Using the Minnesota Multiphasic Personality Inventory–2 Restructured Form to Predict Functioning After Treatment for Borderline Personality Disorder

A Machine Learning Approach

Wibbelink, Carlijn J.M.; Sellbom, Martin; Grasman, Raoul P.P.P.; Arntz, Arnoud; Sinnaeve, Roland; Kamphuis, Jan H.

DOI
10.1037/pas0001385

Publication date
2025

Document Version
Final published version

Published in
Psychological Assessment

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Assessment

ISSN: 1040-3590

2025, Vol. 37, Nos. 6–7, 261–272
https://doi.org/10.1037/pas0001385

Using the Minnesota Multiphasic Personality Inventory–2 Restructured Form to Predict Functioning After Treatment for Borderline Personality Disorder: A Machine Learning Approach

Carlijn J. M. Wibbelink[1], Martin Sellbom[2], Raoul P. P. P. Grasman[3], Arnoud Arntz[1, 4], Roland Sinnaeve[5], and Jan H. Kamphuis[1]

[1] Department of Clinical Psychology, University of Amsterdam
[2] Department of Psychology, University of Otago
[3] Department of Psychological Methods, University of Amsterdam
[4] Academic Center for Trauma and Personality, Amsterdam, the Netherlands
[5] Department of Neurosciences, Mind Body Research, Catholic University Leuven

Insight into predictors of functioning after treatment for borderline personality disorder (BPD) is limited, despite growing recognition that more focus on other aspects of recovery, especially psychosocial functioning, is warranted. The present study explored the utility of a widely used omnibus assessment instrument, the Minnesota Multiphasic Personality Inventory–2 Restructured Form (MMPI-2-RF), to predict change in functioning during treatment for BPD. Data were obtained from a randomized clinical trial into the effectiveness of 2-year evidence-based treatment for BPD (dialectical behavior therapy or schema therapy) among 130 participants diagnosed with BPD. Different machine learning algorithms, including elastic net regression (ENR), random forest, gradient boosting machine, and extreme gradient boosting, were implemented using nested cross-validation. The ENR model had an average explained variance of 42%. A combination of baseline functioning and four MMPI-2-RF scales emerged as key predictors of change in functioning. Baseline functioning was the most important predictor, with lower initial functioning levels related to more improvement. Higher scores on ideas of persecution, somatic complaints, family problems, and disconstraint were associated with less improvement in functioning. Given the risk of overfitting and the lack of an independent data set, future research should focus on the replicability and generalizability of the findings, as well as clarifying the underlying mechanisms. Our study serves as a first step in identifying patients at risk of poor functional outcome after treatment for BPD.

Public Significance Statement

Given the increasing emphasis on improving psychosocial functioning alongside symptom recovery in treatment for borderline personality disorder, the present study employed machine learning algorithms to identify predictors of functional outcome after treatment using the Minnesota Multiphasic Personality Inventory–2 Restructured Form. Baseline functioning was the strongest predictor, while ideas of persecution, somatic complaints, family problems, and disconstraint were also associated with change in functioning, tentatively suggesting the importance of the therapeutic relationship.

Keywords: borderline personality disorder, predictors, machine learning, Minnesota Multiphasic Personality Inventory–2 Restructured Form, treatment response

This article was published Online First April 14, 2025.

Julie Suhr served as action editor.

Carlijn J. M. Wibbelink https://orcid.org/0000-0001-6002-6892

The Borderline Optimal Treatment Selection (BOOTS) study was approved by the Medical Ethics Committee of the Academic Medical Center Amsterdam (Registration Number NL66731.018.18). The BOOTS study was registered at the Netherlands Trial Register (Registration Number NL7699).

Martin Sellbom is a paid consultant to the University of Minnesota Press, publisher of the Minnesota Multiphasic Personality Inventory–2 Restructured Form. The BOOTS study (Wibbelink et al., 2022) received funding from Stichting Achmea Gezondheidszorg, CZ Fonds, and Stichting Volksbond Rotterdam awarded to Carlijn J. M. Wibbelink. The authors thank all students, research assistants, patients, and therapists for their involvement in the study.

Carlijn J. M. Wibbelink played a lead role in formal analysis, methodology, and writing–original draft and an equal role in conceptualization. Martin Sellbom played a supporting role in methodology and writing–review and editing and an equal role in conceptualization. Raoul P. P. P. Grasman played a supporting role in formal analysis and writing–review and editing. Arnoud Arntz played a supporting role in writing–review and editing. Roland Sinnaeve played a supporting role in writing–review and editing. Jan H. Kamphuis played a lead role in conceptualization and a supporting role in methodology, supervision, and writing–review and editing.

Correspondence concerning this article should be addressed to Carlijn J. M. Wibbelink, Department of Clinical Psychology, University of Amsterdam, P.O. Box 15933, 1001 NK Amsterdam, the Netherlands. Email: C.J.M.Wibbelink@uva.nl

Borderline personality disorder (BPD) is a major health problem, affecting 1%–3% of the general population and up to 22% of psychiatric inpatients (Ellison et al., 2018; Leichsenring et al., 2024; ten Have et al., 2016). BPD is characterized by a pervasive pattern of instability in interpersonal relationships, affect, and identity, along with marked impulsivity (American Psychiatric Association, 2013). Evidence-based treatments for BPD are available (Storebø et al., 2020), but treatment effectiveness studies are often focused on symptom improvement and diagnostic remission (Ng et al., 2019; Renneberg et al., 2024). However, it has become increasingly recognized that more focus on other aspects of recovery, especially psychosocial functioning (Larivière et al., 2015; Ng et al., 2016, 2019; Renneberg et al., 2024), is warranted. Several considerations undergird this assertion. First, a qualitative study examining the experiences of individuals diagnosed with BPD found that while patients with BPD value symptom improvement, 88% also desired better psychosocial functioning, including better social and vocational outcomes (Ng et al., 2019). Second, in psychotherapy studies including functioning measures, less favorable results were found for functional outcomes compared to clinical outcomes (Arntz et al., 2022; Harned et al., 2018; McMain et al., 2012). Likewise, two meta-analytic reviews demonstrated that although patients with BPD had higher levels of functioning after psychotherapy compared to treatment as usual, the effect size was below clinical relevance (Storebø et al., 2020), and there was considerable variability among patients (Zahediabghari et al., 2020). Finally, long-term follow-up studies showed that a substantial proportion of patients, more so than in other disorders, continued to experience difficulties in functioning (J. G. Gunderson et al., 2011; Zanarini et al., 2018). These persistent functional impairments are not only detrimental for patients with BPD (Larivière et al., 2015; Zeitler et al., 2020) but also pose a high burden on society, with up to 80% of BPD-related societal costs attributed to occupational impairment (van Asselt et al., 2007; Wagner et al., 2014, 2022). To conclude, more focus on functioning in patients with BPD appears justified, especially in regard to how to improve functional outcomes after treatment.

One approach to improve treatment effectiveness involves identifying patient characteristics that predict treatment response (Huibers et al., 2021). While numerous studies examined factors associated with BPD symptom change during psychotherapy or treatment discontinuation (Arntz et al., 2015; Barnicot et al., 2012; de Freixo Ferreira et al., 2023; McMurran et al., 2010), less attention has been given to predictors of functioning (Wilks et al., 2016). Wilks et al. (2016) investigated two potential predictors of functional outcome (skills use and emotion dysregulation) after BPD treatment, with only emotion dysregulation being predictive of less improvement in functioning. In addition, Herzog et al. (2020) found that age, work disability, perceived emotional and physical role limitations, and bodily pain were predictive of change in functional outcome among BPD inpatients. Furthermore, a history of abuse, prior treatment, functioning levels, comorbidity, and poor family relationships were associated with functional impairment at 2-year follow-up among treatment-seeking patients with BPD (J. G. Gunderson et al., 2006). Finally, in the McLean Study of Adult Development (Zanarini et al., 2003), patients with BPD initially hospitalized were followed over the course of 16 (Zanarini et al., 2014) and 20 (Zanarini et al., 2018) years to examine recovery rates and predictors of recovery. Recovery was defined by symptomatic remission, good social functioning, and good vocational functioning. Several competence features (IQ and vocational record in childhood and adulthood) and clinical features (cluster C personality disorder, extraversion, agreeableness, and neuroticism) were related to recovery.

While these studies provide some clues as to the determinants of functional treatment outcomes among patients with BPD, existing research has examined a diverse range of predictors, resulting in a lack of replication. Moreover, some studies have only considered a limited set of potential predictors. To overcome these shortcomings, research on predictors of functioning would benefit from utilizing a widely used omnibus assessment instrument that measures a broad spectrum of patient characteristics. The Minnesota Multiphasic Personality Inventory–2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2011) is one such instrument and aligns well with contemporary psychopathology and personality models, including the Hierarchical Taxonomy of Psychopathology (Kotov et al., 2017) and the *Diagnostic and Statistical Manual of Mental Disorders, fifth edition* Alternative Model for Personality Disorders (American Psychiatric Association, 2013; Sellbom, 2019). Moreover, its utility in predicting treatment response has been demonstrated across a variety of populations (e.g., Anestis et al., 2015; Scholte et al., 2012; Sellbom et al., 2008; Tarescavage et al., 2015; Tylicki et al., 2019). However, to our knowledge, no studies have yet utilized the MMPI-2-RF to predict functioning after treatment in individuals with BPD.

The aim of the present study was therefore to examine the predictive validity of the MMPI-2-RF for functioning following treatment for BPD. Understanding what factors are associated with limited functional improvement during treatment enables the identification of patients who may be at risk of poor functional outcome and can help in altering treatment strategies to improve treatment effectiveness (Barnicot et al., 2012). We employed machine learning (ML) methods to achieve this aim, as ML allows for the investigation of numerous potential predictors (Aafjes-van Doorn et al., 2021), which applies to the MMPI-2-RF. Moreover, according to Gillan and Whelan (2017), ML can be particularly beneficial for prediction research in the field of psychology due to its ability to generate reproducible, robust, and generalizable results. The present study was part of a multicenter randomized clinical trial into the effectiveness of two evidence-based treatments for BPD, dialectical behavior therapy (DBT; Linehan, 1993, 2015) and schema therapy (ST; Arntz & van Genderen, 2009; Young et al., 2003; the Borderline Optimal Treatment Selection [BOOTS] study; Wibbelink et al., 2022). Due to the exploratory nature of the study, no specific hypotheses were formulated concerning which MMPI-2-RF scales would predict change in functioning after an evidence-based treatment for BPD.

Method

Participants

The participants were drawn from the BOOTS study (Wibbelink et al., 2022). Participants were eligible to participate in the BOOTS study if they (a) were between 18 and 65 years old, (b) had a primary diagnosis of BPD, (c) had a severity score above 20 on the BPD Severity Index, Version 5, (d) were proficient in Dutch, and (e) were motivated and available for a 2-year treatment and assessments over a period of 3 years. Patients were excluded if they (a) had an IQ score below 80, (b) did not live within a 45-min travel distance from the mental health care center (except when the patient lived in the same

city), (c) had no permanent home address, and (d) had received ST or DBT in the past year. Moreover, patients diagnosed with antisocial personality disorder in combination with interpersonal violence in the past 2 years were excluded to reduce the risk of violent events among group members. Additionally, if patients met the criteria in the past year for a psychotic disorder or bipolar I disorder with one or more manic episodes, they were referred to specialized treatments. To select the final data set for the present study, two additional exclusion criteria were specified: (a) an incomplete MMPI-2-RF or invalid MMPI-2-RF profile (Cannot Say ≥ 15, Variable Response Inconsistency ≥ 80, True Response Inconsistency ≥ 80, Infrequent Responses ≥ 120, and/or Infrequent Psychopathology Responses ≥ 100; Ben-Porath & Tellegen, 2011) and (b) absence of a reassessment of the functioning measure.

Procedure

The BOOTS study was approved by the Medical Ethics Committee of the Academic Medical Center Amsterdam (Registration Number NL66731.018.18). Participants were recruited from January 2019 to October 2021 at various Dutch mental health care centers, including Antes (Rotterdam), GGZ inGeest (Amsterdam), GGZ Noord-Holland Noord (Heerhugowaard), GGZ Rivierduinen (Leiden), NPI (Amsterdam), Pro Persona (Ede and Tiel), PsyQ (Rotterdam-Kralingen), and PsyQ/i-psy (Amsterdam). Patients diagnosed with BPD or suspected of BPD were invited to the screening process. After obtaining written informed consent, the screening process involved the assessment of personality disorders with the Structural Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, fifth edition* Personality Disorders (First et al., 2015), if not already administered, and syndrome disorders with the Structural Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, fifth edition* Syndrome Disorders (First et al., 2018). In addition, the BPD Severity Index, Version 5, and a screening interview assessing the participant's motivation and availability were administered. Moreover, if there were uncertainties regarding the participant's intelligence level, an intelligence test was administered. Participants eligible for participation were invited for the baseline assessment, including self-report questionnaires and interviews. Thereafter, participants were randomized to a 2-year treatment (DBT or ST). For more information on the treatment protocols, the reader is referred to Wibbelink et al. (2022).

After the treatment phase started, participants were reassessed every 6 months until 1 year posttreatment using self-report questionnaires and interviews. The interviews were conducted by independent research assistants. In addition, a brief online questionnaire was administered every 3 months during the treatment period. In the present study, only data collected during the 6-month reassessments until the end of treatment were included.

Materials

MMPI-2-RF

The MMPI-2-RF (Ben-Porath & Tellegen, 2011; Dutch version: van der Heijden et al., 2013) is a self-report questionnaire including 338 true–false items. The MMPI-2-RF contains nine Validity Scales measuring various forms of response bias (e.g., overreporting, random reporting) and 42 substantive scales that assess a wide range of personality characteristics and psychopathology symptoms. These scales are organized in three hierarchical sets (three Higher Order Scales, nine Restructured Clinical Scales, and 23 Specific Problem Scales), two Interest Scales, and a revised version of the Personality Psychopathology Five Scales. The Technical Manual provides extensive documentation on the test's reliability and validity across a variety of samples (Tellegen & Ben-Porath, 2011). A recent review also documents extensive construct validity evidence from subsequent research studies (Sellbom, 2019). In the present study, only substantive subscales that were maintained in the latest version of the MMPI (MMPI-3; Ben-Porath & Tellegen, 2020) were considered as potential predictors, resulting in 37 predictors (excluded subscales: Gastrointestinal Complaints, Head Pain Complaints, Multiple Specific Fears, Aesthetic-Literary Interests, and Mechanical-Physical Interests). Moreover, the Validity Scales were only utilized to exclude patients with invalid test responses.

World Health Organization Disability Assessment Schedule 2.0

Functioning was measured by the 36-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) interview version (Üstün et al., 2010; Dutch version: van der Hoeken et al., 2000). The WHODAS 2.0 measures experienced difficulties over the last 30 days in six major life domains, including cognition, mobility, self-care, getting along, life activities, and participation. The items are rated on a 5-point Likert scale (0 = *none* to 4 = *extreme or cannot do*). A total score indicating overall functioning was calculated using the simple scoring method (Üstün et al., 2010). The WHODAS 2.0 showed adequate psychometric properties among patients and nonpatients (Üstün et al., 2010).

Statistical Analyses

Preprocessing and Outcome Measure

Univariate outliers (Z-score > |3.29|; Tabachnick & Fidell, 2013) were winsorized, and the variables in the data set were screened for near-zero variance. Given that the used ML methods are robust against correlated predictors (e.g., Altelbany, 2021; Herawati et al., 2024; Mousa et al., 2019), we only checked for near-perfect correlations (r > |.95|), which were not found. Finally, 15 predictors with a nonnormal distribution were transformed using the Box–Cox method (Box & Cox, 1964), implemented in the R package caret (Kuhn, 2008), and all predictors were standardized.

To compute the outcome variable—individual change in functioning during treatment—a multilevel model was fitted. This model included a fixed effect of time and a random intercept and slope, with an unstructured covariance structure for the repeated part. The WHODAS 2.0 demonstrated a quadratic relationship with time, in line with previous research (Wilks et al., 2016). Therefore, time was modeled as the square root of the number of months since the start of the treatment phase, which provided a better fit compared to models with a linear effect of time or linear and quadratic effects of time. The estimated individual slope coefficients were then extracted to represent the change in functioning during treatment for each participant. We adhered to the intention-to-treat principle, including data from all patients that started treatment, except from patients with only a baseline assessment of the WHODAS 2.0.

Model Building and Evaluation

For predicting change in functioning (i.e., estimated individual slope coefficients), we implemented and compared four ML approaches, including (a) elastic net regression (ENR; Friedman et al., 2010), (b) random forest (RF; Garge et al., 2013), (c) gradient boosting machine (GBM; Natekin & Knoll, 2013), and (d) extreme gradient boosting (XGB; Chen & Guestrin, 2016). ENR is a regularized linear regression method that combines the Lasso and Ridge penalizations (Zou & Hastie, 2005). RF, GBM, and XGB are decision tree ensemble methods that can handle nonlinear relations and interactions and therefore allow for complex model structures. In RF, multiple decision trees are built in parallel to make predictions. In GBM, the decision trees are built sequentially by focusing on minimizing the residual errors from previous trees. Finally, XGB is a regularized and optimized version of gradient boosting (Jacobucci et al., 2023). The default values of the tuning parameters were applied, except for the parameters listed in Table 1. In each model, the 37 MMPI-2-RF scales as well as baseline functioning levels of the participants were included.

To minimize overfitting, we employed a nested cross-validation (CV) approach that combines an inner and outer loop (Bischl et al., 2012; Jacobucci et al., 2023; Pargent et al., 2023). In the outer loop, the full data set was divided into five folds (80% of the data in the training set and 20% as test data). The inner loop operates on the training set of the outer loop. Within the inner loop, we trained the four ML approaches with tenfold CV to test the performance of these approaches and select the optimal method including the optimal tuning parameter values. Model assessment was then conducted by testing the selected model on the test set from the outer loop, providing a more accurate estimation of the true root-mean-square error (RMSE) and explained variance (R^2). This process was repeated four additional times. Subsequently, the four ML methods were evaluated on the whole data set using tenfold CV, which was repeated 30 times. Variables were selected from the best performing method (lowest RMSE), and the average variable importance of the MMPI-2-RF predictors across the 30 trials was visualized using the varImp function in the caret package. In this method, the contribution of each predictor to the model is estimated using a model-specific approach (e.g., absolute value of the *t*-statistic for linear models), with values scaled to a maximum of 100 (Kuhn, 2008).

Table 1
Adapted Values of the Tuning Parameters

ML approach	Parameter	Value
ENR	tuneLength	50
RF	mtry	1–18
GBM	interaction.depth	1, 2, 3
	n.trees	50–5,000, with increments of 50
	shrinkage	0.0001, 0.001, 0.01, 0.1
	n.minobsinnode	5
XGB	nrounds	200–600, with increments of 50
	eta	0.01, 0.025, 0.05, 0.1
	max_depth	2, 3, 4, 5

Note. ML = machine learning; ENR = elastic net regression; RF = random forest; GBM = gradient boosting machine; XGB = extreme gradient boosting; tuneLength = tuning grid for alpha and lambda values; mtry = number of randomly selected predictors; interaction.depth = maximum tree depth; n.trees = number of iterations; n.minobsinnode = minimal terminal node size; nrounds = number of boosting iterations; max_depth = maximum tree depth; eta = shrinkage.

Transparency and Openness

The BOOTS study was registered at the Netherlands Trial Register (Registration Number NL7699), part of the Dutch Cochrane Center. The present secondary analysis was not preregistered. Research materials and analysis code are available by emailing the corresponding author. The analyses were carried out with Rstudio (RStudio Team, 2020), using R packages lme4 (Bates et al., 2015; multilevel analysis) and caret (Kuhn, 2008; preprocessing and ML analyses).

Results

Patient Characteristics

Out of the 291 patients that were referred to the BOOTS study, 87 patients were not eligible for participation ($n = 3$ met the exclusion criteria and $n = 84$ did not meet the inclusion criteria). In addition, 74 participants were excluded due to an incomplete MMPI-2-RF ($n = 1$) or an invalid MMPI-2-RF profile ($n = 29$, with Cannot Say ≥ 15: $n = 0$, Variable Response Inconsistency ≥ 80: $n = 7$, True Response Inconsistency ≥ 80: $n = 6$, Infrequent Responses ≥ 120: $n = 12$, and Infrequent Psychopathology Responses ≥ 100: $n = 8$) or because no reassessment of the functioning measure was available ($n = 44$). Therefore, 130 patients participated in the study. Most patients were female (84.6%), had a Dutch ethnicity (79.2%), and were unemployed (71.5%), with a mean International Standard Classification of Education score (International Standard Classification of Education, 2012) of 4.32 ($SD = 1.66$) and a mean age of 32.24 ($SD = 10.08$). Of the 130 patients, 42 (32.3%) dropped out of treatment, and 38 (29.2%) had at least one missing reassessment. The descriptive statistics of the MMPI-2-RF scales and WHODAS 2.0 at baseline are presented in the Appendix.

Predicting Change in Functioning

Table 2 presents the results of the nested CV procedure in which the four ML methods were tested in the training data set and, subsequently, the optimal method was evaluated in the test data set. Overall, ENR was the best performing model in the training phase, based on the RMSE. In the test phase, an average RMSE of .0139 ($SD = 5.13 \times 10^{-4}$) was found with an average R^2 of 41.6%, representing an estimation of the performance of this model selection procedure in an independent data set.

Next, the four ML methods were tested in the whole data set across 30 trials. In line with the previous results, ENR was consistently the best performing method (average RMSE = .0134, average $R^2 = .489$). Baseline functioning was the main predictor of change in functioning during treatment across all 30 trials, with lower initial functioning levels related to more improvement in functioning. Figure 1 shows the average variable importance of the MMPI-2-RF predictors, based on the absolute values of the *t*-statistics and scaled to a maximum value of 100 (baseline functioning). Of the 37 MMPI-2-RF scales, nine emerged as predictors of change in functioning during treatment, with higher values of ideas of persecution (RC6), somatic complaints (RC1), family problems (FML), disconstraint-revised (DISC-r),

Table 2
Predictive Performances of the Four Machine Learning Methods in the Training Phase and Optimal Method in the Test Phase

	Training phase								Test phase	
	RMSE				R^2				RMSE	R^2
Fold	ENR	RF	GBM	XGB	ENR	RF	GBM	XGB	Optimal method	
1	**.0137**	.0161	.0149	.0157	.535	.390	.465	.432	.0135	.189
2	**.0133**	.0151	.0148	.0140	.495	.329	.371	.448	.0148	.355
3	**.0126**	.0143	.0136	.0143	.476	.324	.401	.333	.0143	.580
4	**.0129**	.0143	.0134	.0156	.455	.353	.413	.299	.0143	.583
5	**.0138**	.0154	.0147	.0163	.508	.342	.428	.324	.0126	.371

Note. The RMSE in bold is the RMSE of the best performing method in the training phase. RMSE = root-mean-square error; ENR = elastic net regression; RF = random forest; GBM = gradient boosting machine; XGB = extreme gradient boosting.

aberrant experiences, cognitive complaints, and suicidal/death ideation related to less improvement in functioning, whereas more anger proneness and negative emotionality/neuroticism were related to more improvement in functioning. However, five predictors (anger proneness, aberrant experiences, negative emotionality/neuroticism, cognitive complaints, and suicidal/death ideation) had a relatively low average variable importance value (<1), suggesting that these predictors were not as important compared to the other predictors.

Discussion

In the present study, a ML approach was applied to identify predictors of functional outcome after treatment in individuals with BPD, using the MMPI-2-RF. The MMPI-2-RF is a comprehensive measure of psychopathology and personality characteristics that is widely used. We applied ML techniques to examine 37 MMPI-2-RF scales in addition to baseline functioning among patients with BPD receiving a 2-year evidence-based treatment (DBT or ST). Overall, we found that baseline functioning, combined with several MMPI-2-RF scales, was predictive of change in functioning during treatment, with an average explained variance of 42%.

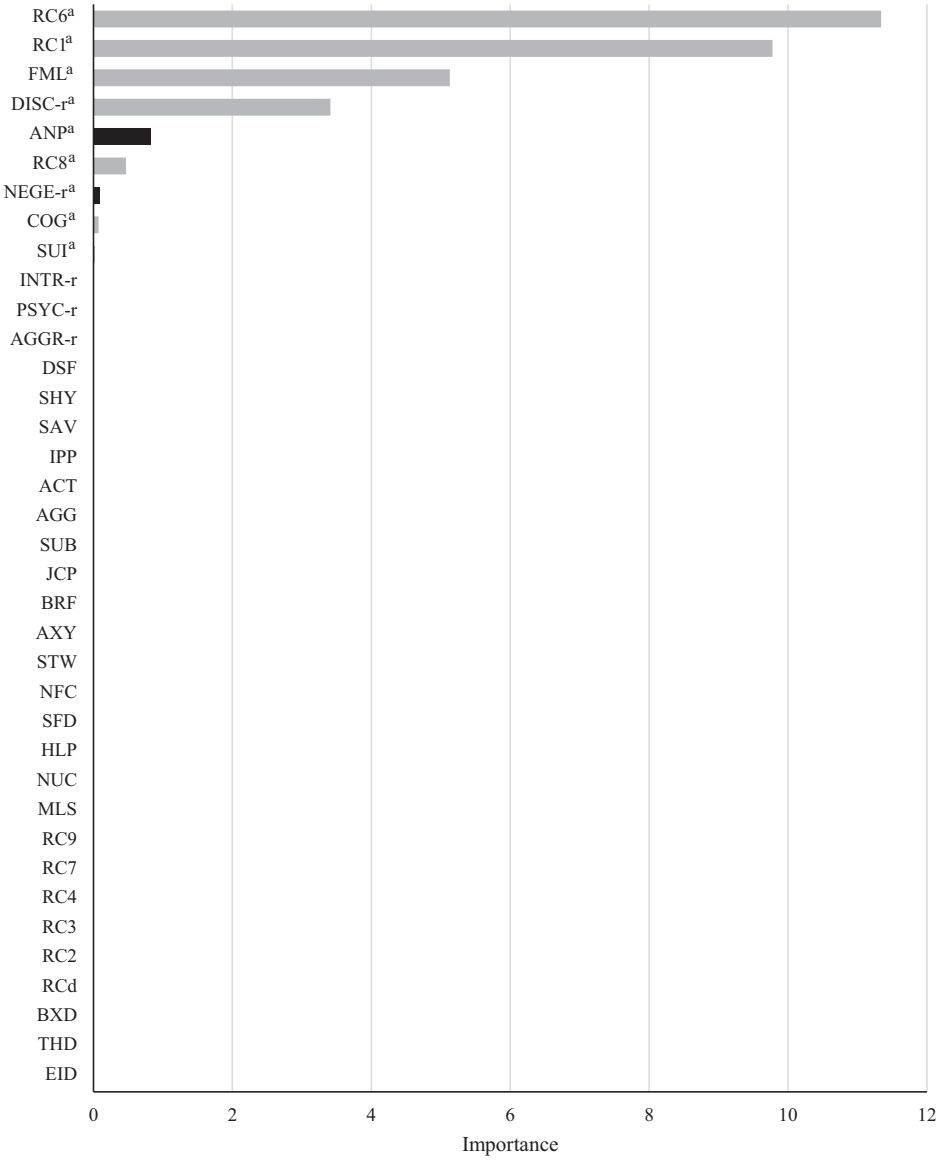
Baseline functioning was by far the strongest predictor of change in functioning, consistent with previous research among diverse patient populations (Bozzatello et al., 2022; Carrión et al., 2013; Cross et al., 2023). Lower initial functioning levels were related to more improvement in functioning, suggesting that the greatest functional gains during treatment may be made by patients with significant levels of impairment at the start of the treatment. Alternatively, this could also reflect statistical artifacts, such as ceiling effects in patients with high initial functional levels and regression to the mean. In the meta-analysis of Barnicot et al. (2012), a similar finding was observed whereby higher initial symptom severity predicted greater symptom improvement among patients with BPD. They concluded that there was minimal evidence of statistical artifacts, suggesting that patients with more symptom severity actually have a greater potential for change. In conclusion, further research is needed on how initial functioning levels affect change in functioning during treatment.

Of the MMPI-2-RF scales, RC6, RC1, FML, and DISC-r were the most important predictors of change in functioning, with higher scores related to less improvement. Although various explanations are possible, a potential underlying mechanism could be the therapeutic relationship, which is considered a crucial element for successful treatment of (BPD) patients (Barnicot et al., 2012; Fahlgren et al., 2020; Flückiger et al., 2018; J. Gunderson et al., 1997; Horvath et al., 2011). According to Ben-Porath and Tellegen (2011), persecutory beliefs (RC6) may impede the development of a therapeutic relationship and affect treatment compliance. Persecutory thinking is related to paranoid delusional thinking, which has been linked to or even equated with epistemic mistrust (Pierre, 2020). Its inverse, epistemic trust, has been defined as the ability to trust others as sources of knowledge about the world and has been posited as a general precondition for any psychotherapeutic intervention to be effective (Fonagy et al., 2017; Kamphuis & Finn, 2019). For social learning (of which psychotherapy is an example) to take place, epistemic hypervigilance needs to be relaxed which likely will be particularly challenging for individuals with higher scores on RC6.

Along with RC6, higher levels of RC1 may also be related to a weakened therapeutic alliance, perhaps through a compromised ability to mentalize psychological problems. In a study among psychiatric outpatients, a negative association between RC1 and therapeutic alliance was found (Patel & Suhr, 2020). Additionally, higher scores on RC1 were related to counterproductive behaviors in treatment, including noncompliance (e.g., arriving late at appointments) and inactivity (not completing homework assignments; Tylicki et al., 2019). Moreover, previous research has demonstrated that somatic problems are associated with poorer functional treatment outcomes, which aligns with the findings of our study. Tylicki et al. (2019) found that more RC1 were linked to reduced functioning at the end of treatment. Similarly, physical role limitations and bodily pain were predictive of less improvement in functioning among BPD inpatients (Herzog et al., 2020). High scores on RC1 indicate a preoccupation with physical health issues, a tendency to develop physical symptoms in response to stress, and a psychological aspect to somatic problems (Ben-Porath & Tellegen, 2011). Patel and Suhr (2020) suggested that a preoccupation with physical health might impede the formation of a therapeutic alliance, as the patient's perception of their difficulties may not align with a psychological formulation of their problems. This misalignment can hinder agreement on the tasks and goals of therapy, which are crucial elements of the therapeutic alliance. Integrating RC1 into therapy goals could therefore enhance the therapeutic alliance (Patel & Suhr, 2020), potentially leading to more treatment success including improved functional outcome.

Figure 1

Average Variable Importance of Minnesota Multiphasic Personality Inventory–2 Restructured Form Predictors Across 30 Iterations



Note. A gray bar indicates a negative relationship, whereas a black bar indicates a positive relationship. ACT = Activation; AGG = Aggression; AGGR-r = Aggressiveness-Revised; ANP = anger proneness; AXY = Anxiety; BRF = Behavior-Restricting Fears; BXD = Behavioral/Externalizing Dysfunction; COG = Cognitive Complaints; DISC-r = disconstraint-revised; DSF = Disaffiliativeness; EID = Emotional/Internalizing Dysfunction; FML = family problems; HLP = Helplessness/Hopelessness; INTR-r = Introversion/Low Positive Emotionality-Revised; IPP = Interpersonal Passivity; JCP = Juvenile Conduct Problems; MLS = Malaise; NEGE-r = negative emotionality/neuroticism-revised; NFC = Inefficacy; NUC = Neurological Complaints; PSYC-r = Psychoticism-Revised; RC1 = somatic complaints; RC2 = Low Positive Emotions; RC3 = Cynicism; RC4 = Antisocial Behavior; RC6 = ideas of persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; RCd = Demoralization; SAV = Social Avoidance; SFD = Self-Doubt; SHY = Shyness; STW = Stress/Worry; SUB = Substance Abuse; SUI = suicidal/death ideation; THD = Thought Dysfunction.

[a] Predictors had an average variable importance larger than 0.

Furthermore, our finding that FML was associated with less improvement in functioning during treatment corresponds to findings from previous studies. J. G. Gunderson et al. (2006) found that poor family relationships predicted reduced functioning after a 2-year follow-up among patients with BPD. Moreover, the FML scale specifically was related to no-show behavior and premature termination in an outpatient sample (Anestis et al., 2015). According to Fruzzetti and colleagues (Fruzzetti et al., 2005; Fruzzetti & Fantozzi, 2008), BPD symptoms and related issues are maintained by distress in close relationships. Close relationships may respond with judgment, express emotions in a dysfunctional manner, and engage in negative behaviors such as criticizing, making demands, or attacking the individual with BPD, which can escalate emotions (Fitzpatrick et al., 2019). Although the empirical evidence is limited, a review by Guillén et al. (2021) suggested that interventions involving family members of patients with BPD can improve family dynamics and, in turn, benefit the patient's recovery. Moreover, the National Institute for Health and Care Excellence guideline on the treatment of BPD recognizes the potential relevance of family interventions for improving the social outcome and welfare of individuals with BPD (National Institute for Health and Clinical Excellence, 2009). To conclude, enhancing family relationships might be a promising approach for improving functional outcomes of patients with BPD, although more research is necessary.

Finally, DISC-r has been repeatedly related to negative treatment outcomes across diverse patient populations (Anestis et al., 2015; Mattson et al., 2012; Patel & Suhr, 2020; Scholte et al., 2012). High scores on DISC-r refer to disinhibited and acting-out behavior, poor impulse control, and sensation seeking (Ben-Porath & Tellegen, 2011). Such externalizing behaviors are associated with more interpersonal problems (Mancke et al., 2015), which, in turn, are linked to a lower quality of the therapeutic relationship (Grosse Holtforth et al., 2014; Strauss et al., 2006). Within a therapeutic context, acting-out behavior toward therapists can lead to conflicts and evoke negative countertransference in therapists (Fahlgren et al., 2020; McMain et al., 2015). Scholte et al. (2012) found that higher DISC-r scores were associated with more disruptive behavior during therapy sessions. Furthermore, disconstrained behavior is assumed to be related to treatment noncompliance and a lack of motivation for treatment (Ben-Porath & Tellegen, 2011). This aligns with findings from several studies linking DISC-r to no-show behavior (Anestis et al., 2015) and dropout (Mattson et al., 2012; Patel & Suhr, 2020).

To summarize, four (RC6, RC1, FML, and DISC-r) of the 37 MMPI-2-RF scales examined emerged as predictors of change in functioning. More insight into the underlying mechanisms of change is however necessary, for example, the role of the therapeutic relationship. In addition, this study has several limitations that need to be considered. First, the primary limitations of our study are the relatively small sample size and the absence of an independent data set. An inadequate sample size increases the risk of overfitting and the probability of producing incorrect results (Rajput et al., 2023). Moreover, the bias–variance trade-off in ML often favors less complex models when data are limited (Pargent et al., 2023). Consequently, our relatively small sample size may be the reason why our comparison of the four ML approaches favored the least complex model (ENR). In addition, the generalizability of our findings might be limited because of our lack of a truly independent data set. To minimize potential bias, we employed a nested CV approach, which, according to several studies (e.g., Vabalas et al., 2019; Varma & Simon, 2006), reduces bias considerably and provides error estimates close to those obtained in independent data sets. Luedtke et al. (2019) also argued that using a smaller sample size is justified if the study aims to develop prediction models that can be tested in future research. We therefore emphasize that the findings of our study, particularly those related to the MMPI-2-RF scales, should serve as input for further cross-validation, with future studies focusing on replicating and generalizing the results.

Second, we utilized the WHODAS 2.0 to assess functioning, which is frequently used in previous research (Federici et al., 2017). Functioning is a multifaceted concept with varying conceptualizations (Habtamu et al., 2016; Ro, 2011). The WHODAS 2.0 follows the definition of functioning according to the International Classification of Functioning, Disability, and Health framework, including functional impairments across six domains (cognition, mobility, self-care, getting along, life activities, and participation; Üstün et al., 2010). However, the WHODAS 2.0 was developed for all mental and physical disorders, resulting in several items that may be less relevant to mental disorders (Konecky et al., 2014). Moreover, vocational problems are not reflected in the functioning score in case individuals are neither working nor studying. This could lead to an underestimation of the disability score for individuals without work or education, which is especially problematic for BPD as individuals with BPD are less frequently employed or in education (Hastrup et al., 2019). It is therefore advisable for future research to also include additional measures of functioning.

Third, our aim was to examine predictors of functional outcome after an evidence-based treatment for BPD. Consequently, this study did not provide insight into moderating factors that predict different treatment responses depending on the treatment (DBT or ST). In future research, it would be interesting to investigate both predictors and moderators of treatment response. Finally, although we adhered to the intention-to-treat principle, patients without a reassessment of our functioning measure could not be included in the analysis, potentially leading to biased results.

Constraints on Generality

Our study sample consisted predominantly of female participants of Dutch ethnicity, recruited from several mental health care institutions. The overrepresentation of female and White participants in our BPD sample aligns with previous studies on the effectiveness of psychotherapy for BPD (Crotty et al., 2024) and is representative of the Dutch BPD outpatient population. However, research on differences between subgroups in treatment effectiveness and predictors of (functional) outcome after treatment among patients with BPD remains inconclusive and scarce (Chang et al., 2023; Cougle & Grubaugh, 2022; Herzog et al., 2020). Herzog et al. (2020) found some indications of differences in predictors of functional outcome after treatment for female and male inpatients with BPD. Unfortunately, our sample size was insufficient to test our models within specific subgroups. It is therefore important for future research to evaluate the generalizability of the results across gender, race, and ethnicity. In addition, patients with a comorbid diagnosis of antisocial personality disorder and a history of interpersonal violence during the last 2 years, as well as patients with a comorbid psychotic disorder or bipolar I disorder with at least one manic episode in the past year, were not eligible to participate in the

BOOTS study. This aligns with clinical practice in the Dutch outpatient mental health care and therefore does not pose a substantial threat to the ecological validity of our study in the Netherlands. However, it may limit the generalizability to other countries and mental health care settings.

Conclusion

Taken together, baseline functioning emerged as a key predictor of functional outcome after an evidence-based treatment for BPD. Additionally, four MMPI-2-RF scales were also predictive of change in functioning. To our knowledge, this is the first study that utilized a ML approach to identify predictors of functional outcome after treatment in individuals with BPD based on a single assessment instrument, the MMPI-2-RF. The widespread use of the MMPI-2-RF in routine clinical settings enhances the feasibility of implementing a prediction model in clinical practice. Future research should focus on the replicability and generalizability of the findings and the underlying mechanisms involved. Our study serves as a first step in identifying patients at risk of poor functional outcome following an evidence-based treatment for BPD. This information could help in refining treatment strategies to enhance treatment efficacy and, consequently, improve the well-being of individuals with BPD.

References

Aafjes-van Doorn, K., Kamsteeg, C., Bate, J., & Aafjes, M. (2021). A scoping review of machine learning in psychotherapy research. *Psychotherapy Research*, *31*(1), 92–116. https://doi.org/10.1080/10503307.2020.1808729

Altelbany, S. (2021). Evaluation of ridge, elastic net and lasso regression methods in precedence of multicollinearity problem: A simulation study. *Journal of Applied Economics and Business Studies*, *5*(1), 131–142. https://doi.org/10.34260/jaebs.517

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

Anestis, J. C., Gottfried, E. D., & Joiner, T. E. (2015). The utility of MMPI-2-RF substantive scales in prediction of negative treatment outcomes in a community mental health center. *Assessment*, *22*(1), 23–35. https://doi.org/10.1177/1073191114536771

Arntz, A., Jacob, G. A., Lee, C. W., Brand-de Wilde, O. M., Fassbinder, E., Harper, R. P., Lavender, A., Lockwood, G., Malogiannis, I. A., Ruths, F. A., Schweiger, U., Shaw, I. A., Zarbock, G., & Farrell, J. M. (2022). Effectiveness of predominantly group schema therapy and combined individual and group schema therapy for borderline personality disorder: A randomized clinical trial. *JAMA Psychiatry*, *79*(4), 287–299. https://doi.org/10.1001/jamapsychiatry.2022.0010

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. (2009). *Schema therapy for borderline personality disorder*. Wiley-Blackwell.

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Ben-Porath, Y. S., & Tellegen, A. (2011). *Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF): Manual for administration, scoring, and interpretation*. University of Minnesota Press. (Original work published 2008).

Ben-Porath, Y. S., & Tellegen, A. (2020). *Minnesota Multiphasic Personality Inventory-3 (MMPI-3): Manual for administration, scoring, and interpretation*. University of Minnesota Press.

Bischl, B., Mersmann, O., Trautmann, H., & Weihs, C. (2012). Resampling methods for meta-model validation with recommendations for evolutionary computation. *Evolutionary Computation*, *20*(2), 249–275. https://doi.org/10.1162/EVCO_a_00069

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Bozzatello, P., Giordano, B., Montemagni, C., Rocca, P., & Bellino, S. (2022). Real-world functioning in psychiatric outpatients: Predictive factors. *Journal of Clinical Medicine*, *11*(15), Article 4400. https://doi.org/10.3390/jcm11154400

Carrión, R. E., McLaughlin, D., Goldberg, T. E., Auther, A. M., Olsen, R. H., Olvet, D. M., Correll, C. U., & Cornblatt, B. A. (2013). Prediction of functional outcome in individuals at clinical high risk for psychosis. *JAMA Psychiatry*, *70*(11), 1133–1142. https://doi.org/10.1001/jamapsychiatry.2013.1909

Chang, C. J., Halvorson, M. A., Lehavot, K., Simpson, T. L., & Harned, M. S. (2023). Sexual identity and race/ethnicity as predictors of treatment outcome and retention in dialectical behavior therapy. *Journal of Consulting and Clinical Psychology*, *91*(10), 614–621. https://doi.org/10.1037/ccp0000826

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785

Cougle, J. R., & Grubaugh, A. L. (2022). Do psychosocial treatment outcomes vary by race or ethnicity? A review of meta-analyses. *Clinical Psychology Review*, *96*, Article 102192. https://doi.org/10.1016/j.cpr.2022.102192

Cross, S. P., Karin, E., Asrianti, L., Walker, J., Staples, L. G., Bisby, M. A., Nielssen, O., Kayrouz, R., Fisher, A., Dear, B. F., & Titov, N. (2023). Predictors of functional impairment at assessment and functional improvement after treatment at a national digital mental health service. *Internet Interventions*, *31*, Article 100603. https://doi.org/10.1016/j.invent.2023.100603

Crotty, K., Viswanathan, M., Kennedy, S., Edlund, M. J., Ali, R., Siddiqui, M., Wines, R., Ratajczak, P., & Gartlehner, G. (2024). Psychotherapies for the treatment of borderline personality disorder: A systematic review. *Journal of Consulting and Clinical Psychology*, *92*(5), 275–295. https://doi.org/10.1037/ccp0000833

de Freixo Ferreira, L., Guerra, C., & Vieira-Coelho, M. A. (2023). Predictors of psychotherapy dropout in patients with borderline personality disorder: A systematic review. *Clinical Psychology & Psychotherapy*, *30*(6), 1324–1337. https://doi.org/10.1002/cpp.2888

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

Fahlgren, M. K., Berman, M. E., & McCloskey, M. S. (2020). The role of therapeutic alliance in therapy for adults with problematic aggression and associated disorders. *Clinical Psychology & Psychotherapy*, *27*(6), 858–886. https://doi.org/10.1002/cpp.2475

Federici, S., Bracalenti, M., Meloni, F., & Luciano, J. V. (2017). World Health Organization disability assessment schedule 2.0: An international systematic review. *Disability and Rehabilitation*, *39*(23), 2347–2380. https://doi.org/10.1080/09638288.2016.1223177

First, M. B., Williams, J. B. W., 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., Karg, R. S., & Spitzer, R. L. (2018). *Gestructureerd klinisch interview voor DSM-5 Syndroomstoornissen (SCID-5-S)*. Nederlandse vertaling van Structural 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) (1st ed.). Boom.

Fitzpatrick, S., Wagner, A. C., & Monson, C. M. (2019). Optimizing borderline personality disorder treatment by incorporating significant others: A review and synthesis. *Personality Disorders: Theory, Research, and Treatment*, *10*(4), 297–308. https://doi.org/10.1037/per0000328

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *55*(4), 316–340. https://doi.org/10.1037/pst0000172

Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2017). What we have changed our minds about: Part 2. Borderline personality disorder, epistemic trust and the developmental significance of social communication. *Borderline Personality Disorder and Emotion Dysregulation*, *4*(1), Article 9. https://doi.org/10.1186/s40479-017-0062-8

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

Fruzzetti, A. E., & Fantozzi, B. (2008). Couple therapy and the treatment of borderline personality and related disorders. In A. S. Gurman (Ed.), *Clinical handbook of couple therapy* (4th ed., pp. 567–590). Guilford Press.

Fruzzetti, A. E., Shenk, C., & Hoffman, P. D. (2005). Family interaction and the development of borderline personality disorder: A transactional model. *Development and Psychopathology*, *17*(4), 1007–1030. https://doi.org/10.1017/S0954579405050479

Garge, N. R., Bobashev, G., & Eggleston, B. (2013). Random forest methodology for model-based recursive partitioning: The mobForest package for R. *BMC Bioinformatics*, *14*(1), Article 125. https://doi.org/10.1186/1471-2105-14-125

Gillan, C. M., & Whelan, R. (2017). What big data can do for treatment in psychiatry. *Current Opinion in Behavioral Sciences*, *18*, 34–42. https://doi.org/10.1016/j.cobeha.2017.07.003

Grosse Holtforth, M., Altenstein, D., Krieger, T., Flückiger, C., Wright, A. G., & Caspar, F. (2014). Interpersonal differentiation within depression diagnosis: Relating interpersonal subgroups to symptom load and the quality of the early therapeutic alliance. *Psychotherapy Research*, *24*(4), 429–441. https://doi.org/10.1080/10503307.2013.829253

Guillén, V., Díaz-García, A., Mira, A., García-Palacios, A., Escrivá-Martínez, T., Baños, R., & Botella, C. (2021). Interventions for family members and carers of patients with borderline personality disorder: A systematic review. *Family Process*, *60*(1), 134–144. https://doi.org/10.1111/famp.12537

Gunderson, J., Najavits, L., Leonhard, C., Sullivan, C., & Sabo, A. (1997). Ontogeny of the therapeutic alliance in borderline patients. *Psychotherapy Research*, *7*(3), 301–309. https://doi.org/10.1080/10503309712331332033

Gunderson, J. G., Daversa, M. T., Grilo, C. M., McGlashan, T. H., Zanarini, M. C., Shea, M. T., Skodol, A. E., Yen, S., Sanislow, C. A., Bender, D. S., Dyck, I. R., Morey, L. C., & Stout, R. L. (2006). Predictors of 2-year outcome for patients with borderline personality disorder. *The American Journal of Psychiatry*, *163*(5), 822–826. https://doi.org/10.1176/ajp.2006.163.5.822

Gunderson, J. G., Stout, R. L., McGlashan, T. H., Shea, M. 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

Habtamu, K., Alem, A., Medhin, G., Fekadu, A., Prince, M., & Hanlon, C. (2016). Development and validation of a contextual measure of functioning for people living with severe mental disorders in rural Africa. *BMC Psychiatry*, *16*(1), Article 311. https://doi.org/10.1186/s12888-016-1022-3

Harned, M. S., Wilks, C. R., Schmidt, S. C., & Coyle, T. N. (2018). Improving functional outcomes in women with borderline personality disorder and PTSD by changing PTSD severity and post-traumatic cognitions. *Behaviour Research and Therapy*, *103*, 53–61. https://doi.org/10.1016/j.brat.2018.02.002

Hastrup, L. H., Kongerslev, M. T., & Simonsen, E. (2019). Low vocational outcome among people diagnosed with borderline personality disorder during first admission to mental health services in Denmark: A nationwide 9-year register-based study. *Journal of Personality Disorders*, *33*(3), 326–340. https://doi.org/10.1521/pedi_2018_32_344

Herawati, N., Sutrisno, A., Nusyirwan, N., & Misgiyati, M. (2024). The performance of ridge regression, LASSO, and elastic-net in controlling multicollinearity: A simulation and application. *Journal of Modern Applied Statistical Methods*, *23*(2). https://doi.org/10.56801/Jmasm.V23.i2.4

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*, Article 103692. https://doi.org/10.1016/j.brat.2020.103692

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, *48*(1), 9–16. https://doi.org/10.1037/a0022186

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*, Article 607508. https://doi.org/10.3389/fpsyt.2020.607508

International Standard Classification of Education. (2012). *International Standard Classification of Education (ISCED) 2011*. UNESCO Institute for Statistics.

Jacobucci, R., Grimm, K. J., & Zhang, Z. (2023). *Machine learning for social and behavioral research*. Guilford Press.

Kamphuis, J. H., & Finn, S. E. (2019). Therapeutic assessment in personality disorders: Toward the restoration of epistemic trust. *Journal of Personality Assessment*, *101*(6), 662–674. https://doi.org/10.1080/00223891.2018.1476360

Konecky, B., Meyer, E. C., Marx, B. P., Kimbrel, N. A., & Morissette, S. B. (2014). Using the WHODAS 2.0 to assess functional disability associated with DSM-5 mental disorders. *The American Journal of Psychiatry*, *171*(8), 818–820. https://doi.org/10.1176/appi.ajp.2014.14050587

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., … Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

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

Larivière, N., Couture, É., Blackburn, C., Carbonneau, M., Lacombe, C., Schinck, S. A., 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

Leichsenring, F., Fonagy, P., Heim, N., Kernberg, O. F., Leweke, F., Luyten, P., Salzer, S., Spitzer, C., & Steinert, C. (2024). Borderline personality disorder: A comprehensive review of diagnosis and clinical presentation,

etiology, treatment, and current controversies. *World Psychiatry*, *23*(1), 4–25. https://doi.org/10.1002/wps.21156

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

Linehan, M. M. (2015). *DBT skills training manual* (2nd ed.). 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

Mancke, F., Herpertz, S. C., & Bertsch, K. (2015). Aggression in borderline personality disorder: A multidimensional model. *Personality Disorders: Theory, Research, and Treatment*, *6*(3), 278–291. https://doi.org/10.1037/per0000098

Mattson, C., Powers, B., Halfaker, D., Akeson, S., & Ben-Porath, Y. (2012). Predicting drug court treatment completion using the MMPI-2-RF. *Psychological Assessment*, *24*(4), 937–943. https://doi.org/10.1037/a0028267

McMain, S. F., Boritz, T. Z., & Leybman, M. J. (2015). Common strategies for cultivating a positive therapy relationship in the treatment of borderline personality disorder. *Journal of Psychotherapy Integration*, *25*(1), 20–29. https://doi.org/10.1037/a0038768

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. *The American Journal of Psychiatry*, *169*(6), 650–661. https://doi.org/10.1176/appi.ajp.2012.11091416

McMurran, M., Huband, N., & Overton, E. (2010). Non-completion of personality disorder treatments: A systematic review of correlates, consequences, and interventions. *Clinical Psychology Review*, *30*(3), 277–287. https://doi.org/10.1016/j.cpr.2009.12.002

Mousa, S. R., Bakhit, P. R., & Ishak, S. (2019). An extreme gradient boosting method for identifying the factors contributing to crash/near-crash events: A naturalistic driving study. *Canadian Journal of Civil Engineering*, *46*(8), 712–721. https://doi.org/10.1139/cjce-2018-0117

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*, Article 21. https://doi.org/10.3389/fnbot.2013.00021

National Institute for Health and Clinical Excellence. (2009). *Borderline personality disorder: Treatment and management* (Clinical guideline No. 78). https://www.nice.org.uk/guidance/cg78

Ng, F. Y. Y., Bourke, M. E., & Grenyer, B. F. (2016). Recovery from borderline personality disorder: A systematic review of the perspectives of consumers, clinicians, family and carers. *PLOS ONE*, *11*(8), Article e0160515. https://doi.org/10.1371/journal.pone.0160515

Ng, F. Y. Y., Carter, P. E., Bourke, M. E., & Grenyer, B. F. S. (2019). What do individuals with borderline personality disorder want from treatment? A study of self-generated treatment and recovery goals. *Journal of Psychiatric Practice*, *25*(2), 148–155. https://doi.org/10.1097/PRA.0000000000000369

Pargent, F., Schedel, R., & Stachl, C. (2023). Best practices in supervised machine learning: A tutorial for psychologists. *Advances in Methods and Practices in Psychological Science*, *6*(3), 1–35. https://doi.org/10.1177/25152459231162559

Patel, K. D., & Suhr, J. A. (2020). The relationship of MMPI-2-RF scales to treatment engagement and alliance. *Journal of Personality Assessment*, *102*(5), 594–603. https://doi.org/10.1080/00223891.2019.1635488

Pierre, J. M. (2020). Mistrust and misinformation: A two-component, socio-epistemic model of belief in conspiracy theories. *Journal of Social and Political Psychology*, *8*(2), 617–641. https://doi.org/10.5964/jspp.v8i2.1362

Rajput, D., Wang, W. J., & Chen, C. C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*, *24*(1), Article 48. https://doi.org/10.1186/s12859-023-05156-9

Renneberg, B., Hutsebaut, J., Berens, A., De Panfilis, C., Bertsch, K., Kaera, A., Kramer, U., Schmahl, C., Swales, M., Taubner, S., Alvarez, M. M., Sieg, J., & the 30 Experts With Lived Experience—Clients, Relatives, Significant Others—From 10 European Countries. (2024). Towards an informed research agenda for the field of personality disorders by experts with lived and living experience and researchers. *Borderline Personality Disorder and Emotion Dysregulation*, *11*(1), Article 14. https://doi.org/10.1186/s40479-024-00257-0

Ro, E. (2011). *Conceptualization of psychosocial functioning: Understanding structure and relations with personality and psychopathology* (Publication No. 3476349) [Doctoral dissertation, University of Minnesota]. ProQuest Dissertations and Theses Global.

RStudio Team. (2020). *RStudio: Integrated development environment for R* [Computer software]. RStudio, PBC. https://www.rstudio.com

Scholte, W., Tiemens, B., Verheul, R., Meerman, A., Egger, J., & Hutschemaekers, G. (2012). The RC scales predict psychotherapy outcomes: The predictive validity of the MMPI-2's restructured clinical scales for psychotherapeutic outcomes. *Personality and Mental Health*, *6*(4), 292–302. https://doi.org/10.1002/pmh.1190

Sellbom, M. (2019). The MMPI-2-Restructured Form (MMPI-2-RF): Assessment of personality and psychopathology in the twenty-first century. *Annual Review of Clinical Psychology*, *15*(1), 149–177. https://doi.org/10.1146/annurev-clinpsy-050718-095701

Sellbom, M., Ben-Porath, Y. S., Baum, L. J., Erez, E., & Gregory, C. (2008). Predictive validity of the MMPI-2 Restructured Clinical (RC) Scales in a batterers' intervention program. *Journal of Personality Assessment*, *90*(2), 129–135. https://doi.org/10.1080/00223890701845153

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*, *2020*(5), Article CD012955. https://doi.org/10.1002/14651858.CD012955.pub2

Strauss, J. L., Hayes, A. M., Johnson, S. L., Newman, C. F., Brown, G. K., Barber, J. P., Laurenceau, J. P., & Beck, A. T. (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial of cognitive therapy for avoidant and obsessive–compulsive personality disorders. *Journal of Consulting and Clinical Psychology*, *74*(2), 337–345. https://doi.org/10.1037/0022-006X.74.2.337

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Tarescavage, A. M., Corey, D. M., & Ben-Porath, Y. S. (2015). Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) predictors of police officer problem behavior. *Assessment*, *22*(1), 116–132. https://doi.org/10.1177/1073191114534885

Tellegen, A., & Ben-Porath, Y. S. (2011). *Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF): Technical manual*. University of Minnesota Press. (Original work published 2008).

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*(1), Article 249. https://doi.org/10.1186/s12888-016-0939-x

Tylicki, J. L., Martin-Fernandez, K. W., & Ben-Porath, Y. S. (2019). Predicting therapist ratings of treatment progress and outcomes with the MMPI-2-RF. *Journal of Clinical Psychology*, *75*(9), 1673–1683. https://doi.org/10.1002/jclp.22795

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.

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLOS ONE*, *14*(11), Article e0224365. https://doi.org/10.1371/journal.pone.0224365

van Asselt, A. D., 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 der Heijden, P., Derksen, J., Egger, J., Rossi, G., Laheij, M., & Bögels, T. (2013). *MMPI-2-RF, Minnesota Multiphasic Personality Inventory-2 Restructured Form: Handleiding voor Afname, scoring en interpretatie* (Manual for administration, scoring and interpretation: Dutch-Flemish translation). PEN Tests Publisher.

van der Hoeken, D., Riet-van Hoof, K., & Hoek, H. W. (2000). *Nederlandse WHODAS II: Vragenlijst voor het vaststellen van beperkingen* [Dutch WHODAS II: Questionnaire to assess disability]. WHO Collaborating Centre WHODAS.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, *7*(1), Article 91. https://doi.org/10.1186/1471-2105-7-91

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

Wagner, T., Fydrich, T., Stiglmayr, C., Marschall, P., Salize, H. J., 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

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), Article 89. https://doi.org/10.1186/s12888-021-03670-9

Wilks, C. R., Korslund, K. E., Harned, M. S., & Linehan, M. M. (2016). Dialectical behavior therapy and domains of functioning over two years. *Behaviour Research and Therapy*, *77*, 162–169. https://doi.org/10.1016/j.brat.2015.12.013

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. Guilford Press.

Zahediabghari, S., Boursiquot, P., & Links, P. (2020). Impact of psychotherapy on psychosocial functioning in borderline personality disorder patients. *International Journal of Environmental Research and Public Health*, *17*(12), Article 4610. https://doi.org/10.3390/ijerph17124610

Zanarini, M. C., Frankenburg, F. R., Hennen, J., & Silk, K. R. (2003). The longitudinal course of borderline psychopathology: 6-year prospective follow-up of the phenomenology of borderline personality disorder. *The American Journal of Psychiatry*, *160*(2), 274–283. https://doi.org/10.1176/appi.ajp.160.2.274

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., Wedig, M. M., Conkey, L. C., & Fitzmaurice, G. M. (2014). Prediction of time-to-attainment of recovery for borderline patients followed prospectively for 16 years. *Acta Psychiatrica Scandinavica*, *130*(3), 205–213. https://doi.org/10.1111/acps.12255

Zanarini, M. C., Temes, C. M., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. M. (2018). Description and prediction of time-to-attainment of excellent recovery for borderline patients followed prospectively for 20 years. *Psychiatry Research*, *262*, 40–45. https://doi.org/10.1016/j.psychres.2018.01.034

Zeitler, M. L., Bohus, M., Kleindienst, N., Knies, R., Ostermann, M., Schmahl, C., & Lyssenko, L. (2020). How to assess recovery in borderline personality disorder: Psychosocial functioning and satisfaction with life in a sample of former DBT study patients. *Journal of Personality Disorders*, *34*(3), 289–307. https://doi.org/10.1521/pedi_2018_32_394

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

(*Appendix follows*)

Appendix

Descriptive Statistics for the Minnesota Multiphasic Personality Inventory–2 Restructured Form Scales and WHODAS 2.0 at Baseline ($N = 130$)

Scale	Raw score		*T*-score	
	M	*SD*	*M*	*SD*
Higher Order Scales				
EID	29.01	5.87	75.22	7.35
THD	4.35	3.26	60.17	12.79
BXD	8.05	3.63	68.66	12.51
Restructured Clinical Scales				
RCd	18.99	3.66	76.66	6.40
RC1	11.21	5.38	66.35	13.42
RC2	10.10	3.19	66.27	11.91
RC3	7.68	3.27	55.33	11.14
RC4	8.69	3.49	75.27	13.30
RC6	3.30	2.38	61.66	14.12
RC7	13.85	4.36	71.44	10.82
RC8	4.62	3.31	62.73	13.32
RC9	12.20	4.62	60.70	12.39
Specific Problem Scales				
MLS	5.83	1.60	68.48	9.02
NUC	3.62	2.22	62.61	14.00
COG	5.84	2.13	76.12	10.79
SUI	2.14	1.62	76.55	19.03
HLP	2.65	1.43	63.05	12.63
SFD	3.45	0.89	72.27	8.17
NFC	5.97	2.20	67.30	11.30
STW	5.38	1.21	72.31	8.10
AXY	1.92	1.50	70.37	15.52
ANP	4.13	2.09	68.72	13.30
BRF	1.82	1.69	61.21	13.68
JCP	2.14	1.65	64.42	13.88
SUB	2.04	1.78	67.11	19.43
AGG	3.87	1.97	68.84	13.56
ACT	4.22	2.06	60.60	14.63
FML	5.02	2.63	70.72	14.64
IPP	4.70	2.52	52.41	10.70
SAV	5.31	2.91	54.15	13.47
SHY	3.64	2.52	54.98	12.17
DSF	1.19	1.37	55.18	11.96
Personality Psychopathology Five Scales, Revised				
AGGR-r	7.63	3.51	51.78	11.23
PSYC-r	4.62	3.38	59.33	12.38
DISC-r	7.01	3.07	64.43	12.31
NEGE-r	13.47	2.56	73.22	8.44
INTR-r	10.36	4.49	55.78	13.56
WHODAS 2.0	1.54	0.61		

Note. WHODAS 2.0 = World Health Organization Disability Assessment Schedule 2.0; EID = Emotional/Internalizing Dysfunction; THD = Thought Dysfunction; BXD = Behavioral/Externalizing Dysfunction; RCd = Demoralization; RC1 = somatic complaints; RC2 = Low Positive Emotions; RC3 = Cynicism; RC4 = Antisocial Behavior; RC6 = ideas of persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; MLS = Malaise; NUC = Neurological Complaints; COG = Cognitive Complaints; SUI = suicidal/death ideation; HLP = Helplessness/Hopelessness; SFD = Self-Doubt; NFC = Inefficacy; STW = Stress/Worry; AXY = Anxiety; ANP = anger proneness; BRF = Behavior-Restricting Fears; JCP = Juvenile Conduct Problems; SUB = Substance Abuse; AGG = Aggression; ACT = Activation; FML = family problems; IPP =Interpersonal Passivity; SAV = Social Avoidance; SHY = Shyness; DSF = Disaffiliativeness; AGGR-r = Aggressiveness-Revised; PSYC-r = Psychoticism-Revised; DISC-r = disconstraint-revised; NEGE-r = negative emotionality/neuroticism-r evised; INTR-r = Introversion/Low Positive Emotionality-Revised.

Received August 30, 2024
Revision received February 17, 2025
Accepted February 20, 2025 ■