

EMPIRICAL PAPER

Investigating predictors of treatment response in Dialectical Behavior Therapy for borderline personality disorder using LASSO regression

QINGQING YIN ¹, MOLLY STERN², EVAN M. KLEIMAN ¹, & SHIREEN L. RIZVI ²

¹Department of Psychology, Rutgers University, New Brunswick, NJ, USA & ²Graduate School of Applied and Professional Psychology, Rutgers University, New Brunswick, NJ, USA

(Received 16 July 2022; revised 12 October 2022; accepted 13 October 2022)

Abstract

Objective Prior studies of Dialectical Behavior Therapy (DBT) for borderline personality disorder (BPD) have yielded heterogeneous findings on what factors differentiate individuals with or without sufficient treatment response, highlighting the need for further research.

Method We investigated a sample of 105 individuals with BPD receiving a 6-month course of DBT. Participants were categorized as sufficient or insufficient responders using clinical and statistical change indices (based on emotion dysregulation, BPD symptom severity, utilization of DBT skills, and functional impairment). Sociodemographic, clinical severity, and treatment process factors were tested as potential predictors of treatment response using a machine learning approach (LASSO regression).

Results Two cross-validated LASSO regression models predicted treatment response (AUCs > .75). They suggested that higher homework completion rate, retention in treatment, and greater baseline severity were the most important predictors of DBT treatment response indicated by BPD symptom severity and utilization of DBT skills. Favorable effects of some aspects of therapeutic alliance during initial sessions were also found.

Conclusions Future research may benefit from consolidating the criteria of treatment response, identifying clinically relevant variables, and testing the generalizability of findings to enhance knowledge of insufficient treatment response in DBT for BPD.

Keywords: borderline personality disorder; Dialectical Behavior Therapy; treatment response; machine learning; LASSO regression

Clinical or Methodological Significance of this Article: With a naturalistic clinical sample and a machine learning approach, the study found that homework completion, retention in treatment, and greater baseline severity increased the likelihood of sufficient treatment response, as indicated by BPD symptom severity and DBT skills use, after six months of Dialectical Behavior Therapy for borderline personality disorder. Such research is critical for early identification of those in need of enhanced treatment and making effective decisions to improve client outcomes in a more timely fashion.

Borderline personality disorder (BPD) is a serious psychological disorder characterized by pervasive emotion instability, interpersonal conflicts, identity disturbance, and impulsive and risky behaviors (American Psychiatric Association [APA], 2013). Individuals with BPD often report suicidal thinking and behaviors and experience severe distress and functional impairments (World Health Organization, 2001). One of the most studied treatment for BPD is

Dialectical Behavior Therapy (DBT), a multi-component behavior therapy that involves teaching change- and acceptance-oriented coping strategies to enhance clients' capabilities for emotion regulation. DBT was originally developed as a treatment for suicidal individuals with BPD, for which it has the most empirical support (Kliem et al., 2010; Panos et al., 2014). Its use has also been expanded broadly to populations with transdiagnostic

*Correspondence concerning this article should be addressed to Qingqing Yin, Department of Psychology, School of Arts and Sciences, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854. Email: qy93@rutgers.edu

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

emotion regulation difficulties (e.g., Chen et al., 2008; Harned et al., 2014; Neacsu et al., 2014).

Despite ample empirical support for DBT for treating BPD, a lack of treatment response in some clients exists as a crucial issue (Rizvi, 2011). Synthesizing knowledge about treatment response is difficult, given no agreed-upon threshold for determining sufficient response. One operationalization of treatment response includes both (1) clinically relevant change, when sufficient improvement from dysfunctional toward functional range is observed, and (2) reliable change, the statistical significance of change accounting for measure reliability (e.g., Jacobson & Truax, 1991). Based on this operationalization, across several inpatient and outpatient studies of DBT for BPD, insufficient response is roughly estimated to range between 48% and 58% (Bohus et al., 2004; Kröger et al., 2013; Rizvi et al., 2017; Schmitgen et al., 2019; 53% in Woodbridge et al., 2022).

The room for further improvement in DBT's efficacy is apparent given the implications of insufficient treatment response, such as continued suffering, decreased quality of life, and increased mortality rates (Gloster et al., 2020). Early identification of those who do not sufficiently respond to treatment can help providers and consumers make more effective decisions about treatment. It may also provide researchers and treatment developers with necessary information to enhance treatment outcomes. However, several limitations characterize the relevant literature, including the few number of studies, heterogeneity in examined variables and findings across studies, and a lack of emphasis on researching modifiable treatment processes (e.g., attendance and dropout) that may inform targets for further intervention.

To date, there is a paucity of research examining factors that prospectively distinguish insufficient responders from responders in DBT for BPD. Bohus et al. (2004) examined an array of sociodemographic characteristics (e.g., current employment, age) and pre-treatment severity variables (e.g., mental health comorbidities, lifetime suicide attempts, general psychopathology, and BPD severity) among individuals with BPD receiving an inpatient DBT program. They found that none of the factors discriminated between responders and insufficient responders. McMain et al. (2018) categorized individuals into latent subgroups based on suicide and self-injurious behaviors over one year of outpatient comprehensive DBT or general psychiatric management for BPD. Consistent with Bohus et al.'s (2004) findings, they did not find significant differences in pre-treatment general psychopathology, BPD symptom severity, or mental health

comorbidities between groups. However, higher levels of depression, more emergency department visits, and unemployment at pre-treatment were found to be associated with the poorest response. Similarly, Berk et al. (2022) identified latent subgroups among adolescents with BPD symptoms who received six months of DBT or individual and group supportive therapy depending on their improvement in suicidal thinking and self-injurious behaviors. Their results suggested that more frequent suicide thinking and self-injury, parent-reported youth externalizing problems, and racially identifying as White were predictors of insufficient treatment response in DBT.

The studies above represent the few prior attempts to investigate what makes an individual with BPD less likely to respond to treatment of DBT. Unfortunately, this literature fails to provide a conclusive picture, due to the limited consistent findings across studies. It is noteworthy that most other studies on treatment response in DBT investigated what factors were positively or negatively correlated with symptom or functioning scores after treatment or the magnitude of improvement. These studies, though providing potential predictors of treatment response (e.g., higher symptom severity, self-injury history, employment), also yielded conflicting findings regarding the influences (i.e., positive or negative) of baseline clinical severity and sociodemographic characteristics (e.g., Adrian et al., 2019; Harned et al., 2010; Herzog et al., 2020). Furthermore, they are limited due to the lack of a cutoff for sufficient or insufficient response, which makes identification of those in need of further intervention difficult.

In addition, fewer studies have examined modifiable treatment processes related to response in DBT, such as attendance and engagement. Edwards et al. (2021) found greater completion of skills use homework among clients who did not engage in various problem behaviors (e.g., substance use, self-injury) during the last month of an inpatient DBT program compared to those who did, highlighting an important role of homework completion in treatment. Regarding treatment dropout, although we acknowledge the practicality of excluding dropouts in evaluations of treatment response, accounting for them may help answer important questions, such as whether individuals who drop out from treatment do not benefit sufficiently.

Taken together, the existing literature on factors associated with treatment response in DBT for BPD has notable limitations. There is a clear need for further research investigating and replicating what factors may predict treatment response in DBT. For this objective, machine learning methods

may offer more robust predictions addressing multicollinearity among a large number of variables examined in a single model and augment generalizability of results (Aafjes-van Doorn et al., 2021). To our knowledge, no study to date has involved machine learning in predicting response in comprehensive DBT for BPD based on an explicit operationalization (or cutoff) of treatment response. Least Absolute Shrinkage and Selection Operator (LASSO) regression is a machine learning method, which mitigates the problems in traditional regression methods with many predictors that are likely correlated in a way to identify the most relevant predictors, thus increasing parsimony (i.e., fewer predictors). It applies a penalty to all regression coefficients to allow some to be shrunk towards zero and some to be discarded, thereby performing variable selection. Using LASSO regression, the current study aimed to identify sociodemographic characteristics, clinical severity markers, and treatment process variables that predict whether an individual with BPD would sufficiently respond to a 6-month outpatient DBT program.

Methods

Participants

The participants in the study were adults enrolled in a university-based Dialectical Behavior Therapy research and training clinic between September 2010 and May 2021. The inclusion criteria for participation were: meeting criteria of BPD diagnosis via structured interview; aged 18 years or older; agreement to complete assessments; consent to session recording; availability to receive treatment for six months; residing within 45 min of the clinic; and agreement to discontinue all other psychotherapy. The sample included a subset of participants ($n = 13$; 12.38% of final sample) that was part of a sub-study on a mobile phone app (see Rizvi et al., 2017) that required a higher severity threshold: two or more instances of either non-suicidal self-injury (NSSI) or suicide attempt in the last five years, with one instance occurring in the past six months; and agreement to carry a mobile device. The exclusion criteria were: mental health problems that require services that were not provided by the clinic (e.g., schizophrenia, life-threatening anorexia); non-English speaking; an indication that the client has an IQ of 70 or below; and inability to consent for research.

A total of 118 individuals were deemed eligible (i.e., comprising the intent-to-treat sample) and offered six months of comprehensive DBT

treatment. Among intent-to-treat individuals, 11.02% ($n = 13$) did not complete post-treatment assessment and were excluded. Due to incompleteness of some but not all measures by some participants, the final sample size differed for each outcome, ranging from 99 to 105 participants, including approximately 11.43% ($n = 12$) who dropped out but completed post-treatment assessment.

Demographics were evaluated for 105 participants. The average age was 29.11 years old ($SD = 10.34$). For gender, 76.19% ($n = 80$) of sample identified as female, 20.95% ($n = 22$) identified as male, and 2.86% ($n = 3$) identified as others. Race/ethnicity breakdown was 79.05% ($n = 83$) White, 8.57% ($n = 9$) Asian, 8.57% ($n = 9$) mixed race, 2.86% ($n = 3$) Black, and 0.95% ($n = 1$) other races/ethnicities. The sample included 72.38% ($n = 76$) of individuals who were single and never married, 16.19% ($n = 17$) who were married or living with partner, and 11.43% ($n = 12$) who were separated, divorced, or widowed. In terms of employment, 16.67% ($n = 17$) were full-time employed, 28.43% ($n = 29$) were part-time employed, 1.96% ($n = 2$) were volunteers, 20.59% ($n = 21$) were students, and 32.35% ($n = 33$) were unemployed. In terms of education, 11.43% ($n = 12$) had high school or less education, 52.38% ($n = 55$) had some college, 25.71% ($n = 27$) graduated from college, and 20.48% ($n = 11$) had more advanced degrees or education. For annual income, 36.58% ($n = 35$) reported \$0-\$9,999, 28.13% ($n = 27$) reported \$10,000-\$39,999, 11.46% ($n = 11$) reported \$40,000-\$69,999, 15.63% ($n = 15$) reported \$70,000-\$99,999, and 8.33% ($n = 8$) reported \$100,000 or more.

At baseline, the sample met criteria for an average of 4.25 ($SD = 2.53$) comorbid lifetime psychological disorders, and an average of 3.48 ($SD = 2.29$) current comorbid disorders. Major depressive, persistent depressive, generalized anxiety, and social anxiety disorders were the most common comorbidities for both lifetime and current presentation. A total of 58.1% ($n = 61$) of individuals made at least one suicide attempt in their lifetime, and the average number of attempts across these individuals was 8.52 ($SD = 20.19$, range = 1-114). Also, 64.76% ($n = 68$) of the sample engaged in NSSI within the year before baseline, and the average number of lifetime NSSI was 380 ($SD = 1125.32$, range = 1-7500).

Procedures

Participants were self-referred or referred by local agencies or clinicians. Interested individuals

completed a brief screening assessment over the phone. Those deemed eligible based on phone screen were then scheduled to complete the informed consent and further assessments, including demographic measures, diagnostic interviews, an interview of suicide and self-injury histories, and other clinical outcome measures, to further assess eligibility and levels of functioning before treatment. Eligible participants then took part in six months of comprehensive DBT and completed a post-treatment assessment at the end of the 6-month treatment that included the same outcome measures as baseline. Additionally, measures on therapeutic alliance were completed after each of the first four individual sessions (see below). To record group homework completion, skills training leaders rated members' homework as complete, partial, or none during each group (see below).

All assessments and treatments were conducted by trained clinical psychology doctoral students and were fully conducted in person until March 2020, when services were moved to telehealth ($n = 5$, 4.76%) due to the Coronavirus-2019 pandemic. The assessors were instructed to confirm the validity of self-report responses based on completeness of data, variation in response, and time spent in appointment to complete, and ask participants to redo the measures if the validity was reasonably questionable. Additionally, participants were compensated up to \$60 for completing the post-treatment assessment. Individuals were informed that the assessments are separate from treatment, and those who dropped out were still contacted to complete post-treatment assessment. The university's Institutional Review Board approved all study procedures.

Treatment. Participants took part in six months of comprehensive DBT (Linehan, 1993, 2014), defined as DBT that includes all treatment components. Specifically, the treatment comprised weekly individual therapy, weekly group skills training, inter-session phone coaching as needed, and therapist consultation team. Prior research suggested that DBT can be delivered by trainee therapists at this clinic to produce client improvements comparable to a gold standard clinical trial (Rizvi et al., 2017). In addition to consultation team, all therapists and assessors participated in a weekly didactic training on advanced topics in DBT (e.g., case conceptualization, secondary targets) and individual supervision. They had also completed a doctoral-level semester-long course in DBT at the university in which the clinic is based, which covers the theories, principles, strategies, and research of DBT. Participants were charged \$10 to \$100 per week for

treatment sessions, based on a sliding scale fee structure according to annual income.

Measures

Treatment response. We categorized participants' levels of treatment response based on ratings to specific measures (see below) at pre- and post-treatment assessments. Reliable change (RCI) and clinically significant change (CSC) indices were obtained through widely adopted criteria (Jacobson & Truax, 1991), and Wise's (2004) method was used to integrate RCI and CSC. RCI was calculated as $X_{post} - X_{pre}/S_{diff}$ where X = scores on any given measure and S = the standard error of the difference between scores while accounting for test-retest reliability. RCI scores greater than 1.96 indicate that the change over time is statistically significant. Due to the lack of non-patient normative data, CSC was defined as greater than two standard deviations away from the pre-treatment sample mean in the direction of functionality.

Using Wise's (2004) method of integrating RCI and CSC, participants were categorized for varying levels of treatment response¹ for further evaluation, such as "recovered" given an RCI greater than 1.96 and a change of two standard deviations in the direction of functionality; "positive response" given an RCI greater than 1.28 and a change of one standard deviation; "minimal positive response" given an RCI greater than 0.84 and a change of 0.5 standard deviations; "minimal negative response" given an RCI lower than -0.84 and a change of -0.5 standard deviations; and "deteriorated" given an RCI lower than -1.96 and a change of -2 standard deviations. For the purpose of analysis, participants categorized to have gained "minimal positive" or greater response were subsequently coded as treatment "responders," while participants categorized to have gained less than "minimal positive" or any negative response were coded as "insufficient responders."

Outcome variables. Four variables relevant to the treatment of DBT for BPD were examined for treatment response considering the importance of both symptom and functional improvements and the multi-dimensional nature of BPD treatment outcomes. Prior research demonstrated the internal consistency and test-retest reliability of scores of all measures used (see below). In the current sample, scores of all measures showed Cronbach's $\alpha > .85$.

Emotion Dysregulation. Emotion regulation difficulty was measured by the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004), a 36-item self-report questionnaire assessing various

Table 1 List of Predictor Variables and Descriptive Statistics

Demographic and Baseline Predictors		
	<i>M (SD) or n (%)</i>	Proportion Missing
Age	29.11 (10.34)	
Gender	female: 80 (76.19%); male: 22 (20.95%); others: 3 (2.86%)	
Employment	full/part-time: 33 (32.35%); volunteer/student: 23 (21.90%); unemployed 46: (45.10%)	2.86%
Education	high school grad/below: 12 (11.43%); some college: 55 (52.38%); college grad/beyond: 38 (36.19%)	
Ever made a suicide attempt	61 (58.10%)	
Engaged in NSSI in past year	68 (64.76%)	1.90%
Received medical treatment for most recent suicide attempt	31 (51.67%)	
Received medical treatment for most recent NSSI in past year	5 (7.04%)	1.90%
Number of BPD criteria met	7.11 (1.26) ^a	2.86%
Number of lifetime comorbidities	4.25 (2.53)	0.95%
Treatment Process Predictors		
	<i>M (SD) or n (%)</i>	
Treatment dropout	12 (11.43)	
Number of individual sessions	23.90 (7.55)	
Number of group sessions attended	19.57 (5.61) ^b	
Homework completion rate	0.73 (0.24)	
WAI-C Goal subscale	24.37 (3.05)	
WAI-C Task subscale	23.97 (3.47)	
WAI-C Bond subscale	23.56 (4.03)	
WAI-T Goal subscale	22.69 (2.57)	
WAI-T Task subscale	21.66 (3.01)	
WAI-T Bond subscale	22.25 (2.74)	

Note. NSSI = non-suicidal self-injury. BPD = borderline personality disorder. WAI = Working Alliance Inventory-Short Form.

^aOut of a total of 9 BPD symptoms.

^bOut of a possible total of 25 group sessions.

dimensions of difficulties in emotion regulation (e.g., “When I’m upset, I lose control over my behaviors”). The items are responded are on a Likert scale of 1 (“almost never”) to 5 (“almost always”).

BPD Symptom Severity. BPD symptomatology was measured by the Borderline Symptoms List—Short Form (BSL-23; Bohus et al., 2004). It is a 23-item self-report questionnaire (e.g., “My mood rapidly cycled in terms of anxiety, anger, and depression”), which participants respond to on a Likert scale of 0 (“not at all”) to 4 (“very strong”).

DBT Skills Utilization. Use of DBT coping skills was measured by the DBT Skills Subscale (DSS) of DBT Ways of Coping Checklist (DBT-WCCL; Neacsiu et al., 2010). DBT-WCCL is a 59-item self-report questionnaire including two subscales: one assessing coping via DBT skills (DSS; e.g., “Changed something about myself so that I could deal with the situation better”) and the other assessing dysfunctional coping (DCS; e.g., “Refused to believe that it had happened”). All items are responded using a scale of 0 (“never use”) to 3 (“always use”).

Impairment in Work and Social Adjustment. Impairment in work and social adjustment was measured by the Work and Social Adjustment Scale (WSAS; Mundt et al., 2002), a 5-item self-report questionnaire of general functioning (e.g., “Because of my [disorder], my ability to work is impaired”). All items are responded using a Likert scale of 0 (“not at all impaired”) to 8 (“very severely impaired”).

Predictor variables. A total of 11 demographic and baseline predictors as well as 10 treatment related predictors were examined for the prediction of each outcome. We selected variables previously shown to be associated with treatment response or outcome (Adrian et al., 2019; Edwards et al., 2021) or yielded mixed findings (Bohus et al., 2004; Herzog et al., 2020; McMain et al., 2018) across multiple studies, including age, employment status, histories of suicide attempts and self-injurious behaviors, mental health comorbidities, pre-treatment BPD symptoms, and skills use homework completion. We also included treatment dropout and

associated variables (Herzog et al., 2020; Kröger et al., 2013; Landes et al., 2016; Perroud et al., 2010; Thomas, 2017), including education level and therapeutic alliance during initial treatment sessions. Additionally, we included other variables theoretically relevant to BPD treatment seeking or engagement that may be associated with treatment response, including gender (Goodman et al., 2010), number of individual and group sessions attended, and baseline rating of the outcome variables tested. See Table 1 for a list of predictor variables.

Demographics. Participants answered a series of demographic questions at baseline, including age, gender identity, current employment status, and education level. For the purpose of analyses, responses to employment status were collapsed into three levels: full- or part-time employed, volunteer or student, and unemployed. Responses to education were collapsed into three levels: high school graduate or below, some college, and college graduate or beyond.

Diagnostic Interviews. Diagnoses based on the Diagnostic and Statistical Manual of Mental Disorders (DSM IV: APA, 2000; or DSM 5th edition: APA, 2013) were assessed at baseline. The Structured Clinical Interview-IV Axis I Disorders (SCID-I; First et al., 2002) for DSM-IV or Structured Clinical Interview for DSM-5 (SCID-5; First et al., 2015) was used for assessing mental health comorbidities. During the course of study, we replaced the use of SCID-I with SCID-5 for the most recent 61 (58.10%) participants. Structured Clinical Interview for DSM-IV Axis II Disorder (SCID-II; First et al., 1997) with excellent interrater reliability (Lobbstaël et al., 2011) for BPD categorization was used to determine BPD diagnosis and the number of criteria met.

Suicide and Self-Injurious Behaviors. The Suicide Attempt Self-Injury Interview (SASII; Linehan et al., 2006) or the Self-Injurious Thoughts and Behaviors Interview (SITBI; Nock et al., 2007) was used at baseline to assess histories of suicide attempt and NSSI. In the study, the SASII was replaced by SITBI for the most recent 86 (81.90%) participants. Equivalent items on SASII and the SITBI were matched to provide consistent measures on whether an individual ever made a suicide attempt, whether medical treatment was received for the most recent attempt, whether NSSI was engaged in the past year, and whether medical treatment was received for the most recent NSSI in the past year. Prior research demonstrated the interrater reliability of SASII (Linehan et al., 2006) and SITBI (Nock et al., 2007).

Treatment Engagement. As specified in the manual for DBT, individuals who missed four consecutive individual or group sessions were considered

dropped out of the program (Linehan, 1993). The numbers of individual and group sessions attended were recorded. Furthermore, completion rate of weekly skills group homework was calculated from ratings by skills training leaders (0 = none; 1 = partial; and 2 = complete), by dividing the sum of ratings across all group sessions attended by the highest total ratings possible.

Therapeutic Alliance. Therapeutic alliance was assessed by the three subscales of Working Alliance Inventory-Short Form (Hatcher & Gillaspay, 2006), a 12-item abbreviated version of the 36-item Working Alliance Inventory (Horvath & Greenberg, 1989). WAI has a client (WAI-C) and a therapist (WAI-T) version and includes three subscales assessing agreement on treatment goals (Goal; e.g., “We are working toward mutually agreed goals”), agreement on treatment tasks (Task; e.g., “I believe the way we are working with my problem is correct” or “My client believes the way we are working with their problem is correct”), and affective bond (Bond; e.g., “I feel that my treatment provider appreciates me as a person” or “I appreciate my client as a person”). Responses to the items are on a Likert scale of 1 (“never”) to 7 (“always”). This measure was completed by clients and therapists after each of the four initial individual sessions, and the average subscale scores across sessions were calculated. The WAI subscales showed internal consistency in prior research (Hatcher & Gillaspay, 2006) and current sample ($\alpha = .75-.92$).

Analytic Strategy

Two LASSO logistic regression models for each outcome variable were ran to select relevant features from 11 baseline and 10 treatment process predictors,² respectively, in R (R core Team, 2021; Glmnet package: Friedman et al., 2010). We tried to limit the number of predictors in each model for power considerations. Missingness in predictor (0.95%–2.86%) and outcome (4.76%–5.7%) variables for each model was addressed via list-wise deletion due to the small amount. We used 5-fold cross-validation, in which observations were randomly divided into five subsets each composed of 20% of observations; 4/5 of data from each of the five subsets was used for training (i.e., developing the model) with the remaining 1/5 of each subset used for testing (i.e., testing model performance). In estimation of model, an optional penalty parameter (lambda) was chosen via cross-validation using the “cv.glmnet” function of Glmnet. The function runs to get a sequence of lambda and then computes the fit with each of the folds tested, with error

accumulated and mean error across folds calculated. The value of lambda that yields the minimum mean cross-validated error was then selected.

The results of cross-validation were evaluated for each of the eight models by comparing observed and predicted classifications (threshold at .50) of each participant as a “responder” or “insufficient responder,” via area under the receiver operating curve (AUC). The better the model performs in classification, the higher the AUC value is with 1 indicating perfect prediction, and an AUC of .70 or higher was considered to indicate acceptable discrimination ability of model. Accuracy, which is the percentage of correct classifications, was also reported. Additionally, the relative importance of variables was evaluated (vip package: Greenwell et al., 2020) by computing standardized absolute coefficients that reflect the relative influence of each predictor on the outcome variable in the prediction model.

Results

Descriptive statistics of predictor variables are presented in Table 1. Evaluation of treatment response is summarized in Table 2. It was observed in the study that the rates of treatment response were approximately 78% for emotion regulation difficulties, 65% for BPD symptoms, 55% for utilization of DBT skills, and 50% for work and social impairments. None of the four baseline models (DERS: AUC = .63, accuracy = 23%; BSL: AUC = .64, accuracy = 64%; WSAS: AUC = .66, accuracy = 48%; DSS: AUC = .69, accuracy = 63%), and neither of the two treatment process models involving DERS (AUC = .60, accuracy = 22%) and WSAS (AUC

= .67, accuracy = 66%) achieved acceptable discrimination ability (AUC > .70) via cross validation, despite that some were close to the threshold. We considered it to indicate that in each of these models, the predictors together were likely not able to differentiate responders from insufficient responders reliably. Thus, only the two remaining models involving BSL-23 and DBT WCCL-DSS scores and treatment process predictors are reported in detail here.

Predicting Outcome of BPD Severity from Treatment Process Factors

Based on BSL-23 scores at baseline and post-treatment, 64.65% of participants were categorized as treatment responders and 35.35% were categorized as insufficient responders. Of the participants who dropped out of treatment, 25% were classified as responders. The mean AUC of the cross-validated LASSO model across five folds was .77 (see ROC curves in Figure 1, left) for predicting whether a participant was responder or insufficient responder based on BSL-23 scores (accuracy = 73%). Predictors retained in the model that increased the likelihood of treatment response based on BSL-23 were higher percentage of homework completion, higher baseline BSL-23 scores, as well as higher client-rated agreement on goals (WAI-C goal) and personal bonding (WAI-C bond) with therapist. Dropout was shown to decrease the likelihood of treatment response. Greater number of individual session was also associated with decreased likelihood of treatment response, albeit a near-zero variable importance coefficient (<.003). The relative importance of these predictors is shown in Figure 2 (left).

Table 2. Reliable and clinically significant change indices.

	DERS (<i>n</i> = 105)	BSL-23 (<i>n</i> = 99)	DBT WCCL-DSS (<i>n</i> = 99)	WSAS (<i>n</i> = 100)
Response	82 (78.1%)	64 (64.65%)	53 (54.64%)	50 (50%)
Recovered	22	20	26	13
Positive response	26	27	19	28
Minimal positive response	34	17	10	9
Insufficient response	23 (21.9%)	35 (35.35%)	44 (45.36%)	50 (50%)
No categorical response	21	30	38	38
Minimal negative response	2	4	4	12
Deteriorated	0	1	2	0
Baseline vs. Post-assessment ^a	122.09 (22.50)	2.01 (0.79)	1.31 (0.49)	23.48 (8.06)
	88.98 (33.68)	1.24 (0.89)	1.77 (0.57)	17.16 (10.21)

Note. DERS = Difficulties in Emotion Regulation Scale. BSL-23 = Borderline Symptoms List–Short Form. DBT WCCL-DSS = DBT Ways of Coping Checklist-DBT Skills Subscale. WSAS = Work and Social Adjustment Scale. Only categories from Wise (2004) that included at least 1 participant in the current study are listed here.

^a*M(SD)*s are reported.

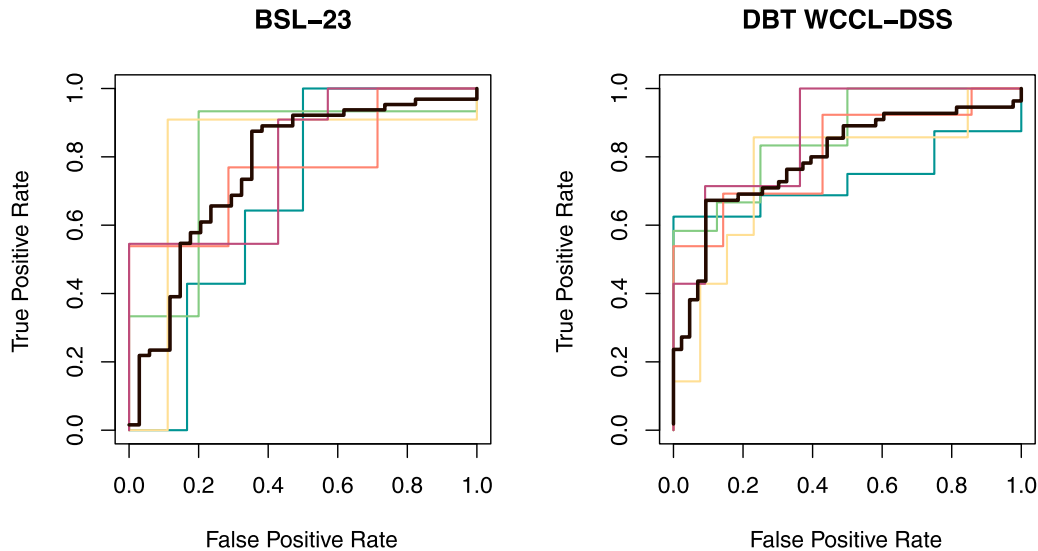


Figure 1. ROC curves of 5-fold cross-validation models for the prediction of treatment response from treatment process factors.

Note. ROC = receiver operating curve. Colored lines represent folds. Black line represents the prediction of cross-validated model of all observations. BSL-23 = Borderline Symptoms List-Short Form. DBT WCCL-DSS = DBT Ways of Coping Checklist-DBT Skills Subscale.

Predicting Outcome of DBT Skills Use from Treatment Process Factors

According to DBT WCCL-DSS (hereafter referred to as “DSS”) scores, 54.64% of participants were categorized as treatment responders and 45.36% were categorized as insufficient responders. Of treatment dropouts, 25% responded. Informed by treatment process predictors, the mean AUC of the cross-validated model across five folds was .81 (see

ROC curves in Figure 1, right) for categorizing participants as treatment responders or insufficient responders (accuracy = 71%). Predictors retained in the model that increased the likelihood of treatment response based on DSS were higher percentage of homework completion, higher client-rated personal bonding with therapist (WAI-C Bond) and agreement on goals (WAI-C Goal), as well as greater number of group sessions attended. Predictors

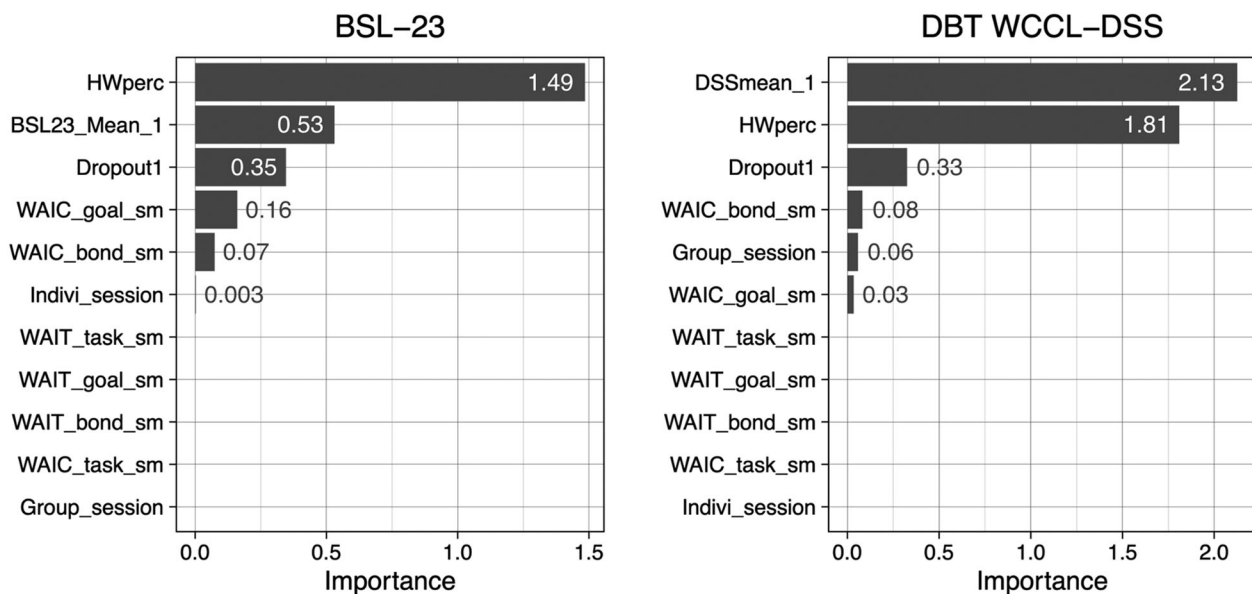


Figure 2. Variable importance of treatment process factors in predicting treatment response.

Note. BSL-23 = Borderline Symptoms List-Short Form. DBT WCCL-DSS = DBT Ways of Coping Checklist-DBT Skills Subscale. WAI = Working Alliance Inventory-Short Form (T = therapist-rated; C = client-rated); sm = mean across sessions. Variable mean 1 = Baseline score.

retained that decreased the likelihood of treatment response were higher baseline DSS scores and dropout. The relative importance of these predictors is shown in [Figure 2](#) (right).

Comparison with Models Including Baseline Ratings of Outcome Only

Due to the intuitive nature of baseline outcome ratings predicting treatment response, we conducted simple linear regressions including only the baseline scores to compare with the two LASSO models. The AUCs were .64 and .66 (accuracy = 67–72%) for models of BSL-23 and DSS, respectively, suggesting that the performance of these models were worse than the LASSO models involving other treatment process predictors (AUCs >.75).

Discussion

The current study aimed to investigate variables that may discriminate between individuals with and without sufficient response in DBT for treatment of BPD. Such information is critically important for early identification of those in need of enhanced treatment and making effective decisions to improve client outcomes in a more timely fashion. Despite increasing interest in machine learning in broader psychotherapy literature (Aafjes-van Doorn et al., 2021), to our knowledge, this is the first study to adopt a machine learning approach to select predictors of treatment response in comprehensive DBT for BPD, which mitigates problems of overfitting in traditional methods with large numbers of variables and provides parsimonious models of prediction. Another main strength of the study lies in the naturalistic sample with BPD receiving DBT in a real-world community setting, where the ecological validity of research is enhanced.

In this study, the rate of response was the highest for emotion regulation outcome, followed by BPD symptom severity, utilization of DBT skills, and work and social adjustment. The current rate of response in DBT indicated by BPD symptoms severity (65%) was higher than that (53%) reported in a recent review (Woodbridge et al., 2022). Given the varying rates of response based on different outcomes measured in the current study (78% for emotion regulation difficulty vs. 50% for work and social impairment), a multidimensional approach of outcome measurement was confirmed to be necessary. This also provides a fairly good indication that improvements linked to dysregulated emotions occurred in the treatment of DBT for BPD and perhaps at a greater or faster rate than symptom amelioration,

skills acquisition, or general functional improvement, though it may not be the one and only mechanism of DBT (Lynch et al., 2006; Mehlum, 2021).

The study had several findings related to the prediction of response in the treatment of DBT for BPD. First, higher homework completion rate, retention in treatment, and greater baseline severity (greater BPD severity or lower endorsement of DBT skills utilization), emerged as the three most important predictors of treatment response evaluated through two metrics (BPD symptom severity and utilization of DBT skills). Although a greater number of skills training groups attended was positively associated with DBT skills utilization, greater homework completion across group sessions, independently from the effect of group attendance, was related to both BPD symptom and DBT skills utilization.

These results may have direct clinical implications as both homework completion and treatment retention are explicit foci in treatment of DBT to be addressed right after eliminating life-threatening behaviors in the treatment hierarchy (Linehan, 1993). The theoretical rationales for prioritizing these therapy-interfering behaviors, namely that individuals with BPD presentations are prone to quitting therapy prematurely and engaging in other avoidance and disengagement behaviors (Chalker et al., 2015; Linehan, 1993), seem to be corroborated by the current findings. It is plausible that indications of premature quitting and failure to complete weekly skills homework, if left insufficiently addressed, are indeed early signs of poorer treatment response, highlighting the need for continuing monitoring and timely intervention. Particularly, the finding regarding homework completion along with prior research showing a similar pattern (Edwards et al., 2021) provide empirical support for the importance of weekly skills homework. Such empirical evidence could also be utilized to enhance the rationale for addressing therapy interfering behaviors that is provided to clients in session. In addition, the finding that individuals who responded to treatment had more severe BPD symptoms and used less DBT coping strategies before starting treatment is informative. Although this could simply mean that there is greater room for improvement for those with higher baseline scores, an alternative interpretation is that the treatment was particularly effective for individuals presenting with difficult-to-treat problems at baseline. While therapists may feel less optimistic about treatment outcomes when working with individuals presenting with severe BPD symptoms and lacking effective coping strategies, these clients may actually show favorable prognosis.

A second important finding was that aspects of therapeutic alliance during initial sessions increased

the likelihood of treatment response in DBT for BPD. Both outcomes of BPD severity and DBT skills utilization were associated with client-rated agreement on goals and bonding with therapist. We further hypothesize that these findings reflect the effects of relevant therapeutic strategies. In DBT, therapists strategically make use of interaction and relationships with clients to facilitate their behavior change. An example of this is a therapist providing responsiveness when a client effectively makes a request (to reinforce effective behavior) and neutral affect when the client doesn't (to extinguish ineffective behavior). Another example is the therapist building a genuine relationship with the client so the therapist's positive regard toward the client functions as a natural reinforcer. Prior research has also highlighted favorable effects of the use of validation or other acceptance strategies in conjunction with behavioral change strategies. Use of validation, which communicates understanding and acceptance, was associated with increases in positive affect and decreases in negative affect (Carson-Wong et al., 2018). Across therapeutic dyads, client-rated therapist understanding and warmth also predicted reduction in self-harm and suicide urges one week later (Shearin & Linehan, 1994).

Third, we studied the effects of several sociodemographic and clinical severity factors, including age, gender, education, employment, BPD and comorbid diagnoses, and histories of suicide attempt in lifetime and self-injury within the past year (and whether medical treatment was received for these behaviors), on treatment response in DBT for BPD. This set of sociodemographic and clinical severity factors at pre-treatment was not found to reliably discriminate between individuals with and without sufficient response across outcome measures. However, this may have provided some clarification for the mixed literature (e.g., Adrian et al., 2019; Bohus et al., 2004; Herzog et al., 2020; McMain et al., 2018). It is possible that we may not be able to accurately predict treatment outcome based on these client characteristics, in which case it could be iatrogenic to hold certain assumptions about treatment response based on client characteristics. The lack of accurate predictions in the LASSO models may also be due to chance or other reasons, such as the imbalanced data for some variables (e.g., only 7.04% participants reported "yes" to receiving medical treatment for NSSI in past year). We believe it is important for future research to continue replicating prior findings to help elucidate the mixed literature.

Additionally, there are important lessons to learn regarding using machine learning to address fundamental questions about psychotherapy treatment

response. First, machine learning models may help us generate theories of DBT response in ways that traditional statistical models do not. Second, there is perhaps a need to rethink about ways we measure treatment response related variables, such as increasing the emphasis on functional outcomes, measuring therapist characteristics and approach (e.g., therapist avoidance, inconsistent treatment targets; Rizvi, 2011), and measuring challenging client behaviors (e.g., frequently calling therapists; Chalker et al., 2015) during treatment. Moreover, synthesizing knowledge about insufficient treatment response in DBT is difficult due to the lack of a consistent definition of response. Wise's (2004) approach of integrating clinical and statistically significant change indices may be a good starting point.

Limitations

The current results should be interpreted in the context of several limitations. The sample size was not determined based on an a-priori power analysis. Further research is needed to provide methods of power analysis for machine learning models like LASSO regression (Riley et al., 2020). Our relatively small sample size could impact our ability to identify some smaller effect sizes (Riley et al., 2021). Also, due to the small sample, the cross-validated models we developed were not further tested in clean holdout datasets. Despite our attempt to reduce model overfitting, we want to acknowledge generalizability as a potential limitation and highlight the need for replications in future research. Another limitation was that our study data were collected as a part of the routine assessment at the clinic. The benefit of this, as mentioned, is the naturalistic setting in which DBT for BPD was examined; and the downside was the lack of control condition, which means that changes in outcome variables over time may not be entirely attributed to the treatment received. Furthermore, some potentially relevant variables were not examined due to practical constraints (e.g., the effects of virtual treatment was not studied because all clients seen virtually during COVID-19 had also received in-person sessions; the effects of minoritized racial identities were not examined due to the small subset), and some of the variables tested were not measured the same way as they were studied in prior literature (e.g., NSSI in the past year instead of a different period). Moreover, it is important to note that post-treatment data were not available for some intent-to-treat individuals who dropped out of treatment resulting in potentially biased results based on available data.

Conclusions

Taken together, higher homework completion rate, retention in treatment, and greater baseline severity appeared to be important predictors of DBT treatment response evaluated by BPD symptom severity and the utilization of DBT skills. Favorable effects of some aspects of therapeutic alliance during initial sessions were also found. Future research may benefit from consolidating the criteria of treatment response, identifying clinically relevant variables for evaluation, and testing generalizability of findings across settings, to enhance knowledge of insufficient treatment response in comprehensive DBT for BPD.

Notes

¹ Note that Wise (2004) specifies a continuum of treatment response categories ranging from recovery to deterioration. Only categories that included at least 1 participant in the current study are described here.

² Note that the baseline outcome rating was also included in model 2 resulting in a total of 11 predictors entered.

Acknowledgements

The current study was not pre-registered. The Rutgers University Institutional Review Board approved this study (Study ID: Pro2019001864). Materials of the study can be accessed by contacting the authors.

Disclosure Statement

Shireen L. Rizvi provides training and consultation for Behavioral Tech, LLC. Qingqing Yin, Molly Stern, and Evan M. Kleiman do not have competing interests.

ORCID

QINGQING YIN  <http://orcid.org/0000-0002-7731-1486>

EVAN M. KLEIMAN  <http://orcid.org/0000-0001-8002-1167>

SHIREEN L. RIZVI  <http://orcid.org/0000-0002-6714-0097>

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>
- Adrian, M., McCauley, E., Berk, M. S., Asarnow, J. R., Korslund, K., Avina, C., Gallop, R., & Linehan, M. M. (2019). Predictors

- and moderators of recurring self-harm in adolescents participating in a comparative treatment trial of psychological interventions. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 60(10), 1123–1132. <https://doi.org/10.1111/jcpp.13099>
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- Berk, M. S., Gallop, R., Asarnow, J. R., Adrian, M., Avina, C., Hughes, J. L., Korslund, K. E., & McCauley, E. (2022). Trajectories of treatment response and nonresponse in youth at high risk for suicide. *Journal of the American Academy of Child & Adolescent Psychiatry*, 61(9), 1119–1130. <https://doi.org/10.1016/j.jaac.2022.01.010>
- Bohus, M., Haaf, B., Simms, T., Limberger, M. F., Schmah, C., Unckel, C., Lieb, K., & Linehan, M. M. (2004). Effectiveness of inpatient dialectical behavioral therapy for borderline personality disorder: A controlled trial. *Behaviour Research and Therapy*, 42(5), 487–499. [https://doi.org/10.1016/S0005-7967\(03\)00174-8](https://doi.org/10.1016/S0005-7967(03)00174-8)
- Carson-Wong, A., Hughes, C. D., & Rizvi, S. L. (2018). The effect of therapist use of validation strategies on change in client emotion in individual DBT treatment sessions. *Personality Disorders: Theory, Research, and Treatment*, 9(2), 165. <https://doi.org/10.1037/per0000229>
- Chalker, S. A., Carmel, A., Atkins, D. C., Landes, S. J., Kerbrat, A. H., & Comtois, K. A. (2015). Examining challenging behaviors of clients with borderline personality disorder. *Behaviour Research and Therapy*, 75, 11–19. <https://doi.org/10.1016/j.brat.2015.10.003>
- Chen, E. Y., Matthews, L., Allen, C., Kuo, J. R., & Linehan, M. M. (2008). Dialectical behavior therapy for clients with binge-eating disorder or bulimia nervosa and borderline personality disorder. *The International Journal of Eating Disorders*, 41(6), 505–512. <https://doi.org/10.1002/eat.20522>
- Edwards, E. R., Kober, H., Rinne, G. R., Griffin, S. A., Axelrod, S., & Cooney, E. B. (2021). Skills-homework completion and phone coaching as predictors of therapeutic change and outcomes in completers of a DBT intensive outpatient programme. *Psychology and Psychotherapy: Theory, Research and Practice*, 94(3), 504–522. <https://doi.org/10.1111/papt.12325>
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured clinical interview for DSM-IV axis I personality disorders (SCID-II)*. American Psychiatric Press, Inc.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. (2002). *Structured clinical interview for DSM-IV-TR Axis I disorders, research version*. Biometrics Research.
- First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2015). *Structured clinical interview for DSM-5-research version*. American Psychiatric Association.
- 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>
- Gloster, A. T., Walder, N., Levin, M. E., Twohig, M. P., & Karekla, M. (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. *Journal of Contextual Behavioral Science*, 18, 181–192. <https://doi.org/10.1016/j.jcbs.2020.09.009>
- Goodman, M., Patil, U., Steffel, L., Avedon, J., Sasso, S., Triebwasser, J., & Stanley, B. (2010). Treatment utilization by gender in patients with borderline personality disorder. *Journal of Psychiatric Practice*, 16(3), 155–163. <https://doi.org/10.1097/01.pra.0000375711.47337.27>

- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 41–54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>
- Greenwell, B. M., Boehmke, B. C., & Gray, B. (2020). Variable importance plots—An introduction to the vip package. *The R Journal*, 12(1), 343. <https://doi.org/10.32614/RJ-2020-013>
- Harned, M. S., Jackson, S. C., Comtois, K. A., & Linehan, M. M. (2010). Dialectical behavior therapy as a precursor to PTSD treatment for suicidal and/or self-injuring women with borderline personality disorder. *Journal of Traumatic Stress*, 23(4), 421–429. <https://doi.org/10.1002/jts.20553>
- Harned, M. S., Korslund, K. E., & Linehan, 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, 7–17. <https://doi.org/10.1016/j.brat.2014.01.008>
- Hatcher, R. L., & Gillaspay, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, 16(1), 12–25. <https://doi.org/10.1080/10503300500352500>
- 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. Advance online publication. <https://doi.org/10.1016/j.brat.2020.103692>
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the working alliance inventory. *Journal of Counseling Psychology*, 36(2), 223–233. <https://doi.org/10.1037/0022-0167.36.2.223>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19. <https://doi.org/10.1037/0022-006X.59.1.12>
- Kliem, S., Kröger, C., & Kosfelder, J. (2010). Dialectical behavior therapy for borderline personality disorder: A meta-analysis using mixed-effects modeling. *Journal of Consulting and Clinical Psychology*, 78(6), 936–951. <https://doi.org/10.1037/a0021015>
- Kröger, C., Harbeck, S., Armbrust, M., & Kliem, S. (2013). Effectiveness, response, and dropout of dialectical behavior therapy for borderline personality disorder in an inpatient setting. *Behaviour Research and Therapy*, 51(8), 411–416. <https://doi.org/10.1016/j.brat.2013.04.008>
- Landes, S. J., Chalker, S. A., & Comtois, K. A. (2016). Predicting dropout in outpatient dialectical behavior therapy with patients with borderline personality disorder receiving psychiatric disability. *Borderline Personality Disorder and Emotion Dysregulation*, 3(1), 9. <https://doi.org/10.1186/s40479-016-0043-3>
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Publications.
- Linehan, M. M. (2014). *DBT skills training manual* (2nd ed.). Guilford Publications.
- Linehan, M. M., Comtois, K. A., Murray, A. M., Brown, M. Z., Gallop, R. J., Heard, H. L., Korslund, K. E., Tutek, D. A., Reynolds, S. K., & Lindenboim, N. (2006). Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Archives of General Psychiatry*, 63(7), 757–766. <https://doi.org/10.1001/archpsyc.63.7.757>
- Lobbstaël, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV axis I disorders (SCID I) and axis II disorders (SCID II). *Clinical Psychology & Psychotherapy*, 18(1), 75–79. <https://doi.org/10.1002/cpp.693>
- Lynch, T. R., Chapman, A. L., Rosenthal, M. Z., Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology*, 62(4), 459–480. <https://doi.org/10.1002/jclp.20243>
- McMain, S. F., Chapman, A. L., Kuo, J. R., Guimond, T., Streiner, D. L., Dixon-Gordon, K. L., Isaranuwatjai, W., & Hoch, J. S. (2018). The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: The feasibility of a shorter treatment and evaluating responses (FASTER) trial protocol. *BMC Psychiatry*, 18(1), 230. <https://doi.org/10.1186/s12888-018-1802-z>
- Mehlum, L. (2021). Mechanisms of change in dialectical behaviour therapy for people with borderline personality disorder. *Current Opinion in Psychology*, 37, 89–93. <https://doi.org/10.1016/j.copsyc.2020.08.017>
- Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. H. (2002). The work and social adjustment scale: A simple measure of impairment in functioning. *The British Journal of Psychiatry: The Journal of Mental Science*, 180(5), 461–464. <https://doi.org/10.1192/bjp.180.5.461>
- Neacsiu, A. D., Eberle, J. W., Kramer, R., Wiesmann, T., & Linehan, M. M. (2014). Dialectical behavior therapy skills for transdiagnostic emotion dysregulation: A pilot randomized controlled trial. *Behaviour Research and Therapy*, 59, 40–51. <https://doi.org/10.1016/j.brat.2014.05.005>
- Neacsiu, A. D., Rizvi, S. L., Vitaliano, P. P., Lynch, T. R., & Linehan, M. M. (2010). The dialectical behavior therapy ways of coping checklist: Development and psychometric properties. *Journal of Clinical Psychology*, 66(6), 563–582. <https://doi.org/10.1002/jclp.20685>
- Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-Injurious thoughts and behaviors interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment*, 19(3), 309–317. <https://doi.org/10.1037/1040-3590.19.3.309>
- Panos, P. T., Jackson, J. W., Hasan, O., & Panos, A. (2014). Meta-Analysis and systematic review assessing the efficacy of dialectical behavior therapy (DBT). *Research on Social Work Practice*, 24(2), 213–223. <https://doi.org/10.1177/1049731513503047>
- Perroud, N., Uher, R., Dieben, K., Nicastro, R., & Huguelet, P. (2010). Predictors of response and drop-out during intensive dialectical behavior therapy. *Journal of Personality Disorders*, 24(5), 634–650. <https://doi.org/10.1521/pedi.2010.24.5.634>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL <https://www.R-project.org/>.
- Riley, R. D., Ensor, J., Snell, K. I., Harrell, F. E., Martin, G. P., Reitsma, J. B., Moons, K. G. M., Collins, G., & Van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ*, 368, m441. <https://doi.org/10.1136/bmj.m441>
- Riley, R. D., Snell, K., 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. (2011). Treatment failure in dialectical behavior therapy. *Cognitive and Behavioral Practice*, 18(3), 403–412. <https://doi.org/10.1016/j.cbpra.2010.05.003>

- Rizvi, S. L., Hughes, C. D., Hittman, A. D., & Vieira Oliveira, P. (2017). Can trainees effectively deliver dialectical behavior therapy for individuals With borderline personality disorder? Outcomes from a training clinic. *Journal of Clinical Psychology*, 73(12), 1599–1611. <https://doi.org/10.1002/jclp.22467>
- Schmitgen, M. M., Niedtfeld, I., Schmitt, R., Mancke, F., Winter, D., Schmahl, C., & Herpertz, S. C. (2019). Individualized treatment response prediction of dialectical behavior therapy for borderline personality disorder using multimodal magnetic resonance imaging. *Brain and Behavior*, 9(9), e01384. <https://doi.org/10.1002/brb3.1384>
- Shearin, E. N., & Linehan, M. M. (1994). Dialectical behavior therapy for borderline personality disorder: Theoretical and empirical foundations. *Acta Psychiatrica Scandinavica*, 89, 61–68.
- Thomas, M. C. (2017). Working Alliance Factors and Their Relation to Emotion, Engagement and Dropout in Dialectical Behavior Therapy (Order No. 10757832). Available from Dissertations & Theses @ Rutgers University; ProQuest Dissertations & Theses Global. (1994113125). <https://login.proxy.libraries.rutgers.edu/login?url=https%3A%2F%2Fwww.proquest.com%2Fdissertations-theses%2Fworking-alliance-factors-their-relation-emotion%2Fdocview%2F1994113125%2Fse-2%3Faccountid%3D13626>
- Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, 82(1), 50–59. https://doi.org/10.1207/s15327752jpa8201_10
- Woodbridge, J., Townsend, M., Reis, S., Singh, S., & Grenyer, B. F. (2022). Non-response to psychotherapy for borderline personality disorder: A systematic review. *Australian & New Zealand Journal of Psychiatry*, 56(7), 771–787. <https://doi.org/10.1177/00048674211046893>
- World Health Organization. (2001). *International classification of functioning, disability and health: ICF*.