EUROPEAN JOURNAL OF PSYCHOTRAUMATOLOGY
2025, VOL. 16, NO. 1, 2497161
https://doi.org/10.1080/20008066.2025.2497161

CLINICAL RESEARCH ARTICLE

OPEN ACCESS Check for updates

Predicting the outcome of psychological treatments for borderline personality disorder and posttraumatic stress disorder: a machine learning approach to predict long-term outcome of Narrative Exposure Therapy vs. Dialectical Behavioral Therapy based treatment

Jakob Blaß [a], Benjamin Iffland[b], Philipp Herzog[c,d], Tim Kaiser[e], Thomas Elbert[f] and Carolin Steuwe[a]

[a]Department of Psychiatry and Psychotherapy, Evangelisches Klinikum Bethel, Universitätsklinikum OWL of Bielefeld University, Bielefeld, Germany; [b]Department of Clinical Psychology and Psychotherapy, Bielefeld University, Bielefeld, Germany; [c]Department of Psychology, Harvard University, Cambridge, MA, USA; [d]Department of Psychology, Pain and Psychotherapy Research Lab, University of Kaiserslautern–Landau (RPTU), Landau, Germany; [e]Department of Methods and Evaluation, Freie Universität Berlin, Berlin, Germany; [f]Department of Psychology, University of Konstanz, Konstanz, Germany

ABSTRACT

Background: A comorbidity between Borderline Personality Disorder (BPD) and Posttraumatic Stress Disorder (PTSD) is common, severely disabling, and hard to treat. The choice of an optimal psychotherapy based on patient characteristics remains challenging.
Objective: This study develops models to predict the outcome of two psychotherapies for comorbid BPD and PTSD.
Method: Data from two trials comparing Narrative Exposure Therapy (NET, $N = 40$) with Dialectical Behavior Therapy-based treatment (DBT-bt, $N = 40$) was analysed. A cross-validated genetic algorithm was used to detect baseline predictors of change in PTSD symptoms.
Results: In the NET group higher education, more baseline PTSD symptoms, more traumatic experiences, fewer baseline BPD symptoms, and not taking antipsychotic medication predicted better treatment outcome. This model (RMSE = 8.98) outperformed the prediction of PTSD symptom reduction with baseline PTSD symptoms alone (RMSE = 10.07) or with all available predictor variables (RMSE = 12.97). Only more baseline PTSD symptoms were selected to predict a better treatment outcome after DBT-bt. This model (RMSE = 9.41) outperformed the prediction of change in PTSD symptoms with all available predictor variables (RMSE = 14.43).
Conclusion: Differences in treatment outcome between NET and DBT-bt may be predictable at baseline, to identify which one of both treatments may be most beneficial for individual patients. The small sample size may restrict the generalizability of the results.

Predicción del resultado de los tratamientos psicológicos para el trastorno límite de la personalidad y el trastorno de estrés postraumático: un enfoque de aprendizaje automático para predecir el resultado a largo plazo de la terapia de exposición narrativa frente al tratamiento basado en la terapia dialéctica conductual

Antecedentes: La comorbilidad entre el Trastorno Límite de la Personalidad (TLP) y el Trastorno de Estrés Postraumático (TEPT) es frecuente, gravemente incapacitante y difícil de tratar. La elección de una psicoterapia óptima basada en las características del paciente sigue siendo un desafío.
Objetivo: Este estudio desarrolla modelos para predecir el resultado de diferentes psicoterapias para el TLP y el TEPT.
Método: Se analizaron los datos de dos ensayos que compararon la Terapia de Exposición Narrativa (NET, $N = 40$) con el tratamiento basado en la Terapia Dialéctica Conductual (DBT-bt, $N = 40$). Se utilizó un algoritmo genético validado para detectar los predictores basales de cambio en los síntomas del TEPT.
Resultados: En el grupo NET, un mayor nivel educacional, más síntomas iniciales de TEPT, más experiencias traumáticas, menos síntomas iniciales de TLP y no tomar medicación antipsicótica predijeron un mejor resultado del tratamiento. Este modelo (RMSE = 8.98) superó la predicción de reducción de los síntomas de TEPT con los síntomas basales de TEPT solos (RMSE = 10.07) o con todas las variables predictoras disponibles (RMSE = 12.97). Solo se seleccionaron más síntomas basales de TEPT para predecir un mejor resultado del tratamiento después de la

ARTICLE HISTORY
Received 2 August 2024
Revised 15 April 2025
Accepted 17 April 2025

KEYWORDS
Borderline personality disorder; posttraumatic stress disorder; Dialectical Behavior Therapy-based treatment; Narrative Exposure Therapy; treatment outcome; prediction models; internal validation; personalization

PALABRAS CLAVE
Trastorno límite de la personalidad; trastorno de estrés postraumático; Tratamiento basado en Terapia Dialéctica Conductual; Terapia de Exposición Narrativa; resultado del tratamiento; modelos de predicción; validación interna; personalización

HIGHLIGHTS

- Combined BPD and PTSD is a severely disabling, chronic, and hard to treat condition. We explored variables that might be able to predict differential treatment outcomes in a PTSD-specific and a BPD-specific treatment.
- Higher education, more baseline PTSD symptoms, more traumatic experiences, fewer baseline BPD symptoms, and not taking antipsychotic medication predicted better treatment outcome in the PTSD-specific treatment.
- Only more baseline PTSD

CONTACT Jakob Blaß jakob.blass@hotmail.com, jakob.blass@evkb.de Department of Psychiatry and Psychotherapy, Evangelisches Klinikum Bethel, Universitätsklinikum OWL of Bielefeld University, Remterweg 69/71, 33617 Bielefeld, Germany
Supplemental data for this article can be accessed online at https://doi.org/10.1080/20008066.2025.2497161.

DBT-bt. Este modelo (RMSE = 9.41) superó la predicción del cambio en los síntomas de TEPT con todas las variables predictoras disponibles (RMSE = 14.43).
Conclusión: Las diferencias en los resultados del tratamiento entre NET y DBT-bt pueden ser predecibles al inicio, para identificar cuál de ambos tratamientos puede ser más beneficioso. El pequeño tamaño de la muestra puede restringir la generalización de los resultados.

symptoms predicted better treatment outcome in the BPD-specific treatment.

1.1. Comorbidity of borderline personality disorder and posttraumatic stress disorder

Borderline Personality Disorder (BPD) has a multi-layered relationship with traumatic experiences. While a history of childhood trauma is common for patients with BPD and is likely to play an important aetiological role (Herzog et al., 2022), BPD phenomena such as re-victimization and high-risk behaviours increase the likelihood of traumatic experiences later in life. Accordingly, comorbidity with posttraumatic stress disorder (PTSD) is high: depending on the observed sample, a proportion of approximately 30 - 50% of people with BPD fulfil criteria for PTSD (Harned et al., 2010; Pagura et al., 2010). Both disorders are severely disabling on their own (Alonso et al., 2004; Le et al., 2020), but their combination leads to an increased and particularly complex symptom burden (Harned et al., 2010; Scheiderer et al., 2015) resulting in severe suffering for patients and also high costs for health care systems (von der Warth et al., 2020; Wagner et al., 2022). The presence of a PTSD diagnosis reduces the chances of BPD remission (Zanarini et al., 2006) and the course of BPD and PTSD symptoms predict each other in the long-term (Barnicot & Crawford, 2018; Masland et al., 2019), often leading to a chronic course of the condition. This highlights the importance of long-term outcomes and PTSD symptoms in research and treatment for this patient group (Zeifman et al., 2021).

1.2. Treatment approaches

Phase-based programmes comprised of high intensity psychological treatments have been developed to meet the needs of these complex and severely ill patients (e.g. Bohus et al., 2013; Harned, 2014). These programmes start with a stabilizing phase, e.g. using dialectical behaviour therapy (DBT; Linehan, 1993), to target BPD symptoms, followed by a trauma-focused treatment (TFT), e.g. using trauma-focused cognitive–behavioural therapy (TF-CBT), to reduce PTSD symptoms. In view of the available treatment studies, phase-based treatment appears to be status quo in this patient group (Harned et al., 2014; Bohus et al., 2020), as are mostly recommended and show encouraging results with large effect sizes that outperformed TFT or DBT alone. However, phase-based treatments are resource intensive (Neuner, 2008). Furthermore, Cuijpers et al. (2024) showed that currently only 50% of all patients with PTSD sufficiently respond to their treatment. Herzog and Kaiser (Herzog & Kaiser, 2022; Kaiser & Herzog, 2023) showed that there is considerable variance in treatment response in patients with BPD as well as in patients with PTSD. This opens up the possibility to optimize the treatment of patients with BPD and PTSD. For example, although phase-based treatments may be the most effective option on average, some individual patients may benefit more from other treatment options, or may benefit equally from less resource-intensive treatments (Cohen & DeRubeis, 2018). Therefore, studying how individual patients respond to different treatments could therefore improve the outcomes for patients and further optimize the cost-effectiveness for healthcare systems. Demonstrating this principle in a similar scenario, Delgadillo et al. (2022) showed that stratification to different treatment approaches could be more effective than a stepped care approach alone in patients with depression. Based on this notion, stratifying patients to a TFT or a skills-based treatment such as DBT may be a promising avenue to explore. For this to be possible, clinicians would first need to know whether the treatments ('phases') are viable options for treating the comorbidity, and secondly, they would need to be able to predict which patients would benefit from each treatment (TFT, DBT and a combination of both) and to what extent. Third, a comparison of stratification with a phase-based approach alone would be needed to determine whether stratification would add benefit beyond an already effective treatment. Regarding the first point, two studies (Pabst et al., 2014; Steuwe et al., 2021) compared the effectiveness of Narrative Exposure Therapy (NET; Schauer et al., 2011), which is a TFT, and Dialectical Behavior Therapy-based treatment (DBT-bt; detailed explanations in the method section) alone in reducing PTSD and BPD symptoms. Both treatments were effective in reducing long-term PTSD and BPD symptoms with medium to large effect sizes, but there was no significant difference in average symptom reduction between the treatments. This coincides with other findings that TFTs are effective in reducing BPD symptoms (De Jongh et al., 2020; Kolthof et al., 2022), as are BPD-specific treatments in reducing PTSD symptoms (Masland et al., 2019). In the study by Steuwe et al. (2021), significantly more patients remitted from PTSD after

NET than after DBT-bt, and in all cases, PTSD remission was accompanied by remission of BPD, which is consistent with the findings of Zanarini et al. (2006). Taken together, a combination of TFT and DBT is most effective in treating patients with comorbid BPD and PTSD, while TFT and DBT alone are also effective treatment options, but less so. However, there is considerable variability in response to these treatments between patients (Herzog & Kaiser, 2022; Kaiser & Herzog, 2023), leaving room for optimization of treatment selection. Therefore, the question remains to what extent an individual patient would benefit from each treatment.

1.3. Prediction of treatment outcome and personalization

Relying on clinical intuition for treatment decisions is known to lead to poorer outcomes compared to statistical decision algorithms (Ægisdóttir et al., 2006) and may therefore result in longer and even more expensive treatments. Establishing predictive statistical models to ultimately guide clinical decision-making is the main focus in the field of precision mental health (Cohen & DeRubeis, 2018). Previous research based on conventional statistical hypothesis testing has identified a number of possible predictors of treatment outcome in PTSD, but without consistent results across studies, limiting the usefulness of the predictors for prediction (Barawi et al., 2020; Keyan et al., 2024). Compared to statistical hypothesis testing, machine learning is a data-driven approach offering more flexibility regarding modelling and the underlying data. It is designed to optimize the utility of models to predict in unseen data. Recent studies have demonstrated the potential use of data-driven personalization(Moggia et al., 2023; Zainal et al., 2024). In a prospective trial Delgadillo et al. (2022) compared a stepped care treatment approach (patients sequentially access low-intensity guided self-help followed by high-intensity psychotherapy) with stratification (patients are matched to low – or high-intensity treatment after initial assessment) based on a statistical prognosis in a naturalistic transdiagnostic sample. Their results showed that stratification was more successful and that patients with complex psychopathology benefited from being directly stratified to high-intensity treatments without going through stepped-care (Delgadillo et al., 2022). Applying this to patients with BPD and PTSD, a statistical prognosis of success of the given treatment options may allow for efficient stratification. In recent years, there has been an increasing number of studies that used machine learning (ML) algorithms to predict treatment outcome in BPD (e.g. Herzog et al., 2020; Keefe et al., 2020) and PTSD samples (Bremer-Hoeve et al., 2023; Deisenhofer et al., 2018; Held et al., 2022; Herzog et al., 2021; Hoeboer et al., 2021; Stirman et al., 2021; Tait et al., 2024). Some of the above-mentioned studies used RCT data and thus were able to compare differential predictions of treatment outcome between treatment groups. For example, Hoeboer and colleagues (2021) built two different prognostic models that helped to differentiate between patients with childhood-abuse-related PTSD regarding their benefit from TFT with and without a prior stabilization. Although a comorbidity of BPD and PTSD can be assumed for a substantial proportion in the sample used by Hoeboer et al. (2021), there is no study to date that predicts a comparable treatment outcome in both treatments while taking the comorbidity into account, thus allowing differential predictions for TFT or BPD-specific treatment. Developing clinical prediction models with sufficient accuracy to stratify clients to treatment options is a multi-step process that requires extensive data (Collins et al., 2024). In the absence of externally validated models, data-driven stratification to treatment options for patients with BPD and PTSD is not yet possible. However, the present study aims to take the first steps in the model development process, namely feature selection and internal validation (Collins et al., 2024). An ML approach will be applied to identify predictors of long-term outcome in DBT-bt (Linehan, 1993) and NET (Schauer et al., 2011) for patients with BPD and PTSD using the aggregated data from Steuwe et al. (2021) and Pabst et al. (2014). In addition, internal cross-validation will be used to test the emerging models against comparison models. It is hypothesized that the ML-based models will outperform the comparator models.

2. Method

2.1. Transparency and openness

We report all data exclusions, all manipulations, and all measures in the study. All data, analysis code, and research materials can be shared upon request. This study's design and its analysis were not preregistered.

2.2. Study sites

Datasets from two trials (Pabst et al., 2014; Steuwe et al., 2021) comparing NET and DBT-bt were used. Data were collected from Pabst et al. (2014; $n = 22$) between 2009 and 2011 in residential and outpatient settings at the University Medical Center Schleswig-Holstein in Kiel, Germany, and from Steuwe et al. (2021; $n = 58$) between 2013 and 2020 in a residential setting at the Clinic of Psychiatry and Psychotherapy, Ev. Klinikum Bethel, Universitätsklinikum OWL, Bielefeld University, Germany. Both studies were approved by the respective ethics committee and all

subjects gave informed consent to participate in the study. For details on study sites and ethics approval see the original studies.

2.3. Participants and procedures

Pabst and colleagues (2014) assigned patients with BPD and PTSD to receive NET or expert treatment of BPD (consisting mainly of DBT components). Steuwe et al. (2021) randomized patients with BPD and PTSD to either NET or DBT-bt. For the present study, the intention-to-treat (ITT) samples from both trials were combined in a sample of $N = 80$ patients. Out of these, $n = 40$ patients were assigned to the NET treatment condition and $n = 40$ patients were assigned to the DBT-bt treatment condition. All participants were female. Further demographics can be seen in Table 1.

No patients were removed when data from the two studies were combined. Data were available at pre-treatment, post-treatment, and 12-month follow-up (12-M FU). Up to the 12-M FU, patients did not receive any booster sessions but were free to seek treatment on their own. At 12-M FU patients were asked about critical and traumatic life events and further therapy after the post-treatment assessment. No patients from Pabst et al. (2014) sought therapy until 12-M FU. 18 patients in the NET-group and 13 patients in the DBT-bt-group started an outpatient psychotherapy between the post- and 12-M FU assessment. The inclusion and exclusion criteria in both studies were highly comparable, the main exception being that Steuwe et al. (2021) explicitly included severely ill patients (a score of at least 50 points on the Clinician-Administered PTSD Scale (Schnyder & Moergeli, 2002) based on the DSM-IV and at least one unsuccessful previous outpatient treatment for BPD). In addition, Steuwe et al. (2021) included patients only if they were female and had not received DBT-based or trauma-focused treatment within the previous 12 months. Compared to Pabst et al. (2014), Steuwe et al. (2021) included patients with a lower BMI (cut-off for inclusion 18 and 16.5, respectively). Noteworthy, Steuwe et al. (2021) used randomization, whereas Pabst et al. (2014) approximately matched participants for age and types of traumatic experiences in both conditions. More detailed descriptions of the inclusion and exclusion criteria can be found in the original studies (Pabst et al., 2014; Steuwe et al., 2021).

2.4. Treatments

NET uses an autobiographical, narrative approach to integrate highly emotional trauma memories with sensory, affective, and cognitive features of the experience and their appropriate situational and temporal contextual information. Clients are guided through the process of constructing a chronological and coherent narrative of their most arousing life experiences, focusing on emotions, cognitions and bodily sensation during the traumatic events and when telling their story. NET is concluded by reviewing the narrative and providing brief cognitive interventions as needed (Schauer et al., 2011). DBT is a modular treatment enabling patients to adaptively deal with current stress, emotions and interpersonal difficulties by focusing on acceptance and change/problem-solving. DBT consists of individual psychotherapy, group skills training, telephone coaching, and a therapist consultation team to achieve symptom improvement by focusing on acceptance and problem-solving. DBT

Table 1. Descriptive statistics of raw data of all possible baseline predictors.

Variable Name (Instrument)	Cronbach's alpha	Mean (SD) or % NET	Mean (SD) or % DBT-bt	$t(p)$ / $\chi^2(p)$
PTSD symptoms (PDS)	.79	34.30 (8.80)	29.92 (15.25)	1.572 (.121)
BPD symptoms (BSL-23)	.91	2.35 (0.77)	2.50 (0.68)	−0.870 (.387)
Anxiety symptoms (BSI-18)	.76	0.49 (0.19)	0.51 (0.19)	−0.569 (.571)
Depressive symptoms (BSI-18)	.72	0.60 (0.17)	0.63 (0.17)	−0.798 (.428)
Dissociative Symptoms (DES)	.97	20.30 (17.58)	20.30 (18.20)	0.024 (.981)
Number of diagnoses (SCID/MINI)	-	4.38 (1.14)	4.31 (1.15)	0.297 (.767)
Number of traumatic experiences (PDS)	-	14.18 (4.83)	11.55 (6.57)	**2.035 (.046)**
Self-harming behaviours (BSL-23)	-	2.20 (2.69)	3.38 (3.49)	−1.686 (.096)
Current mental state (BSL-23)	-	31.67 (15.51)	29.46 (14.99)	0.518 (.607)
Onset of PTSD symptoms 6 months after event (PDS)	-	45.00	27.50	0.685 (.408)
Symptom impact on overall satisfaction with life (PDS)	-	87.50	70.00	0.072 (.789)
Higher Education*	-	20.00	22.50	0.000 (1.000)
Age	-	30.45 (8.35)	30.97 (10.35)	−0.246(.806)
Any medication	-	85.00	67.50	1.481 (0.224)
Anti-depressive medication	-	67.50	60.00	0.027 (.869)
Anti-psychotic medication	-	55.00	42.50	0.462 (.497)
Number of individual sessions	-	15.28 (5.62)	10.18 (4.53)	**4.47 (<.001)**

Note: NET = Narrative Exposure Therapy, DBT-bt = Dialectical Behavior Therapy – based treatment, PTSD = Posttraumatic Stress Disorder, BPD = Borderline Personality Disorder, PDS = Posttraumatic Distress Scale (Foa et al., 1997), BSL-23 = Borderline Symptom List 23 (Bohus et al., 2001), BSI-18 = Brief Symptom Inventory 18 (Spitzer et al., 2011), DES = German version of the Dissociative Experiences Scale (Freyberger et al., 1998), SCID = Structural Clinical Interview for DSM IV (Wittchen et al., 1997), MINI = Mini International Neuropsychiatric Interview (Sheehan et al., 1998). *A patient was dichotomized as having received 'higher' education if she had a high-school diploma or higher. Significant p-values are printed in bold.

helps to cope with stress, regulate emotions and to improve interpersonal competence and mindfulness (Linehan, 1993). Steuwe et al. (2021) used a fixed dose and duration for both treatments (10 weeks of treatment; 17 individual sessions (1,330 minutes) in NET; 10 individual sessions (500 min) and 2850 minutes of group sessions in DBT-bt) matching both treatments in terms of dose (for more details see the supplemental material at Steuwe et al., 2021), whereas groups varied in Pabst et al.'s (2014) study (17.2 ± 7.4 individual sessions in NET; 14.4 ± 3.6 individual sessions in DBT-bt). Otherwise, the content of the treatment conditions in both studies was highly comparable. Because the requirements for DBT certification were not consistently met in either trial, the treatment is referred to as DBT-based treatment (DBT-bt). For more detailed description of treatments and dosage see the original studies (Pabst et al., 2014; Steuwe et al., 2021).

3. Measures

3.1. Merging both datasets

Only those measures were used, for which items with identical or at least highly similar wording were available in both data sets. We used items from the Symptom Checklist 90 revised (SCL-90r; Derogatis, 1975) used in Steuwe et al. (2021) and items from the Hopkins Symptom Checklist 25 (HSCL-25; Glaesmer et al., 2014) used in Pabst et al. (2014) to build the depression and anxiety scales of the Brief Symptom Inventory 18 (BSI 18; Spitzer et al., 2011). Because the SCL-90r and the HSCL-25 use different Likert scales, we scaled all items between 0 and 1 before building the scales. A detailed description of the items that were used to construct the variables can be seen in the supplementary material.

3.2. Outcome measure

FU data were used rather than post-treatment data because of their greater relevance for the ongoing clinical process and to patients' long-term daily lives (for further details see Van Bronswijk et al. (2021)). We chose the PDS change score between pre-treatment and 12-M FU as the outcome measure of the prediction models for two reasons. First, PTSD symptoms were the primary outcome in Steuwe et al. (2021) and Pabst et al. (2014). Second, the reduction of PTSD symptoms is relevant in the treatment of comorbid BPD and PTSD, because BPD remission is most often achieved after PTSD remission (Steuwe et al., 2021; Zanarini et al., 2006). However, we also report the feature selection and performance measures of models predicting the change in the BSL score in the supplementary material.

3.3. Potential predictor variables

All available pre-treatment variables were considered to be candidate predictors of outcome. To account for differences in treatment dose or duration between trials and participants, the number of individual sessions as an estimate value of treatment dose was also included in the selection procedure. Table 1 shows descriptive statistics of the raw pre-treatment data, as well as t-tests and Chi-square-test comparing the two treatment arms. There was a significant difference between groups regarding the number of traumatic events reported. However, this difference was no longer present after imputation. Although patients in the DBT-bt group received fewer individual sessions in accordance with the treatment description, the overall treatment dose is thought to be approximately balanced due to the group component of DBT (see supplemental material in Steuwe et al., 2021).

4. Statistical analyses

All statistical analyses were performed with R (version 4.1.3, R Core Team, 2021) and R Studio (RStudio Team, 2022).

4.1. Preparation for predictor selection

To prepare data for predictor selection, missing values in the potential predictor variables as well as the outcome measure were imputed with the *missForest* method (Stekhoven & Bühlmann, 2012). *missForest* builds on the *randomForest* (RF) algorithm and is able to impute both metrical and ordinal data. It has been shown to outperform other imputation methods and impute reliably even when 30% of datapoints are missing (Stekhoven & Bühlmann, 2012). Furthermore, all categorical variables were dichotomized. In both treatment groups, dichotomous variables which represented less than 10% of the sample in one category were excluded from predictor selection. The percentage of missing values in this dataset was 14.16%. For the outcome measure, the percentage of the missing data was higher (22.5%) yet lower than the threshold proposed by the developers of *missForest* (Stekhoven & Bühlmann, 2012). For the imputation, a normalized RMSE of 0.60 for dimensional variables and a proportion of false classification (PFC) of 0.26 for categorical variables were computed as measures of prediction errors.

4.2. Predictor selection with GA

Luedtke et al. (2019) demonstrated that reliable prescriptive models (i.e. including moderator relationships that differentiate predictions between treatments) require sample sizes of at least 300 patients per

treatment arm. Therefore, in the present study, only one prognostic model (containing predictive information for only one treatment) was built in both treatment groups, leaving a sample size of 40 patients per model. Many of the ML studies described above used algorithms based on RF or regularized regression. In smaller samples ($N < 50$), regularized regression is known to perform poorly (Riley et al., 2021), and decision trees in RF algorithms are restricted to very few nodes. Accordingly, Bain and Shi (2023) showed that the Genetic Algorithm (GA) outperforms RF and regularized regression under such conditions. Therefore, we decided to use GA implemented in the R package *glmulti* (Calcagno & de Mazancourt, 2010) for predictor selection. The GA assumes that there are not only one but many possible models to predict an outcome which all have a certain probability to be the best model. The set of possible models is considered the 'population.' Over multiple iterations (or 'generations'), the population of models is optimized by principles of evolution (e.g. fitness and mutation) regarding a specified criterion of model fit. New generations are developed until the average model fit of the population does not improve further. In this study, GA was combined with a cross-validation technique using the R package *caret* (Kuhn, 2008) to enhance stability and generalizability and to overcome limitations of GA rooted in initializing values (Lee et al., 2022). The data set was split multiple times in a 9–1 ratio to create 30 pairs of training- and test-sets. The GA was performed in each training-set. It was set to consider only the main effects and select models based on the sample size corrected Akaike Information Criterion (AICc). For the ten best models chosen in each training-set, their ability to predict in the test-set was evaluated based on the Root Mean Square Error (RMSE). For each training set, the five models with the most accurate out-of-sample-prediction were chosen. A total of 150 models was selected in each treatment group (30 splits with 5 models selected for each split). Finally, variables were selected for the two final models if they were present in at least 60% of all selected models. 60% was set as the a priori criterion based on the recommendations for a similar algorithm (Austin & Tu, 2004). For exploratory purposes, we report the predictive performance of models derived from other thresholds in the supplemental material. Similarly, we report model parameters for models that include session number as a predictor even if this variable should not be selected by the GA.

4.3. Internal validation of the models

The gold standard to validate the predictive performance of a model is to test its prediction on unseen data (Collins et al., 2024). As the present sample was too small to reserve a hold-out sample for this purpose, the analysis was restricted to internal validation. To this end, we compared the GA-based models with two other models. The first comparison was with baseline models, which predicted the change in the PDS score only with the pre-treatment PDS scores. This was done in order to examine whether or not the predictors selected with the GA enhanced predictive accuracy over and above the prediction with baseline symptoms. Second, we compared the GA-based models with full models that used all potential predictor variables available in the dataset in order to test whether or not the predictor selection process was successful in reducing the number of predictors without decreasing predictive performance. For all models (GA-based, baseline, and full models), RMSE, Mean Absolute Error (MAE), and R^2 were averaged over the 10,000 predictions from 1000 repetitions of 10-fold cross-validation. 95% confidence intervals and t-tests using a Bonferroni-corrected significance threshold were applied to test the statistical difference in performance measures across the 10,000 predictions between models. As RMSE and MAE are indicators of predictive inaccuracy, smaller values are preferable for these measures. Conversely, in the context of cross-validation, R^2 describes the proportion of variance that a model explains in the test sets, so that higher values are preferable.

5. Results

5.1. Predictor selection with GA

Table 2 depicts the percentage of the number of times each variable was selected in a model during the cross-validated GA procedure.

Table 2. Details of feature selection with the cross-validated GA.

Baseline Variable	% selected in NET	% selected in DBT-bt
PDS Score	**100.00**	**94.00**
BSL-23 Score	**91.33**	47.33
BSI-18 anxiety score	20.00	50.67
BSI-18 depression score	15.33	56.00
FDS dissociation score	1.33	58.67
Number of diagnoses	0.00	0.00
Number of traumatic experiences	**77.33**	4.00
Any medication	14.00	6.67
Anti-depressive medication	28.67	30.67
Anti-psychotic medication	**62.67**	18.67
Self-harming behaviour	1.33	11.33
Current mental state	5.33	27.33
Onset of PTSD symptoms	4.00	2.67
Overall satisfaction with life	0.67	3.33
Education	**100.00**	14.00
Age	42.67	3.33
Sessions	5.33	28.67

Note: NET = Narrative Exposure Therapy, DBT-bt = Dialectical Behavior Therapy – based treatment, PDS = Posttraumatic Distress Scale (Foa et al., 1997), BSL-23 = Borderline Symptom List 23, BSI-18 = Brief Symptom Inventory 18, FDS = German version of the Dissociative Experiences Scale. Percentages equal to or larger than 60% are printed in bold.

Table 3. Final linear models for the change in PDS score for both treatment conditions.

NET	Estimate	*SE*	*t*-Value	*p*
Higher education	−1.03	0.35	−2.93	.006
PDS score	−0.56	0.16	−3.42	.001
Anti-psychotic medication	0.41	0.28	1.48	.149
BSL-23 score	0.40	0.16	2.47	.019
Number of traumatic experiences	−0.30	0.14	−2.06	.047
	Adj. $R^2 = .29$, F (5, 34) = 4.20, $p = .004$			
DBT-bt				
PDS score	−0.45	0.14	−3.13	.003
	Adj. $R^2 = .18$, F (1, 38) = 9.82, $p = .003$			

Note: NET = Narrative exposure therapy, DBT-bt = dialectical behaviour therapy – based treatment, PDS = Posttraumatic Distress Scale, BSL-23 = Borderline Symptom List – 23 items version.

A higher PDS score and a lower BSL score, being prescribed antipsychotic medicine, having a higher level of education, and a higher number of traumatic experiences at pre-treatment were selected as predictors of a larger reduction in PTSD symptoms after NET. For DBT-bt, only a higher pre-treatment PDS score was selected as a predictor of a larger reduction in PTSD symptoms. Table 3 describes the fitted models with standardized regression weights. 29% and 18% of the variance in the PDS change score was explained by the models in the NET and DBT-bt groups, respectively. All predictors reached a significance level of $p \leq .05$, with the exception of use of anti-psychotic medication in the NET group. A detailed description of the fitted models with the BSL as the outcome, as well as models including the number of individual sessions, can be found in the supplementary material.

5.2. Internal validation of the models

Tables 4 and 5 present a comparison of the GA-based models, the baseline models, and the full models on the RMSE, R^2 and MAE. Since the GA-based and baseline model were identical in the DBT-bt group, this comparison was not applicable. The GA-based models in both groups (NET_{GA}: RMSE = 8.98; $DBT\text{-}bt_{GA}$: RMSE = 9.41) clearly outperformed the other models ($NET_{baseline}$: RMSE = 10.07; NET_{full}: RMSE = 12.97; $DBT\text{-}bt_{full}$: RMSE = 14.43), with significantly lower RMSE and MAE and higher R^2 in both treatment groups. Both GA-based models had RMSE-values smaller than the standard deviation of the PDS-change in the respective groups (NET: $SD = 10.49$; DBT-bt: $SD = 11.01$), indicating that the model predictions are superior to a prediction with the sample mean. This was not true for the models using all potential predictors and the RMSE of the baseline model in the NET group was smaller but very close to the standard deviation of the outcome in that group. Predictive performance of models derived from other selection cut-offs are presented in the supplementary material. Using a cut-off of 70% in the NET group or 50% in the DBT-bt group would not have had an impact on prediction accuracy.

Table 4. Predictive performance of GA-based, baseline, and full models in the NET group.

	GA-based model	Baseline model (Pre-treatment PDS)	Full model (All available predictor variables)
RMSE [95% CI]	8.98 [8.92; 9.04]	10.07 [9.99; 10.14]	12.97 [12.82; 13.00]
t (*p*)	-	−22.38 (>.001)	−40.13 (>.001)
MAE [95% CI]	7.57 [7.52; 7.63]	8.21 [8.15; 8.26]	9.30 [9.24; 9.37]
t (*p*)	-	−15.68 (>.001)	−38.78 (>.001)
R^2 [95% CI]	0.47 [0.46; 0.47]	0.39 [0.39; 0.40]	0.42 [0.42; 0.43]
t (*p*)	-	16.33 (>.001)	10.27 (>.001)

Note: GA = genetic algorithm, PDS = Posttraumatic Diagnostic Scale, RMSE = Root Mean Square Error, MAE = Mean Average Error.

Table 5. Predictive performance of GA-based and full models in the DBT-bt group.

	GA-based model	Full model (All available predictor variables)
RMSE [95% CI]	9.41 [9.34; 9.47]	14.43 [14.34; 15.52]
t (*p*)	-	−89.14 (>.001)
MAE [95% CI]	8.08 [8.03; 8.14]	12.50 [12.41; 12.58]
t (*p*)	-	−85.81 (>.001)
R^2 [95% CI]	0.44 [0.44; 0.45]	0.32 [0.31; 0.32]
t (*p*)	-	28.84 (>.001)

Note: GA = genetic algorithm, PDS = Posttraumatic Diagnostic Scale, RMSE = Root Mean Square Error, MAE = Mean Average Error.

6. Discussion

The aim of the present study was to identify predictors of treatment outcome in complex patients with comorbid BPD and PTSD who underwent high-intensive treatment with NET vs. DBT-bt. Using a cross-validated GA, different sets of predictors were identified for both treatments: the GA-models explained substantial proportions of variance in PTSD symptom change and outperformed comparison models, supporting that the GA was successful in finding the optimal set of predictors.

6.1. Predictor selection

The cross-validated GA found different clinical prognostic models for each treatment, both of which outperformed comparison models. This speaks to the predictive potential of these models. As the GA selected different predictors in both treatment groups, prediction of differential treatment effects as well as the prescriptive potential of some of these variables seem likely. In the NET group, besides higher self-reported symptoms of PTSD, a higher degree of education, lower self-reported BPD symptoms, a higher number of reported traumatic experiences and not using anti-psychotic medication were identified as positive predictors. In the DBT-bt group only higher

self-reported symptoms of PTSD was predictive of a larger decrease of PTSD symptoms. The level of education or other aspects of socioeconomic status have been found to predict outcomes of TFT in previous studies (Bremer-Hoeve et al., 2023; Deisenhofer et al., 2018; Herzog et al., 2021; Rizvi et al., 2009). To date, there is to our knowledge no study that examines the impact of anti-psychotic medication on the outcome of psychotherapy in PTSD or BPD. As anti-psychotics can have an emotionally numbing effect and also reduce neuroplasticity, they may interfere with the exposure rationale of NET and thereby hinder improvement. More self-reported BPD symptoms hindered the reduction of PTSD symptoms in the NET group. While one study found that the presence of a BPD diagnosis reduced the degree of PTSD symptom reduction after TFT (De Jongh et al., 2020), the majority of studies found no such influence of BPD symptomatology (Zeifman et al., 2021). A higher number of different traumatic experiences predicted a larger reduction of PTSD symptoms after NET. In order to interpret this result, the role of possible confounding variables must first be investigated. These might include the age at which the traumatic experience occurred, or the social recognition of the injustice that occurred. For both treatment groups, more self-reported PTSD symptoms at pre-treatment predicted a larger decrease in PTSD symptoms. In the DBT-bt group this was the only relevant predictor. PTSD symptoms at pre-treatment are the most common predictor of treatment success in previous ML-studies (Herzog et al., 2021; Hoeboer et al., 2021; Stirman et al., 2021). However, the direction of the association varies from study to study (positive association between baseline and outcome PTSD severity in Herzog et al., 2021; negative association between baseline and outcome PTSD severity in Hoeboer et al., 2021 and Stirman et al., 2021). No other predictors were selected for the model in the DBT-bt group in our sample. It remains to be seen, whether other potential prognostic variables predict PTSD symptom reduction after DBT-bt in larger samples. Fittingly, the variance of outcomes in the NET group was significantly larger than in the DBT-bt group in the trial by Steuwe et al. (2021). While DBT-bt seems to result in moderate improvement of PTSD symptoms (but rarely leads to remission) for most patients, NET can lead to major improvements (including remissions) in a number of patients (Steuwe et al., 2021; Zeifman et al., 2021). Others, however, experienced only small improvements (or even worsening) during the one-year follow-up period (Steuwe et al., 2021). Our research suggests that in DBT-bt patients baseline characteristics (with the exception of baseline PTSD severity) play a minor role compared to NET, where a number of baseline characteristics differentiate between patients who benefit more or less. If further validated, the predictors found in this study might help to identify patients that would be particularly likely to improve through NET.

6.2. Limitations

Although we combined the data from Pabst et al. (2014) and Steuwe et al. (2021), which are the only trials comparing the treatment of complex and severely ill patients BPD and PTSD with DBT-bt vs. NET (i.e. TFT without stabilizing components), the sample size was small, thus possibly leading to a comparably high imputation error. The imputation error could in turn have had its own undesirable consequences (e.g. the concealment of group differences), which needs to be considered when interpreting the findings. Furthermore, the sample size did not allow us to build a clinical prescriptive model (i.e. including moderators of treatment outcome), nor to validate the prognostic models in a hold-out-sample. Prospective tests of the models are needed. Even though Pabst et al. (2014) approximately matched patients to both conditions, systematic differences in this part of our sample cannot be ruled out. Furthermore, treatment dose varied between the two original studies, between the two treatments and between individuals. We included the number of individual sessions to account for this. Because of the differences between treatments (DBT-bt including group therapy) this can only be seen as an estimate value of treatment dose. Due to the merging of data sets, the number of available predictor variables was limited. We included variables in the final models that were present in at least 60% of all models selected by the GA. The supplementary material shows that using a cut-off of 70% in the NET group or 50% in the DBT-bt group would not have had an impact on predictive accuracy. Yet, future studies should consider different cut-off scores and different algorithms. Another important limitation is the choice of follow-up measurements as the outcome. We chose the follow-up measurement because of its value in assessing the long-term benefits of treatment. However, this is associated with a considerable loss of patients to follow-up and uncertainty regarding the attribution of symptom reduction to the trial or psychotherapy, which a significant number of patients received after the trial. The dropout rate to follow-up was similar to that found in a recent meta-analysis (Weber et al., 2021). It is also part of the clinical reality that severely disabled patients such as those investigated in our study, rely on long-term psychological therapy, which underlines the need to optimize the therapeutic interventions in this patient group. However, the impact of this ongoing therapy on the follow-up measures must be taken into account when interpreting our results, which therefore must be done with caution.

The original studies were conducted using the ICD-10 criteria. In the meantime, the new diagnosis of Complex Posttraumatic Stress Disorder (CPTSD) was introduced into ICD-11. There has been debate about the possible overlap of CPTSD with the comorbidity of BPD and PTSD (De Jongh et al., 2016), especially when comparing ICD-11 CPTSD with DSM-IV/DSM-5 BPD and PTSD. There is a likelihood that a substantial proportion of the subjects in our study would have met the criteria for a CPTSD diagnosis had diagnostic tools for this condition been available at the time. However, because CPTSD and the comorbidity of BPD and PTSD share similar symptoms and clinical challenges, the consideration of trauma-focused versus skills-focused treatment is relevant regardless of the labels used.

6.3. Implications

The predictors detected in the present study are not yet applicable in clinical practice. Findings need to be replicated in bigger samples ($N \geq 300$ per treatment group; e.g. Luedtke et al., 2019) including ethnical and gender-diverse participants. It is likely that the inclusion of more predictors that were not available in the present dataset (e.g. trauma-related variables, other sociodemographic variables, biological or psychophysiological information) could improve the predictions. Furthermore, external validation and prospective tests of the models in unseen data, possibly from naturalistic treatment samples, are needed. Thereafter, the models could be used to inform personalized treatment decisions by recommending the treatment to an individual patient for which the patient has the most beneficial prognosis (Cohen & DeRubeis, 2018). This personalization should then be tested against different control conditions. These would include a randomized treatment selection, sham personalization, patient preference and, most importantly, a phase-based treatment, which is currently the best established treatment option. If such tests support the model's potential to assign patients to the most promising treatment or improve cost-effectiveness, then implementation in clinical practice would be reasonable.

7. Conclusion

Individuals suffering from combined BPD and PTSD are severely ill and disabled for extended periods of time, resulting in severe suffering and high costs to healthcare systems and society (von der Warth et al., 2020; Wagner et al., 2022). As our knowledge about the treatment of these patients is still incomplete, stratifying individuals to their optimal treatments could be particularly promising for this complex group (Delgadillo et al., 2017; Hoeboer et al., 2021; Keefe et al., 2021). The study demonstrated that it is possible to use a robust statistical algorithm to identify predictors of long-term treatment outcome after two different high-intensity psychological treatments. Different models predicted the outcome after NET and DBT-bt, both of which were effective in reducing BPD and PTSD symptoms. This holds promise for predicting differential treatment outcomes and for personalizing treatment decisions in the future. Some of these predictors have been linked to treatment success in past studies and could be disseminated into clinical practice after further validation. Yet, the results need to be interpreted with caution; however, the study provides initial information for a road to personalized treatment of patients with BPD and PTSD.

Acknowledgements

The authors would like to thank Zachary Cohen for his input and guidance in the preparation of this article.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

We acknowledge support for the publication costs by the Open Access Publication Fund of Bielefeld University and the Deutsche Forschungsgemeinschaft (DFG). Beyond this, our work was not supported by any funding.

Data availability statement

The data that support the findings of this study can be shared upon request by Carolin Steuwe (carolin.steuwe@evkb.de). Restrictions apply to the availability of these data, which were used under licence for this study.

Author contributions statement

Conceptualization: JB, CS, BI, PH, TK
Data curation: JB, CS, TE
Formal analysis: JB
Funding acquisition: CS, TE
Investigation: CS, TE
Methodology: JB, CS, BI, PH, TK
Project administration: JB, CS, BI, TE
Resources: CS, TE
Software: JB
Supervision: CS, BI
Validation: JB
Visualization: JB
Writing – original draft: JB
Writing – review and editing: JB, BI, PH, TK, TE, CS

ORCID

Jakob Blaß http://orcid.org/0009-0004-4169-8449

References

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-Six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*(3), 341–382. https://doi.org/10.1177/0011000005285875

Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H., De Girolamo, G., De Graaf, R., Demyttenaere, K., Gasquet, D., Haro, J. M., Katz, S. J., Kessler, R. C., Kovess, Y., Lépine, J. P., Ormel, J., Polidori, G., Russo, L. J., Vilagut, G., … Vollebergh, W. A. M. (2004). Disability and quality of life impact of mental disorders in Europe: Results from the European study of the epidemiology of mental disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*, *109*(420), 38–46. https://doi.org/10.1111/j.1600-0047.2004.00327.x

Austin, P. C., & Tu, J. V. (2004). Bootstrap methods for developing predictive models. *American Statistician*, *58*(2), 131–137. https://doi.org/10.1198/0003130043277

Bain, C., & Shi, D. (2023). Investigating variable selection techniques under missing data: A simulation study. In *The Annual Meeting of the Psychometric Society* (pp. 109–119). Springer Nature Switzerland.

Barawi, K. S., Lewis, C., Simon, N., & Bisson, J. I. (2020). A systematic review of factors associated with outcome of psychological treatments for post-traumatic stress disorder. *European Journal of Psychotraumatology*, *11*(1). https://doi.org/10.1080/20008198.2020.1774240

Barnicot, K., & Crawford, M. (2018). Posttraumatic stress disorder in patients with borderline personality disorder: Treatment outcomes and mediators. *Journal of Traumatic Stress*, *31*(6), 899–908. https://doi.org/10.1002/jts.22340

Bohus, M., Dyer, A. S., Priebe, K., Krüger, A., Kleindienst, N., Schmahl, C., Niedtfeld, I., & Steil, R. (2013). Dialectical behaviour therapy for post-traumatic stress disorder after childhood sexual abuse in patients with and without borderline personality disorder: A randomised controlled trial. *Psychotherapy and Psychosomatics*, *82*(4), 221–233. https://doi.org/10.1159/000348451

Bohus, M., Kleindienst, N., Hahn, C., Müller-Engelmann, M., Ludäscher, P., Steil, R., Fydrich, T., Kuehner, C., Resick, P. A., Stiglmayr, C., Schmahl, C., & Priebe, K. (2020). Dialectical behavior therapy for posttraumatic stress disorder (DBT-PTSD) compared with cognitive processing therapy (CPT) in complex presentations of PTSD in women survivors of childhood abuse: A randomized clinical trial. *JAMA Psychiatry*, *77*(12), 1235–1245. https://doi.org/10.1001/jamapsychiatry.2020.2148

Bremer-Hoeve, S., van Vliet, N. I., Van Bronswijk, S. C., Huntjens, R. J. C., De Jongh, A., & van Dijk, M. K. (2023). Predictors of treatment dropout in patients with posttraumatic stress disorder due to childhood abuse 1. *Frontiers in Psychiatry, August*, *14*, 1194669. https://doi.org/10.3389/fpsyt.2023.1194669

Calcagno, V., & de Mazancourt, C. (2010). Glmulti: An R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software*, *34*(12), 1–29. doi:10.18637/jss.v034.i12

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology*, *14*(1), 209–236. https://doi.org/10.1146/annurev-clinpsy-050817-084746

Collins, G. S., Dhiman, P., Ma, J., Schlussel, M. M., Archer, L., Van Calster, B., Harrell, F. E., Martin, G. P., Moons, K. G. M., van Smeden, M., Sperrin, M., Bullock, G. S., & Riley, R. D. (2024). Evaluation of clinical prediction models (part 1): From development to external validation. *British Medical Journal*, *384*, e074819. https://doi.org/10.1136/bmj-2023-074819

De Jongh, A., Resick, P. A., Zoellner, L. A., Van Minnen, A., Lee, C. W., Monson, C. M., Foa, E. B., Wheeler, K., Broeke, E. T., Feeny, N., Rauch, S. A. M., Chard, K. M., Mueser, K. T., Sloan, D. M., Van Der Gaag, M., Rothbaum, B. O., Neuner, F., De Roos, C., Hehenkamp, L. M. J., … Bicanic, I. A. E. (2016). Critical analysis of the current treatment guidelines for complex PTSD in adults. *Depression and Anxiety*, *33*(5), 359–369. https://doi.org/10.1002/da.22469

Cuijpers, P., Miguel, C., Ciharova, M., Harrer, M., Basic, D., Cristea, I. A., de Ponti, N., Driessen, E., Hamblen, J., Larsen, S. E., Matbouriahi, M., Papola, D., Pauley, D., Plessen, C. Y., Pfund, R. A., Setkowski, K., Schnurr, P. P., van Ballegooijen, W., Wang, Y., … Karyotaki, E. (2024). Absolute and relative outcomes of psychotherapies for eight mental disorders: A systematic review and meta-analysis. *World Psychiatry*, *23*(2), 267–275. https://doi.org/10.1002/wps.21203

De Jongh, A., Groenland, G. N., Sanches, S., Bongaerts, H., Voorendonk, E. M., & Van Minnen, A. (2020). The impact of brief intensive trauma-focused treatment for PTSD on symptoms of borderline personality disorder. *European Journal of Psychotraumatology*, *11*(1), https://doi.org/10.1080/20008198.2020.1721142

Deisenhofer, A. K., Delgadillo, J., Rubel, J. A., Böhnke, J. R., Zimmermann, D., Schwartz, B., & Lutz, W. (2018). Individual treatment selection for patients with posttraumatic stress disorder. *Depression and Anxiety*, *35*(6), 541–550. https://doi.org/10.1002/da.22755

Delgadillo, J., Ali, S., Fleck, K., Agnew, C., Southgate, A., Parkhouse, L., Cohen, Z. D., Derubeis, R. J., & Barkham, M. (2022). Stratified care vs stepped care for depression: A cluster randomized clinical trial. *JAMA Psychiatry*, *79*(2), 101–108. https://doi.org/10.1001/jamapsychiatry.2021.3539

Delgadillo, J., Huey, D., Bennett, H., & McMillan, D. (2017). Case complexity as a guide for psychological treatment selection. *Journal of Consulting and Clinical Psychology*, *85*(9), 835–853. https://doi.org/10.1037/ccp0000231

Derogatis, L. R. (1975). *The SCL-90-R*. Clinical Psychometric Research.

Glaesmer, H., Braehler, E., Grande, G., Hinz, A., Petermann, F., & Romppel, M. (2014). The German version of the Hopkins Symptoms Checklist-25 (HSCL-25) – factorial structure, psychometric properties, and population-based norms. *Comprehensive Psychiatry*, *55*(2), 396–403. https://doi.org/10.1016/j.comppsych.2013.08.020

Harned, K., Linehan, K. E., & M, M. (2014). A pilot randomized controlled trial of Dialectical Behavior Therapy with and without the Dialectical Behavior Therapy prolonged exposure protocol for suicidal and self-injuring women with borderline personality disorder and PTSD. *Behaviour Research and Therapy*, *55*(1), 7–17. https://doi.org/10.1016/j.brat.2014.01.008

Harned, M. S. (2014). The combined treatment of PTSD with borderline personality disorder. *Current Treatment Options in Psychiatry*, *1*(4), 335–344. https://doi.org/10.1007/s40501-014-0025-2

Harned, R., Linehan, S., & M, M. (2010). Impact of co-occurring posttraumatic stress disorder on suicidal women with borderline personality disorder. *American Journal of Psychiatry*, *167*(10), 1210–1217. doi:10.1176/appi.ajp.2010.09081213

Held, P., Schubert, R. A., Pridgen, S., Kovacevic, M., Montes, M., Christ, N. M., Banerjee, U., & Smith, D. L. (2022). Who will respond to intensive PTSD treatment? A machine learning approach to predicting response prior to starting treatment. *Journal of Psychiatric Research*, *151*(March), 78–85. https://doi.org/10.1016/j.jpsychires.2022.03.066

Herzog, P., Feldmann, M., Voderholzer, U., Gärtner, T., Armbrust, M., Rauh, E., Doerr, R., Rief, W., & Brakemeier, E. L. (2020). Drawing the borderline: Predicting treatment outcomes in patients with borderline personality disorder. *Behaviour Research and Therapy*, *133*, 103692. https://doi.org/10.1016/j.brat.2020.103692

Herzog, P., & Kaiser, T. (2022). Is it worth it to personalize the treatment of PTSD? – A variance-ratio meta-analysis and estimation of treatment effect heterogeneity in RCTs of PTSD. *Journal of Anxiety Disorders*, *91*, 102611. doi:10.1016/j.janxdis.2022.102611

Herzog, P., Kube, T., & Fassbinder, E. (2022). How childhood maltreatment alters perception and cognition – the predictive processing account of borderline personality disorder. *Psychological Medicine*, *52*(14), 2899–2916. https://doi.org/10.1017/S0033291722002458

Herzog, P., Voderholzer, U., Gärtner, T., Osen, B., Svitak, M., Doerr, R., Rolvering-Dijkstra, M., Feldmann, M., Rief, W., & Brakemeier, E. L. (2021). Predictors of outcome during inpatient psychotherapy for posttraumatic stress disorder: A single-treatment, multi-site, practice-based study. *Psychotherapy Research*, *31*(4), 468–482. https://doi.org/10.1080/10503307.2020.1802081

Hoeboer, C. M., Oprel, D. A. C., De Kleine, R. A., Schwartz, B., Deisenhofer, A. K., Schoorl, M., Van Der Does, W. A. J., van Minnen, A., & Lutz, W. (2021). Personalization of treatment for patients with childhoodabuse-related posttraumatic stress disorder. *Journal of Clinical Medicine*, *10*(19), 1–15. https://doi.org/10.3390/jcm10194522

Kaiser, T., & Herzog, P. (2023). Is personalized treatment selection a promising avenue in BPD research ? A meta-regression estimating treatment effect heterogeneity in RCTs of Is personalized treatment selection a promising avenue in BPD research ? A meta-regression estimating treatment. *Journal of Consulting and Clinical Psychology*, *91*(3), 165–170. doi:10.1037/ccp0000803

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*, *51*(11), 1829–1837. https://doi.org/10.1017/S0033291720000550

Keefe, J. R., Kim, T. T., Derubeis, R. J., Streiner, D. L., Links, P. S., & McMain, S. F. (2021). Treatment selection in borderline personality disorder between dialectical behavior therapy and psychodynamic psychiatric management. *Psychological Medicine*, *51*(11), 1829–1837. https://doi.org/10.1017/S0033291720000550

Keyan, D., Garland, N., Choi-christou, J., Tran, J., Donnell, M. O., & Bryant, R. A. (2024). A systematic review and meta-analysis of predictors of response to trauma-focused psychotherapy for posttraumatic stress disorder. *Psychological Bulletin*, *150*(7), 767–797. https://doi.org/10.1037/bul0000438.supp

Kolthof, K. A., Voorendonk, E. M., Van Minnen, A., & De Jongh, A. (2022). Effects of intensive trauma-focused treatment of individuals with both post-traumatic stress disorder and borderline personality disorder. *European Journal of Psychotraumatology*, *13*(2). https://doi.org/10.1080/20008066.2022.2143076

Kuhn, M. (2008). Builduing predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. doi:10.18637/jss.v028.i05

Le, H., Hashmi, A., Czelusta, K. L., Matin, A., Moukaddam, N., & Shah, A. A. (2020). Is borderline personality disorder a serious mental illness? *Psychiatric Annals*, *50*(1), 8–13. https://doi.org/10.3928/00485713-20191203-02

Lee, S., Yang, M. S., Kang, J., & Shin, S. J. (2022). Ensemble variable selection using genetic algorithm. *Communications for Statistical Applications and Methods*, *29*(6), 629–640. https://doi.org/10.29220/CSAM.2022.29.6.629

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Press.

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

Masland, S. R., Cummings, M. H., Null, K. E., Woynowskie, K. M., & Choi-Kain, L. W. (2019). Changes in post-traumatic stress disorder symptoms during residential treatment for borderline personality disorder: A longitudinal cross-lagged study. *Borderline Personality Disorder and Emotion Dysregulation*, *6*(1), 1–9. https://doi.org/10.1186/s40479-019-0113-4

Moggia, D., Saxon, D., Lutz, W., Hardy, G. E., & Barkham, M. (2023). Applying precision methods to treatment selection for moderate/severe depression in person-centered experiential therapy or cognitive behavioral therapy. *Psychotherapy Research*, *34*(8), 1035–1050. https://doi.org/10.1080/10503307.2023.2269297

Neuner, F. (2008). Stabilisierung vor konfrontation in der traumatherapie - grundregel oder mythos? *Verhaltenstherapie*, *18*(2), 109–118. https://doi.org/10.1159/000134006

Pabst, A., Schauer, M., Bernhardt, K., Ruf-Leuschner, M., Goder, R., Elbert, T., Rosentraeger, R., Robjant, K., Aldenhoff, J., & Seeck-Hirschner, M. (2014). Evaluation of Narrative Exposure Therapy (NET) for borderline personality disorder with comorbid posttraumatic stress disorder. *Clinical Neuropsychiatry*, *11*(3–5), 108–117. http://nbn-resolving.de/urn:nbn:de:bsz:352-0-279326.

Pagura, J., Stein, M. B., Bolton, J. M., Cox, B. J., Grant, B., & Sareen, J. (2010). Comorbidity of borderline personality disorder and posttraumatic stress disorder in the U.S. Population. *Journal of Psychiatric Research*, *44*(16), 1190–1198. https://doi.org/10.1016/j.jpsychires.2010.04.016

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/.

Riley, R. D., Snell, K. I. E., Martin, G. P., Whittle, R., Archer, L., Sperrin, M., & Collins, G. S. (2021). Penalization and shrinkage methods produced unreliable clinical prediction models especially when sample size was small. *Journal of Clinical Epidemiology*, *132*, 88–96. https://doi.org/10.1016/j.jclinepi.2020.12.005

Rizvi, S. L., Vogt, D. S., & Resick, P. A. (2009). Cognitive and affective predictors of treatment outcome in cognitive processing therapy and prolonged exposure for posttraumatic stress disorder. *Behaviour Research and Therapy*, *47*(9), 737–743. https://doi.org/10.1016/j.brat.2009.06.003

Schauer, M., Neuner, F., & Elbert, T. (2011). *Narrative exposure therapy: A short-term treatment for traumatic stress disorders*. Hogrefe Publishing GmbH.

Scheiderer, E. M., Wood, P. K., & Trull, T. J. (2015). The comorbidity of borderline personality disorder and posttraumatic stress disorder: Revisiting the prevalence and associations in a general population sample. *Borderline Personality Disorder and Emotion Dysregulation*, *2*(1), 1–16. https://doi.org/10.1186/s40479-015-0032-y

Schnyder, U., & Moergeli, H. (2002). German version of clinician-administered PTSD scale. *Journal of Traumatic Stress*, *15*(6), 487–492. https://doi.org/10.1023/A:1020922023090

Spitzer, C., Hammer, S., Löwe, B., Grabe, H. J., Barnow, S., Rose, M., Wingenfeld, K., Freyberger, H. J., & Franke, G. H. (2011). Die Kurzform des Brief Symptom Inventory (BSI-18): erste Befunde zu den psychometrischen Kennwerten der deutschen Version. *Fortschritte Der Neurologie Psychiatrie*, *79*(9), 517–523. https://doi.org/10.1055/s-0031-1281602

Stekhoven, D. J., & Bühlmann, P. (2012). Missforest-Nonparametric missing value imputation for mixed-type data. *Bioinformatics (Oxford, England)*, *28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Steuwe, C., Berg, M., Beblo, T., & Driessen, M. (2021). Narrative exposure therapy in patients with posttraumatic stress disorder and borderline personality disorder in a naturalistic residential setting: A randomized controlled trial. *Frontiers in Psychiatry*, *12*(November), 1–11. https://doi.org/10.3389/fpsyt.2021.765348

Stirman, S. W., Cohen, Z. D., Lunney, C. A., DeRubeis, R. J., Wiley, J. F., & Schnurr, P. P. (2021). A personalized index to inform selection of a trauma-focused or non-trauma-focused treatment for PTSD. *Behaviour Research and Therapy*, *142*, 103872.

Tait, J., Kellett, S., Saxon, D., Deisenhofer, A. K., Lutz, W., Barkham, M., & Delgadillo, J. (2024). Individual treatment selection for patients with post-traumatic stress disorder: External validation of a personalised advantage index. *Psychotherapy Research*, 1–14. https://doi.org/10.1080/10503307.2024.2360449

RStudio Team. (2022). *RStudio: Integrated Development Environment for R*. http://www.rstudio.com/.

Van Bronswijk, S. C., DeRubeis, R. J., Lemmens, L. H. J. M., Peeters, F. P. M. L., Keefe, J. R., Cohen, Z. D., & Huibers, M. J. H. (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

von der Warth, R., Dams, J., Grochtdreis, T., & König, H. H. (2020). Economic evaluations and cost analyses in posttraumatic stress disorder: A systematic review. *European Journal of Psychotraumatology*, *11*(1). https://doi.org/10.1080/20008198.2020.1753940

Wagner, T., Assmann, N., Köhne, S., Schaich, A., Alvarez-Fischer, D., Borgwardt, S., Arntz, A., Schweiger, U., & Fassbinder, E. (2022). The societal cost of treatment-seeking patients with borderline personality disorder in Germany. *European Archives of Psychiatry and Clinical Neuroscience*, *272*(4), 741–752. https://doi.org/10.1007/s00406-021-01332-1

Weber, M., Schumacher, S., Hannig, W., Barth, J., Lotzin, A., Schäfer, I., Ehring, T., & Kleim, B. (2021). Long-term outcomes of psychological treatment for posttraumatic stress disorder: A systematic review and meta-analysis – corrigendum (Psychological medicine (2021) 51 9 (1420-1430)). In *Psychological medicine* (Vol. 51, Issue 16, p. 2946). https://doi.org/10.1017/S0033291721003214.

Zainal, N. H., Bossarte, R. M., Gildea, S. M., Hwang, I., Kennedy, C. J., Liu, H., Luedtke, A., Marx, B. P., Petukhova, M. V., Post, E. P., Ross, E. L., Sampson, N. A., Sverdrup, E., Turner, B., Wager, S., & Kessler, R. C. (2024). Developing an individualized treatment rule for veterans with major depressive disorder using electronic health records. *Molecular Psychiatry*, *29*(8), 2335–2345. https://doi.org/10.1038/s41380-024-02500-0

Zanarini, M. C., Frankenburg, F. R., Hennen, J., Reich, D. B., & Silk, K. R. (2006). Prediction of the 10-year course of borderline personality disorder. *American Journal of Psychiatry*, *163*(5), 827–832. https://doi.org/10.1176/ajp.2006.163.5.827

Zeifman, R. J., Landy, M. S. H., Liebman, R. E., Fitzpatrick, S., & Monson, C. M. (2021). Optimizing treatment for comorbid borderline personality disorder and posttraumatic stress disorder: A systematic review of psychotherapeutic approaches and treatment efficacy. *Clinical Psychology Review*, *86*(March 2020), 102030. https://doi.org/10.1016/j.cpr.2021.102030