to learn
Self-discipline
Level of Personality Functioning
Capacity and skills
Cognitive abilities
Mentalizing capacity
Insight
Link childhood history to present problems (unable / unwilling)
Coping skills
Social skills
Skills to deal with problems and stresses
Context
Psychosocial and environmental problems
Problems related to the social environment
Economic problems
Housing problems
Occupational and/or educational problems
Living in a destructive environment
Child living at home
History
Trauma
Adoption background
Demographics
Age
Treatment
Motivation / willingness to change
Treatment history
Willingness and/or ability to engage in a therapeutic relationship
Commitment to treatment
Confidence in treatment
Preference for treatment
Request for help
Other
Hope
Degree of suffering
Separation problems
Disruption of day-night rhythm
Positive mental health
Self-stigma

Note. BPD = borderline personality disorder.

Table S2. *Candidate Predictors Based on the Literature and Suggestions of a Patient Representative of the Borderline Foundation of the Netherlands*

Domains and (sub)categories
Personality and psychopathology
Bipolar disorder
Schizophrenia
Severity of BPD manifestations
Personality traits
Paranoid
Dependent
Healthy adult mode
Self-criticism
Querulous person
Hypochondriac
Context
Daytime activities
History
Absence of an attachment figure
Divorce of parents
Family with history of mental illness
Severely bullied
Intimate relationship for at least 6 months
Demographics
Ethnicity
Gender
Level of education
Socioeconomic status
Pregnancy
Other
Dishonesty
Medication use
Unwilling to do homework
Pet

Note. BPD = borderline personality disorder.