UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Towards optimal treatment for borderline personality disorder

Comparative effectiveness, treatment selection, and assessment

Wibbelink, C.J.M.

Publication date
2026

Link to publication

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

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

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

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

Download date:10 Nov 2026

Chapter 10

General discussion

This thesis aimed to contribute to the optimization of treatment outcomes for patients with BPD by (1) evaluating the comparative effectiveness of DBT and ST, (2) advancing personalized treatment selection, and (3) strengthening assessment in BPD. These aims were addressed through the BOOTS trial, a multicenter RCT comparing DBT and ST for BPD. This final chapter situates the main findings of the preceding chapters within the three overarching aims of the thesis. The chapter begins with a brief reconsideration of the design and rationale of the BOOTS trial, after which the findings are discussed in relation to each aim. Subsequently, the methodological strengths and limitations are considered, followed by the implications for clinical practice and future research.

DESIGN AND RATIONALE OF THE BOOTS STUDY

The BOOTS trial, described in **Chapter 2**, was designed to address an important gap in the literature by directly comparing DBT and ST for outpatients with BPD in a multicenter RCT, thereby enhancing our understanding of which treatment works best for whom and why. Both treatments were delivered over 2 years in a combined group–individual format, with assessments every 6 months up to the 1-year post-treatment follow-up. The comparative effectiveness was evaluated using both clinical and economic outcome measures. The primary clinical outcome was the rate of change in DSM-5 BPD severity, assessed with the Borderline Personality Disorder Severity Index, fifth version (BPDSI-5). Secondary clinical outcomes included DSM-5 BPD criteria, general symptom severity, functioning, quality of life, well-being, sleep problems, and treatment dropout. Economic outcomes included quality-adjusted life-years (QALYs; cost-utility outcome) and BPD severity (BPDSI-5; cost-effectiveness outcome), both evaluated in relation to societal costs over a 3-year time horizon. In addition, a range of potential predictors of differences in treatment effectiveness between DBT and ST were assessed to help identify factors associated with differential treatment response.

Efforts to improve treatment outcomes for BPD are particularly important given the substantial burden associated with the disorder. **Chapter 3** demonstrated that Dutch treatment-seeking outpatients with BPD experience markedly reduced quality of life and high societal costs. Quality of life was considerably lower than in the general Dutch population (Versteegh et al., 2016). Moreover, quality of life was lower and annual societal costs higher than in individuals with other mental or somatic disorders (Franklin & Hernández Alava, 2023; Gustavsson et al., 2011). These findings highlight the importance of effective treatment for BPD. Identifying and evaluating effective treatment options for this population therefore remains a key priority.

COMPARATIVE TREATMENT EFFECTIVENESS

In the BOOTS trial, the comparative clinical and economic effectiveness of DBT and ST was evaluated. **Chapter 4** revealed no significant differences between DBT and ST in clinical outcomes, including BPD severity and all secondary outcomes, with substantial improvements across outcomes in both treatments. Treatment retention also did not differ significantly, with dropout rates of 34% for DBT and 29% for ST after 1 year, and 52% for DBT and 46% for ST after 2 years. **Chapter 5** showed minimal, non-significant differences between treatments in total societal costs, QALYs, and BPD severity. However, the probability of cost-effectiveness differed depending on the outcome measure. From a cost-utility perspective, the probability that DBT was cost-effective increased with higher willingness-to-pay (WTP) thresholds, reaching 59% at a WTP threshold of €50,000 per QALY. In contrast, when effectiveness was defined in terms of BPD severity, ST showed a somewhat higher probability (60%) of being cost-effective across WTP thresholds. These findings suggest that neither treatment demonstrated a clear economic advantage over the other for outpatients with BPD.

Several aspects of these findings merit further discussion. First, the trial did not include a no-treatment or non-specialized control group. As a result, improvements observed within both treatment conditions cannot be definitively attributed to the specific effects of the treatments (Harrer et al., 2023). In the absence of a control condition, it cannot be ruled out that part of the observed improvements may reflect non-specific factors, such as therapeutic attention, the passage of time (maturation), or patients' expectations of improvement (Cuijpers et al., 2017). However, including a no-treatment control group was considered ethically problematic given the clinical severity of BPD, particularly the elevated risk of self-harm and suicidal behavior. In addition, a no-treatment or non-specialized control condition might have reduced the willingness to participate in the trial, thereby affecting recruitment and the representativeness of the study sample. Importantly, the within-group improvements observed in both treatments were comparable to those reported in previous studies of DBT and ST (e.g., Arntz et al., 2022; Assmann et al., 2024). Previous trials have also demonstrated that both DBT and ST are more effective than treatment as usual (e.g., Arntz et al., 2022; Farrell et al., 2009; Stoffers-Winterling et al., 2022). These findings suggest that the observed improvements likely reflect, at least in part, meaningful treatment effects, although the design of the BOOTS trial does not allow specific treatment effects to be disentangled from non-specific influences.

Beyond questions of treatment attribution, the pattern of improvements across outcomes also warrants closer consideration. Specifically, improvements in participation in work or study were notably smaller (small-to-medium) than improvements observed across other outcomes. This pattern is consistent with previous DBT and ST trials in BPD showing that improvements in functioning tend to be smaller than reductions in BPD manifestations and psychopathological symptoms in general (Arntz et al., 2022; Harned et al., 2018; McMain

et al., 2012). Similarly, meta-analytic findings indicate that gains in functioning following psychotherapy for BPD often fall below thresholds for clinical relevance (Storebø et al., 2020). Long-term prospective studies examining the course of BPD further suggest that a substantial proportion of individuals with BPD continue to experience difficulties in functioning, even after symptomatic improvement (Gunderson et al., 2011; Zanarini et al., 2018). Persistent functional impairments are not only detrimental to individuals with BPD (Larivière et al., 2015; Zeitler et al., 2020), but also place a substantial burden on society, with a large proportion of BPD-related societal costs attributable to occupational impairment (van Asselt et al., 2007; Wagner et al., 2014, 2022). These findings underscore the importance of specifically targeting functional outcomes within DBT and ST.

A third aspect to consider is treatment dropout. Nearly half of the participants discontinued treatment during the 2-year treatment period, with approximately one-third discontinuing treatment within the first year. Although the 1-year dropout rate was lower than the rate reported in a meta-analysis on dropout from psychological treatment for BPD (40%; Arntz et al., 2023), it was higher than the dropout rate observed in the other DBT-ST trial (Assmann et al., 2024). Dropout rates in DBT were broadly comparable to those reported in previous meta-analyses (Arntz et al., 2023; Dixon & Linardon, 2020; Kliem et al., 2010; Swift & Greenberg, 2014), whereas dropout in ST was markedly higher than in an earlier study examining a comparable treatment format (Arntz et al., 2022). In addition, dropout varied considerably across participating centers (33–75%).

Several contextual factors may help explain the observed dropout rates. The implementation and execution of the treatments in some participating centers faced several challenges, including high therapist turnover and organizational instability, factors that have previously been associated with dropout (Bales, Timman, et al., 2017; De Salve et al., 2025; Steuwe et al., 2023; Werbart et al., 2014). Supervisors also expressed concerns regarding the extent to which supervisory instructions and the treatment protocol were followed. Furthermore, the trial was conducted during the COVID-19 pandemic, which led to treatment interruptions or a temporary transition to telehealth for some patients. Telehealth has been associated with reduced group cohesion and, consequently, higher dropout rates in DBT (Lopez et al., 2020; Zalewski et al., 2021). Moreover, illness related to COVID-19 and fear of the pandemic may have also posed barriers to treatment retention (Gao et al., 2023; Hellstern & Robinson, 2023). Importantly, however, there were no indications that COVID-19-related treatment interruptions affected treatment effectiveness in terms of BPD severity. Nevertheless, broader impacts of the COVID-19 pandemic, including possible effects on treatment dropout and, consequently, on treatment effectiveness, cannot be ruled out.

Fourth, no clear advantage emerged for either DBT or ST across a broad range of outcome measures. Our findings align with those of the previous DBT-ST trial (Assmann et al., 2024), which highlighted the potential role of shared therapeutic factors, such as a structured

treatment framework, a focus on emotional processes, and a consistent explanatory model. At the same time, DBT and ST differ substantially in their theoretical rationales and therapeutic techniques. For example, DBT emphasizes the acquisition of skills and primarily focuses on the present, whereas ST places greater emphasis on early development and experiential techniques (Fassbinder et al., 2018). However, distinct treatment techniques may lead to comparable outcomes through different pathways or shared therapeutic mechanisms. These findings may be considered in relation to the so-called "Dodo Bird Verdict" (Rosenzweig, 1936), which suggests comparable effectiveness across bona fide psychotherapies. However, these findings do not provide evidence of equivalence between DBT and ST, as the trial was not designed or powered to test non-inferiority or equivalence. Moreover, other studies have reported indications of differential effects between treatments, although these were based on indirect comparisons rather than head-to-head trials (e.g., Arntz et al., 2023; Setkowski et al., 2023).

Finally, although no straightforward economic preference for DBT or ST emerged due to the substantial uncertainty and non-significant differences in health outcomes and societal costs, the divergence between BPD severity and QALYs in the probability of cost-effectiveness warrants attention. A similar pattern was reported in a previous economic evaluation of individual ST for BPD (van Asselt et al., 2008). The authors hypothesized that ST may be relatively burdensome early in treatment, resulting in smaller early gains in quality of life that, although no longer evident at follow-up, may translate into slightly lower QALYs. A comparable finding was observed in the present study. This may not have affected cost-effectiveness based on BPD severity, as the economic evaluation relied on the follow-up assessment of BPD severity, whereas QALYs capture changes in quality of life across the entire study period.

PERSONALIZED TREATMENT SELECTION

In the absence of a clear preference between DBT and ST, optimizing treatment selection may represent an important avenue for enhancing treatment effectiveness. Although both treatments were associated with substantial improvements, individuals with BPD represent a highly heterogeneous population in terms of clinical presentation (Bohus et al., 2021; Leichsenring et al., 2024), suggesting that patients may benefit from different treatment approaches. A central aim of this thesis was therefore to examine whether specific patient characteristics could inform personalized treatment selection to optimize treatment outcomes. Chapters 6 and 7 addressed this aim using two complementary approaches.

Chapter 6 examined whether patient characteristics predicted differential treatment effectiveness of DBT versus ST. No evidence was found for prescriptive predictors of differential early treatment dropout, although two characteristics (somatic symptom

disorder and paranoid personality disorder) were associated with dropout risk irrespective of treatment. For treatment response, however, the findings suggested heterogeneity in treatment effects. While preliminary and open to alternative explanations—including the possibility that the findings reflect prognostic rather than truly prescriptive effects—the results indicated that a subgroup of patients may derive greater benefit from DBT relative to ST. Identity disturbance, clinician-rated alliance capacity, and self-reflection emerged as potential indicators of differential treatment benefit. **Chapter 7** investigated treatment outcomes among clinically relevant subgroups: patients with and without comorbid SUD. Both DBT and ST were associated with improvements in clinical and psychosocial outcomes in patients with and without comorbid SUD, with generally no evidence that treatment type moderated these effects. However, results regarding substance use outcomes were mixed, with only limited improvements in substance use observed across treatments and subgroups, emphasizing the need to address substance use more explicitly within BPD treatments.

Taken together, the findings provide only modest support for the use of patient characteristics to guide treatment selection between DBT and ST. Even though some indicators of differential treatment benefit were identified (**Chapter 6**), the evidence for robust prescriptive predictors remained preliminary. In addition, the subgroup analysis in **Chapter 7** did not indicate that comorbid SUD moderated the relative effectiveness of DBT and ST. These findings are consistent with previous research suggesting that identifying reliable predictors of differential treatment response in psychotherapy may prove difficult (Chekroud et al., 2021, 2024; Eilertsen & Eilertsen, 2023). One explanation is that treatment outcomes in BPD are influenced by a complex interplay of patient characteristics, treatment processes, and contextual factors, which may be difficult to capture adequately using baseline predictors alone. In addition, therapist-related factors may also substantially influence treatment outcomes. For example, Lutz et al. (2007) found that therapist effects account for substantial variability in psychotherapy outcomes. As a result, models based solely on patient-level baseline characteristics may capture only a limited proportion of the variability in treatment effectiveness (Kaiser & Herzog, 2023).

A further consideration concerns the practical utility of ML-based prediction models for guiding treatment allocation, as applied in **Chapter 6**. Although such models can identify complex multivariate patterns that may go undetected using traditional statistical methods, their clinical applicability remains uncertain (Chekroud et al., 2021). Many ML models operate as “black boxes” (Rudin, 2019), meaning that the internal decision-making processes of the algorithm are difficult to interpret. This lack of transparency complicates the interpretation of predictions and makes it difficult to determine whether identified predictors reflect accurate clinical mechanisms or spurious associations (Rudin, 2019), thereby creating a barrier to implementation in clinical practice (Chekroud et al., 2021; Hameed et al., 2025). The difficulty of interpreting such models is illustrated by the pattern observed in the evaluation of the model’s clinical utility in **Chapter 6**. Patients classified as ST-non-optimal

showed a numerically greater decrease in BPD severity than those classified as ST-optimal, but this difference was not statistically significant. One possible explanation is that the model primarily captures a prognostic signal instead of a purely prescriptive effect, meaning that it identifies patients with better outcomes regardless of treatment rather than those who benefit differentially from DBT versus ST. At the same time, the relatively favorable outcomes observed in the DBT-optimal subgroup suggest that the model may capture at least some treatment-specific information. These findings underscore the difficulty of interpreting complex prediction models and highlight the challenges of translating such models into clear clinical decision rules.

Despite these challenges in interpreting ML-based prediction models, tentative explanations may be proposed for the potential associations between the identified predictors and treatment effectiveness. For example, somatic symptom disorder and paranoid personality disorder may increase the risk of treatment dropout by hindering the development of a therapeutic alliance (Berghuis et al., 2021; Breivik Øvstebø et al., 2024; Patel & Suhr, 2020). Additional speculative explanations may help clarify why DBT might confer greater relative benefit than ST among patients with lower identity disturbance, higher clinician-rated alliance capacity, and lower—and possibly very high—levels of self-reflection. Interestingly, these predictors differed from the prescriptive factors that clinicians most frequently mentioned during semi-structured interviews as indicative of differential treatment response, such as substance abuse and self-destructive behavior. First, DBT may be particularly beneficial for patients with lower identity disturbance, as it targets identity-related difficulties only indirectly (Bonavitacola & Buerger, 2022). Explicit identity-focused interventions may therefore be less helpful for these patients. Second, DBT strategies and procedures rely heavily on a positive patient–therapist relationship, underscoring the importance of early alliance formation (Linehan, 1993). Patients with greater baseline alliance capacity may therefore respond especially well to DBT, consistent with findings showing that alliance capacity is associated with positive outcomes in DBT but not in the comparison treatment (Bedics et al., 2015). Finally, DBT's structured, skills-based approach may be particularly helpful for individuals with limited self-reflective capacity, but also for those with very high levels of self-reflection, indicative of maladaptive rumination or excessive self-focus (Grant et al., 2002). These interpretations, however, rely on the assumption that the model captures treatment-specific effects. If this assumption does not hold, the identified predictors may reflect general treatment responsiveness rather than differential benefit from DBT versus ST.

Finally, evidence from other studies suggests that prediction models developed in a single dataset frequently show reduced performance when applied to independent samples (Chekroud et al., 2021), raising concerns about the external validity of the present prediction models. Although external validation is considered the gold standard for evaluating the performance of prediction models, it remains rare in psychotherapy research. A review of ML approaches in psychotherapy research found that only 6% of studies had conducted external

validation (Aafjes-van Doorn et al., 2021). Studies that did test the predictive accuracy and generalizability of multivariate prediction models in independent datasets often found that the results did not generalize to external validation samples (Chekroud et al., 2024; Tait et al., 2025; van Bronswijk et al., 2021). Identifying predictors that consistently generalize across samples may be particularly challenging in psychotherapy research. As emphasized by Eilertsen and Eilertsen (2023), human lives and psychological processes are highly complex, and the number of potential predictors is extremely large. Moreover, predictors often interact with each other, and psychotherapy is characterized by complex and unique interactions between therapist and patient. This complexity may make it difficult to identify predictors that consistently generalize across samples and settings. However, some studies using large naturalistic samples that included external validation found evidence that prescriptive effects were retained (e.g., Delgadillo & Salas Duhne, 2020; Schwartz et al., 2021). Nevertheless, the ultimate test of the clinical utility of a treatment selection procedure is a prospective study evaluating whether model-informed treatment recommendations improve outcomes. Although such trials are scarce, Delgadillo et al. (2022) demonstrated that model-informed allocation to either high- or low-intensity psychotherapy was associated with higher remission rates compared with care as usual.

Overall, multivariate ML prediction models may offer promising tools for exploring treatment heterogeneity and informing treatment selection procedures, even though their utility for guiding treatment selection between DBT and ST appeared limited in the present context. At the same time, these findings highlight that improving treatment outcomes may not only depend on optimizing treatment selection, but also on ensuring high-quality treatment delivery. Consistent with broader psychotherapy research emphasizing the role of therapist and contextual factors in treatment outcomes (Bales, Timman, et al., 2017; Johns et al., 2019; Knox & Hill, 2021; Werbart et al., 2014), investments in therapist training, supervision, and supportive organizational contexts may therefore represent important avenues for improving treatment effectiveness.

ASSESSMENT

The third aim of this thesis was to advance the assessment of BPD by evaluating instruments that measure BPD symptoms and treatment-related processes. Accurate and efficient assessment is essential for both research and clinical practice, as it enables reliable monitoring of symptom severity, treatment progress, and mechanisms of change. Chapters 8 and 9 therefore focused on the psychometric evaluation of two instruments relevant to BPD assessment.

Chapter 8 described the psychometric evaluation of the uBPDc, a brief measure designed for efficient screening and outcome monitoring in BPD. The uBPDc demonstrated satisfactory

psychometric properties in individuals with BPD, individuals with Cluster C personality disorder, and non-clinical controls, including good reliability and validity, as well as satisfactory diagnostic accuracy. Moreover, the uBPDc proved to be sensitive to change, and support was found for a unidimensional construct with partial measurement invariance. **Chapter 9** examined the Dutch version of the DBT-WCCL, an instrument designed to assess DBT skills use and dysfunctional coping. The hypothesized factor model was not supported, although separate factor models yielded the expected dimensional structure. Partial measurement invariance was established only for the subscale measuring DBT skills use. Reliability and known-group validity were satisfactory, while results for concurrent validity were inconclusive. Finally, the DBT-WCCL also proved sensitive to change.

These findings contribute to the further development of assessment tools for BPD by providing evidence for the psychometric properties of instruments assessing symptom severity and treatment-related processes. Nevertheless, several observations regarding the adaptation of instruments, comparability across groups, and potential refinement of these instruments warrant further reflection. A first issue concerns differences in psychometric properties observed between the evaluated instruments and their original versions. Differences in psychometric performance may partly be related to variations in the length of instruments, as internal consistency generally increases when the number of items increases (Tavakol & Dennick, 2011). This is consistent with the slightly lower, though still satisfactory, internal consistency found for the uBPDc, which is a shortened version of the original questionnaire. In addition, the translation of instruments may introduce challenges, including subtle shifts in meaning, particularly when items contain idiomatic expressions (Herdman et al., 1997). For example, item 59 of the DBT-WCCL ("Tried not to burn my bridges behind me, but leave things open somewhat") required substantial adaptation in the Dutch version to reflect the intended meaning ("Geprobeerd om het probleem op te lossen zonder vriendschappen/ relaties kwijt te raken"). Such adaptations may influence how respondents interpret items and may partly contribute to differences in factor structure or validity across language versions (van de Vijver & Leung, 2021). The hypothesized factor structure was not supported for the Dutch version of the DBT-WCCL, although methodological differences may also partly explain these discrepancies.

A second issue concerns the extent to which the evaluated instruments function equivalently across groups, as indicated by the findings on measurement invariance. Although partial measurement invariance was supported for the uBPDc and the DBT skills use subscale of the DBT-WCCL, several parameters had to be freed to achieve acceptable model fit. This suggests that some items functioned differently across groups, indicating that individuals with BPD and non-clinical controls may interpret or endorse certain items differently. For the uBPDc, differences were observed in factor loadings, suggesting that some items contributed differently to the latent construct across groups. Such differences may arise when certain items are less applicable to a particular group (Chen, 2008). For example, the substantially

lower factor loading observed for item 5 (non-suicidal self-injury and suicidality) in the non-clinical control group likely reflects the lower prevalence of these behaviors among non-clinical individuals (Pompili et al., 2005). When an item is rarely endorsed, correlations with the latent construct may be attenuated, resulting in lower factor loadings (Chen, 2008). This highlights a broader challenge in measurement invariance research when comparing clinical and non-clinical samples: some phenomena occur frequently in clinical populations but rarely in non-clinical groups, making it difficult to determine whether observed non-invariance reflects genuine differences in item functioning or statistical artifacts resulting from restricted variability. For the DBT skills use subscale, differences were observed in item thresholds, indicating variation in the likelihood of endorsing specific items despite similar levels of the latent construct. This may relate to different reference points or response tendencies across groups (Chen, 2008). By allowing the non-invariant parameters to vary across groups, partial measurement invariance was established. Many researchers consider partial measurement invariance to be acceptable when only a small proportion of parameters is non-invariant (Bowen & Masa, 2015). Nevertheless, the use of partial measurement invariance has also been criticized (Robitzsch & Lüdtke, 2023), and clear guidelines regarding how many parameters can be freed without introducing substantial bias remain lacking (Dimitrov, 2010).

Finally, the results suggest that the DBT-WCCL could benefit from further refinement. First, some items may be ambiguous in their interpretation, as certain behaviors can reflect either adaptive skill use or dysfunctional behavior. For example, ensuring that one responds in a way that does not alienate others (item 6) could reflect effective interpersonal functioning, but may also represent appeasing behavior driven by a fear of rejection (Burmeister et al., 2017). Second, the current distribution of items across the four DBT skills domains may be suboptimal, as half of the items represent distress tolerance skills, which could influence sensitivity to change over the course of treatment. In addition, several items required correlated residuals in the model, indicating conceptual overlap and potentially reflecting redundancy and an overrepresentation of specific DBT subskills. Future revisions of the instrument may therefore benefit from eliminating redundant items (Arteaga-de-Luna et al., 2023). A further consideration concerns the instrument's focus on the frequency of DBT skills use, whereas emerging evidence suggests that the quality, effectiveness, and timing of skills are likely to be more important (Southward et al., 2023).

METHODOLOGICAL STRENGTHS AND LIMITATIONS

Strengths

Several methodological strengths of the BOOTS trial are worth noting. First, treatment outcomes were assessed using a comprehensive multimethod approach, combining self-report

and observer-rated instruments, and adopting a broad perspective on treatment response by including outcomes beyond symptom reduction, such as psychosocial functioning and well-being. Effectiveness studies in BPD have often focused on symptom improvement and diagnostic remission (Ng, Townsend, et al., 2019; Renneberg et al., 2024), whereas research has increasingly recognized the importance of other aspects of recovery (Larivière et al., 2015; Ng, Carter, et al., 2019; Ng, Townsend, et al., 2019; Renneberg et al., 2024). For example, qualitative research indicates that individuals with BPD consider improvements in psychosocial functioning highly important (Ng, Carter, et al., 2019). The relatively smaller improvements observed in functioning outcomes in the clinical effectiveness analyses further underscore the importance of assessing multiple outcome domains. By including measures that captured a broad spectrum of recovery, the BOOTS trial provided a more comprehensive evaluation of treatment outcomes.

In addition, the identification of potential predictors was based on a multimethod strategy that integrated evidence from the literature, input from a patient representative, and semi-structured interviews with expert clinicians. This strategy is consistent with Meehl's observation that actuarial prediction models are unlikely to outperform clinical judgment when they are based on incomplete knowledge of relevant variables (Meehl, 1954). At the same time, clinical expertise can play an important role in identifying meaningful predictors (Westen & Weinberger, 2004). While statistical models can help estimate the relative contribution of predictors based on the data, clinical expertise can ensure that potentially relevant variables are considered during model development. Integrating empirical evidence with clinical expertise and patient perspectives therefore increased the likelihood that relevant and clinically meaningful predictors were identified and included, thereby supporting the development of a well-performing prediction model.

Second, the trial was conducted by a research group with balanced allegiance with respect to the treatments under investigation (DBT and ST) and included an independent primary researcher. This reduced the likelihood that the findings were influenced by researcher preferences or expectations regarding a particular treatment approach (i.e., allegiance bias; Leykin & DeRubeis, 2009). Third, the multicenter design of the trial, combined with routine clinical settings and minimal exclusion criteria, enhances the ecological validity of the study and improves the generalizability of the findings to real-world clinical practice.

Limitations

Although limitations of the individual studies have been addressed in the respective chapters, two overarching limitations warrant further discussion: sample size and missing data. A primary limitation of the BOOTS trial concerns the relatively modest sample size ($N = 204$). While the intended sample size of 200 participants was achieved and the trial was relatively large compared to several previous BPD studies (e.g., Assmann et al., 2024; McMain et al., 2009; Walton et al., 2020), the study was not powered to detect equivalence or small differences

between the two treatments. The trial was powered to detect medium-sized treatment effects. Because both DBT and ST are evidence-based treatments, potential differences between the treatments may be small and could therefore have remained undetected (**Chapter 4**). This limitation also applies to the economic evaluation (**Chapter 5**) and the study examining treatment outcomes among patients with and without comorbid SUD (**Chapter 7**). Even though more than one-third of the participants had a comorbid SUD and the current study included a larger sample than most previous studies, the statistical power remained limited to reliably detect subtle interaction effects.

Moreover, the sample size also constitutes an important limitation for the prediction study (**Chapter 6**). Developing multivariable prediction models requires sufficiently large samples to reduce the risk of overfitting and unstable estimates (Rajput et al., 2023). Although repeated nested cross-validation was applied to minimize potential bias (Rajput et al., 2023; Varma & Simon, 2006), the absence of an independent hold-out sample precludes formal external validation of the prediction model. According to Luedtke et al. (2019), a sample size of at least 300 participants per treatment condition is generally recommended for the development of multivariable prediction models. Nevertheless, smaller samples may still be informative when the aim is to generate prediction models that can be tested in future research. Taken together, the findings of the prediction study should be considered hypothesis-generating and require replication in future studies. More broadly, the findings of the BOOTS trial contribute to the growing body of evidence on treatments for BPD and may support future efforts to synthesize data across trials. Approaches such as individual participant data meta-analyses may help overcome the sample size limitations of individual trials and allow more robust examination of treatment effects and predictors of treatment effectiveness (Cuijpers et al., 2022; Storebø et al., 2021).

The second limitation concerns the substantial proportion of missing data at the repeated reassessments. Different approaches were used to handle missing data in the clinical and economic evaluations. In the clinical evaluation and subgroup analysis (**Chapters 4 and 7**), mixed regression models were applied without imputing missing values, in line with common practice and recommendations in longitudinal data analysis (e.g., Fitzmaurice et al., 2018; Molenberghs et al., 2004; Twisk et al., 2013). In the economic evaluation (**Chapter 5**), multiple imputation was applied, which is widely recommended for handling missing data in economic evaluations (Ben et al., 2023; Burton et al., 2007; El Alili et al., 2022). These two approaches yielded slightly different estimates for the BPDSI-5 at follow-up, with a small numerical advantage for DBT in the clinical evaluation and for ST in the economic evaluation. However, these differences were minor and did not affect the overall conclusions, as no statistically significant differences between treatments were observed.

Importantly, both mixed regression models and multiple imputation assume that missing data are missing completely at random (MCAR) or missing at random (MAR) (Cummings, 2013; Twisk et al., 2013). Data are considered MCAR when missingness is unrelated to

observed or unobserved variables, MAR when missingness depends on observed variables, and missing not at random (MNAR) when missingness depends on unobserved factors, including the missing value itself (Twisk et al., 2013). In practice, it is often difficult to determine whether missing data follow a MAR or MNAR mechanism. When data are MNAR, both mixed regression analyses and multiple imputation may yield biased estimates because they rely on the observed data to estimate missing values (Twisk et al., 2013). To address this uncertainty, sensitivity analyses were conducted to examine the robustness of the findings under different assumptions about missingness. For the clinical evaluation, the impact of treatment dropout—highly correlated with missing data—on the primary outcome (BPD severity) was examined in a sensitivity analysis. The results were consistent with the main analysis, with similar improvements over time and no indication that dropout affected the estimated treatment difference between DBT and ST. For the economic evaluation, an MNAR sensitivity analysis was conducted in which missing outcomes (societal costs and health outcomes) were assumed to be worse (worst-case scenario), which likewise produced similar results. Although missing data may still have influenced the estimates to some extent, these analyses suggest that the main findings remained largely consistent under different assumptions about missingness.

Finally, beyond the potential bias related to uncertainty about the missingness mechanism, missing data also complicated the analysis of predictors of differential treatment effectiveness (**Chapter 6**). Early treatment dropout resulted in missing BPDSI-5 reassessments, which necessitated a two-step procedure in which treatment response was examined only among participants who remained in treatment for at least 6 months. Consequently, the findings of this study apply to a selected subgroup and may not fully generalize to the entire sample.

IMPLICATIONS FOR RESEARCH AND PRACTICE

Clinical implications

The findings of this thesis have important implications for clinical practice. First, improving the availability of specialized treatments for BPD is essential. Access to treatments remains limited, resulting in a substantial proportion of individuals with BPD not receiving psychotherapy (Choi-Kain et al., 2016; Hermens et al., 2011; Iliakis et al., 2019; Tusiani-Eng & Yeomans, 2018). Limited availability is largely attributable to structural constraints in healthcare resources, including limited time, staff, and budget (Drummond et al., 2015; Hermens et al., 2011; Tusiani-Eng & Yeomans, 2018). Furthermore, the implementation and long-term sustainability of specialized treatments such as DBT and ST in routine clinical practice can be challenging due to high staff turnover, the need for extensive therapist training, and the intensity and duration of these treatments (Bales, Verheul, et al., 2017; Iliakis et al., 2019; King et al., 2018). These challenges are also reflected in the Dutch mental

healthcare system, where waiting times for specialized mental healthcare remain substantial: in 2025, more than half of the approximately 100,000 individuals on waiting lists exceeded the recommended maximum waiting time of 14 weeks, and individuals requiring treatment for personality disorders waited on average 32 weeks before treatment could start. Prolonged waiting times are concerning, as delays in treatment are associated with worsening symptoms, increased treatment complexity, and a negative impact on relatives (Schutjens et al., 2024). The burden-of-disease study presented in **Chapter 3** further highlights the substantial individual and societal burden associated with BPD, underscoring the need to expand access to specialized, evidence-based treatments for this patient group. Expanding treatment availability, however, also requires evidence regarding which specialized treatments can be effectively delivered in routine clinical practice.

Second, the findings of the BOOTS trial indicate that both DBT and ST can be considered appropriate treatment options for outpatients with BPD. Both treatments were associated with substantial improvements across a broad range of outcomes, suggesting that DBT and ST can lead to meaningful clinical change when implemented in routine clinical settings. Moreover, the analyses did not reveal significant differences in clinical outcomes (**Chapter 4**) or economic outcomes (**Chapter 5**), indicating that both treatments represent viable options for the treatment of BPD. Nevertheless, the findings highlight several areas for improvement, including participation in work or study, substance use-related outcomes, and treatment retention. Addressing these domains in clinical practice may further strengthen the effectiveness of specialized treatments for BPD.

Third, in the absence of clear differences between treatments, treatment selection in clinical practice may be guided by patient preference, therapist expertise, and the availability of treatments in routine clinical practice. At the same time, the findings of the prediction study (**Chapter 6**) suggest that treatment selection strategies may hold promise for further improving outcomes. In particular, the results indicate that a subgroup of patients may derive greater benefit from DBT in terms of reductions in BPD severity, with three specific predictors emerging as potentially relevant indicators. However, the current findings do not yet support the use of prediction models to guide treatment allocation in clinical practice. The absence of robust ST-specific predictors, the possibility that the model reflects prognostic improvement potential rather than treatment-specific effects, and the lack of external validation indicate that further evidence is needed before model-informed treatment selection can be implemented in routine care. Moreover, it bears noting that the differential benefit observed for BPD severity did not generalize to secondary outcomes. This suggests that model-informed treatment allocation based on symptom severity may primarily influence symptom trajectories rather than broader aspects of recovery. Given the importance of functional outcomes and quality of life for individuals with BPD (Larivière et al., 2015; Ng, Carter, et al., 2019), the clinical applicability of the prediction model identified in **Chapter 6** for guiding individualized treatment decisions remains uncertain, although the model may inform future efforts toward more personalized treatment selection.

Fourth, in addition to treatment-related implications, the findings of this thesis also have implications for the assessment of BPD symptoms and treatment-related processes in clinical practice. The results presented in **Chapter 8** suggest that the uBPDc may serve as a useful screening and monitoring instrument for BPD. Due to constraints in time and clinical expertise required to administer diagnostic interviews, several authors have recommended a two-step diagnostic procedure in which a screening instrument is followed by a semi-structured diagnostic interview for individuals scoring above a clinical threshold (Siefert, 2010; Widiger et al., 2024). The uBPDc appears well suited for this purpose. Furthermore, its brief format facilitates repeated assessments and routine outcome monitoring in both clinical practice and research. The findings presented in **Chapter 9** highlight the potential clinical utility of the DBT-WCCL for assessing DBT-specific processes. Skills use represents a central mechanism of change within DBT (Neacsiu, Rizvi, Vitaliano, et al., 2010), and the DBT-WCCL enables clinicians and researchers to systematically assess the extent to which patients apply these skills in daily life. The instrument can provide insight into how patients manage difficult situations and help identify dysfunctional coping patterns that may serve as targets for treatment (Neacsiu, Rizvi, Vitaliano, et al., 2010). As such, the DBT-WCCL may be a valuable tool for monitoring treatment processes and evaluating treatment progress in individuals undergoing DBT. Taken together, the uBPDc and DBT-WCCL enable the systematic assessment of symptom change and treatment-related processes, thereby supporting both monitoring in clinical practice and research aimed at elucidating mechanisms of change in the treatment of BPD. Understanding these mechanisms represents an important future direction of the BOOTS trial and will be discussed in the next paragraph.

Future research

A key direction for future research concerns the investigation of mechanisms of change in the treatment of BPD. In addition to identifying predictors of differential treatment effectiveness, another important avenue for improving treatment outcomes is to examine the processes through which treatments bring about change (Huibers et al., 2021; Kazdin & Nock, 2003; Kraemer et al., 2002), which also represents one of the central aims of the BOOTS trial. The BOOTS trial offers a unique opportunity to investigate potential mechanisms of change in both DBT and ST. The rich longitudinal dataset allows the examination of candidate mediators reflecting treatment-specific processes (i.e., skills use and emotion regulation in DBT; schema modes and maladaptive beliefs in ST), as well as more generic and non-specific therapeutic processes (i.e., therapeutic environment, therapeutic alliance, and attachment). Examining these processes may help clarify the extent to which treatment effects are driven by treatment-specific versus more general mechanisms of change. A study aimed at investigating these mechanisms of change is currently in preparation.

A second priority for future research is the replication and external validation of the prediction model identified in **Chapter 6**. Replication and external validation in independent

samples constitute essential next steps for evaluating the robustness and generalizability of the model (Chekroud et al., 2021). External validation is considered a crucial phase in the development of clinical prediction models, following model development and preceding prospective studies that examine whether model-informed treatment allocation improves outcomes (Moons et al., 2009). Prospective studies are required to determine whether using prediction models for treatment selection leads to meaningful improvements in treatment outcomes. These steps are necessary before prediction models can be reliably used to guide treatment decisions in routine clinical practice. Beyond the model developed in this thesis, future research may also benefit from developing prediction models based on widely used omnibus assessment instruments, such as the Minnesota Multiphasic Personality Inventory-Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008), as their routine use in clinical practice may enhance the feasibility of prospective testing and implementation in naturalistic clinical settings (Wibbelink et al., 2025). In the longer term, this line of research on prediction-guided treatment selection may help ensure that individuals with BPD receive the treatment from which they are most likely to benefit.

A final direction for future research relates to the high dropout rates observed in the BOOTS trial. High rates of premature treatment termination are common in treatments for BPD (Arntz et al., 2023) and tend to be higher than those observed in psychotherapies for other mental disorders, such as major depressive disorder, generalized anxiety disorder, and posttraumatic stress disorder (Iliakis et al., 2021). Treatment dropout is associated with poorer clinical outcomes, including increased hospitalization rates (McMurran et al., 2010). In addition, high dropout rates may have negative consequences for other patients (e.g., disruption of therapy groups), therapists (e.g., demoralization), mental healthcare organizations (e.g., longer waiting lists), and society more broadly (e.g., higher costs) (de Freixo Ferreira et al., 2023; McMurran et al., 2010; Webb & McMurran, 2009). Future research should therefore focus on identifying factors associated with treatment dropout and developing strategies to improve treatment retention. As dropout often occurs early in treatment (Arntz et al., 2023; Cooper et al., 2023), particular attention should be given to factors that influence treatment engagement during the initial phases of treatment. Most studies on dropout have focused primarily on patient characteristics, whereas the role of therapist-related and organizational factors remains relatively understudied (Werbart et al., 2014). However, emerging evidence suggests that such factors may substantially influence treatment outcomes. For example, Bales et al. (2017) found that major organizational changes were associated with a substantial reduction in the effectiveness of Mentalization-Based Treatment for BPD, underscoring the potential impact of organizational factors on treatment outcomes. Future studies should therefore examine not only patient-level predictors of dropout, but also therapist-, team-, and organizational-level factors that may influence treatment retention in BPD. Enhancing treatment retention is likely an important step toward maximizing the effectiveness of specialized treatments for individuals with BPD.

CONCLUSION

Despite extensive efforts to detect potential differences between DBT and ST, no clear advantage emerged for either treatment. These findings support the use of both DBT and ST as viable treatment options in clinical practice, while underscoring the need to further improve treatment outcomes for individuals with BPD. The BOOTS trial also provides important insights into heterogeneity in treatment response, including preliminary indications of the potential utility of model-informed treatment selection. At the same time, these findings suggest that improving outcomes may depend not only on optimizing treatment selection but also on addressing other factors contributing to variability in treatment outcomes. In addition, the results highlight the potential value of psychometrically sound assessment tools for monitoring symptoms and treatment processes. Taken together, the BOOTS trial provides a strong foundation for future research and innovation in clinical practice aimed at optimizing treatment outcomes for patients 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

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. *JAMA Psychiatry*, *79*(4), 287–299. https://doi.org/10.1001/jamapsychiatry.2022.0010

Arntz, A., Mensink, K., Cox, W. R., Verhoef, R. E. J., van Emmerik, A. A. P., Rameckers, S. A., Badenbach, T., & Grasman, R. P. P. P. (2023). Dropout from psychological treatment for borderline personality disorder: A multilevel survival meta-analysis. *Psychological Medicine*, *53*(3), 668–686. https://doi.org/10.1017/S0033291722003634

Arteaga-de-Luna, S. M., Rodríguez, O. F., Hermosillo-de-la-Torre, A. E., Garduño, M. D. L. V., & González-Betanzos, F. (2023). Adaptación, validación y propiedades psicométricas de la DBT-WCCL-DSS en población mexicana [Adaptation, validation, and psychometric properties of the DBT-WCCL-DSS in the Mexican population]. *Revista Iberoamericana de Diagnóstico y Evaluación–e Avaliação Psicológica, 70*, 175–185.

Assmann, N., Schaich, A., Arntz, A., Wagner, T., Herzog, P., Alvarez-Fischer, D., Sipos, V., Jauch-Chara, K., Klein, J. P., Hüppe, M., Schweiger, U., & Fassbinder, E. (2024). The effectiveness of dialectical behavior therapy compared to schema therapy for borderline personality disorder: A randomized clinical trial. *Psychotherapy and Psychosomatics*, *93*(4), 249–263. https://doi.org/10.1159/000538404

Bales, D. L., Timman, R., Luyten, P., Busschbach, J., Verheul, R., & Hutsebaut, J. (2017). Implementation of evidence-based treatments for borderline personality disorder: The impact of organizational changes on treatment outcome of mentalization-based treatment. *Personality and Mental Health*, *11*(4), 266–277. https://doi.org/10.1002/pmh.1381

Bales, D. L., Verheul, R., & Hutsebaut, J. (2017). Barriers and facilitators to the implementation of mentalization-based treatment (MBT) for borderline personality disorder. *Personality and Mental Health*, *11*(2), 118–131. https://doi.org/10.1002/pmh.1368

Bedics, J. D., Atkins, D. C., Harned, M. S., & Linehan, M. M. (2015). The therapeutic alliance as a predictor of outcome in dialectical behavior therapy versus nonbehavioral psychotherapy by experts for borderline personality disorder. *Psychotherapy*, *52*(1), 67–77. https://doi.org/10.1037/a0038457

Ben, Â. J., van Dongen, J. M., El Alili, M., Heymans, M. W., Twisk, J. W. R., MacNeil-Vroomen, J. L., de Wit, M., van Dijk, S. E. M., Oosterhuis, T., & Bosmans, J. E. (2023). The handling of missing data in trial-based economic evaluations: should data be multiply imputed prior to longitudinal linear mixed-model analyses? *European Journal of Health Economics*, *24*(6), 951–965. https://doi.org/10.1007/s10198-022-01525-y

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

Berghuis, H., Bandell, C. C., & Krueger, R. F. (2021). Predicting dropout using DSM–5 Section II personality disorders, and DSM–5 Section III personality traits, in a (day)clinical sample of personality disorders. *Personality Disorders: Theory, Research, and Treatment*, *12*(4), 331–338. https://doi.org/10.1037/per0000429

Bohus, M., Stoffers-Winterling, J., Sharp, C., Krause-Utz, A., Schmahl, C., & Lieb, K. (2021). Borderline personality disorder. *The Lancet*, *398*(10310), 1528–1540. https://doi.org/10.1016/S0140-6736(21)00476-1

Bonavitacola, L., & Buerger, W. (2022). Identity Disturbance: Clinical implications for DBT treatment of self-dysfunction in BPD. *DBT Bulletin*, *6*(2), 20–25. https://static1.squarespace.com/static/62e933212ea2a15419c196cd/t/63763102725422641fc7b7c8/1668690179847/DBTBulletin_V7.pdf

Bowen, N. K., & Masa, R. D. (2015). Conducting measurement invariance tests with ordinal data: A guide for social work researchers. *Journal of the Society for Social Work and Research*, *6*(2), 229–249. https://doi.org/10.1086/681607

Breivik Øvstebø, R., Pedersen, G., Wilberg, T., Røssberg, J. I., Johnsen Dahl, H.-S., & Kvarstein, E. H. (2024). Countertransference, alliance, and outcome in the treatment of patients with personality disorder: A longitudinal naturalistic study. *Frontiers in Psychiatry*, *15,* 1490056. https://doi.org/10.3389/fpsyt.2024.1490056

Burmeister, K., Höschel, K., Barenbrügge, J., Chrysanthou, C., Rist, F., Neacsiu, A. D., & Pedersen, A. (2017). Die Dialectical Behavior Therapy Ways of Coping Checklist (DBT-WCCL). *Diagnostica*, *63*(1), 29–41. https://doi.org/10.1026/0012-1924/a000164

Burton, A., Billingham, L. J., & Bryan, S. (2007). Cost-effectiveness in clinical trials: Using multiple imputation to deal with incomplete cost data. *Clinical Trials*, *4*(2), 154–161. https://doi.org/10.1177/1740774507076914

Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, *20*(2), 154–170. https://doi.org/10.1002/wps.20882

Chekroud, A. M., Hawrilenko, M., Loho, H., Bondar, J., Gueorguieva, R., Hasan, A., Kambeitz, J., Corlett, P. R., Koutsouleris, N., Krumholz, H. M., Krystal, J. H., & Paulus, M. (2024). Illusory generalizability of clinical prediction models. *Science*, *383*(6679), 164–167. https://doi.org/10.1126/science.adg8538

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005–1018. https://doi.org/10.1037/a0013193

Choi-Kain, L. W., Albert, E. B., & Gunderson, J. G. (2016). Evidence-based treatments for borderline personality disorder: Implementation, integration, and stepped care. *Harvard Review of Psychiatry*, *24*(5), 342–356. https://doi.org/10.1097/HRP.0000000000000113

Cooper, A. A., Kline, A. C., Baier, A. L., & Feeny, N. C. (2023). Rethinking research on prediction and prevention of psychotherapy dropout: A mechanism-oriented approach. *Behavior Modification*, *47*(6), 1195–1218. https://doi.org/10.1177/01454455187922 51

Cuijpers, P., Ciharova, M., Quero, S., Miguel, C., Driessen, E., Harrer, M., Purgato, M., Ebert, D., & Karyotaki, E. (2022). The Contribution of "Individual Participant Data" Meta-Analyses of Psychotherapies for Depression to the Development of Personalized Treatments: A Systematic Review. *Journal of Personalized Medicine*, *12*(1), 93. https://doi.org/10.3390/jpm12010093

Cuijpers, P., Weitz, E., Cristea, I. A., & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences*, *26*(4), 364–368. https://doi.org/10.1017/S2045796016000809

Cummings, P. (2013). Missing Data and Multiple Imputation. *JAMA Pediatrics*, *167*(7), 656–661. https://doi.org/10.1001/jamapediatrics.2013.1329

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 and Psychotherapy*, *30*(6), 1324–1337. https://doi.org/10.1002/cpp.2888

De Salve, F., Rossi, C., Gioacchini, E., Messina, I., & Oasi, O. (2025). Dropout in psychotherapy for personality disorders: A systematic review of predictors. *Clinical Psychology & Psychotherapy*, *32*(3), e70080. https://doi.org/10.1002/cpp.70080

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

Delgadillo, J., & Salas Duhne, P. G. (2020). Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. *Journal of Consulting and Clinical Psychology*, *88*(1), 14–24. https://doi.org/10.1037/ccp0000476

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(2), 121–149. https://doi.org/10.1177/0748175610373459

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

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

Eilertsen, S. E. H., & Eilertsen, T. H. (2023). Why is it so hard to identify (consistent) predictors of treatment outcome in psychotherapy? – Clinical and research perspectives. *BMC Psychology*, *11*(1), 198. https://doi.org/10.1186/s40359-023-01238-8

El Alili, M., van Dongen, J. M., Esser, J. L., Heymans, M. W., van Tulder, M. W., & Bosmans, J. E. (2022). A scoping review of statistical methods for trial-based economic evaluations: The current state of play. *Health Economics*, *31*(12), 2680–2699. https://doi.org/10.1002/hec.4603

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

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

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2018). *Applied Longitudinal Analysis* (2nd ed.). John Wiley & Sons Inc.

Franklin, M., & Hernández Alava, M. (2023). Enabling QALY estimation in mental health trials and care settings: mapping from the PHQ-9 and GAD-7 to the ReQoL-UI or EQ-5D-5L using mixture models. *Quality of Life Research*, *32*(10), 2763–2778. https://doi.org/10.1007/s11136-023-03443-9

Gao, L., Keller, F. M., Becker, P., Dahmen, A., & Lippke, S. (2023). Predictors of dropout among psychosomatic rehabilitation patients during the COVID-19 pandemic: Secondary analysis of a longitudinal study of digital training. *Journal of Medical Internet Research*, *25*(1), e43584. https://doi.org/10.2196/43584

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

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

Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., Dodel, R., Ekman, M., Faravelli, C., Fratiglioni, L., Gannon, B., Jones, D. H., Jennum, P., Jordanova, A., Jönsson, L., Karampampa, K., Knapp, M., Kobelt, G., Kurth, T., ... CDBE2010 study group. (2011). Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology, 21*(10), 718–779. https://doi.org/10.1016/j.euroneuro.2011.08.008

Hameed, S., Nauman, M., Akhtar, N., Fayyaz, M. A. B., & Nawaz, R. (2025). Explainable AI-driven depression detection from social media using natural language processing and black box machine learning models. *Frontiers in Artificial Intelligence*, *8,* 1627078. https://doi.org/10.3389/frai.2025.1627078

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

Harrer, M., Cuijpers, P., Schuurmans, L. K. J., Kaiser, T., Buntrock, C., van Straten, A., & Ebert, D. (2023). Evaluation of randomized controlled trials: A primer and tutorial for mental health researchers. *Trials*, *24*(1), 562. https://doi.org/10.1186/s13063-023-07596-3

Hellstern, R. B., & Robinson, W. D. (2023). The impact of COVID-19 and telehealth services on attrition rates in psychotherapy. *Contemporary Family Therapy*, *45*(3). https://doi.org/10.1007/s10591-023-09661-0

Herdman, M., Fox-Rushby, J., & Badia, X. (1997). "Equivalence" and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research*, *6*(3), 237–247. https://doi.org/10.1023/a:1026410721664

Hermens, M. L. M., van Splunteren, P. T., van Den Bosch, A., & Verheul, R. (2011). Barriers to implementing the clinical guideline on borderline personality disorder in the Netherlands. *Psychiatric Services*, *62*(11), 1381–1383. https://doi.org/10.1176/ps.62.11.pss6211_1381

Huibers, M. J. H., Lorenzo-Luaces, L., Cuijpers, P., & Kazantzis, N. (2021). On the road to personalized psychotherapy: A research agenda based on cognitive behavior therapy for depression. *Frontiers in Psychiatry*, *11,* 607508. https://doi.org/10.3389/fpsyt.2020.607508

Iliakis, E. A., Ilagan, G. S., & Choi-Kain, L. W. (2021). Dropout rates from psychotherapy trials for borderline personality disorder: A meta-analysis. *Personality Disorders: Theory, Research, and Treatment*, *12*(3), 193–206. https://doi.org/10.1037/per0000453

Iliakis, E. A., Sonley, A. K. I., Ilagan, G. S., & Choi-Kain, L. W. (2019). Treatment of borderline personality disorder: Is supply adequate to meet public health needs? *Psychiatric Services*, *70*(9), 772–781. https://doi.org/10.1176/appi.ps.201900073

Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to Baldwin and Imel's (2013) review. *Clinical Psychology Review*, *67*, 78–93. https://doi.org/10.1016/j.cpr.2018.08.004

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

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

King, J. C., Hibbs, R., Saville, C. W. N., & Swales, M. A. (2018). The survivability of dialectical behaviour therapy programmes: A mixed methods analysis of barriers and facilitators to implementation within UK healthcare settings. *BMC Psychiatry*, *18*(1), 302. https://doi.org/10.1186/s12888-018-1876-7

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

Knox, S., & Hill, C. E. (2021). Training and supervision in psychotherapy: What we know and where we need to go. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (7th ed., pp. 327–349). John Wiley & Sons, Inc.

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

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

Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in psychotherapy outcome research: Separating association from bias. *Clinical Psychology: Science and Practice, 16*(1), 54–65. https://doi.org/10.1111/j.1468-2850.2009.01143.x

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

Lopez, A., Rothberg, B., Reaser, E., Schwenk, S., & Griffin, R. (2020). Therapeutic groups via video teleconferencing and the impact on group cohesion. *MHealth*, *6,* 13. https://doi.org/10.21037/mhealth.2019.11.04

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

Lutz, W., Martinovich, Z., Lyons, J. S., Leon, S. C., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, *54*(1), 32–39. https://doi.org/10.1037/0022-0167.54.1.32

McMain, S. F., Guimond, T., Streiner, D. L., Cardish, R. J., & Links, P. S. (2012). Dialectical behavior therapy compared with general psychiatric management for borderline personality disorder: Clinical outcomes and functioning over a 2-year follow-up. *American Journal of Psychiatry*, *169*(6), 650–661. https://doi.org/10.1176/appi.ajp.2012.11091416

McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *American Journal of Psychiatry*, *166*(12), 1365–1374. https://doi.org/10.1176/appi.ajp.2009.09010039

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

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* University of Minnesota Press. https://doi.org/10.1037/11281-000

Molenberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M. G., Mallinckrodt, C., & Carroll, R. J. (2004). Analyzing incomplete longitudinal clinical trial data. *Biostatistics*, *5*(3), 445–464. https://doi.org/10.1093/biostatistics/kxh001

Moons, K. G. M., Royston, P., Vergouwe, Y., Grobbee, D. E., & Altman, D. G. (2009). Prognosis and prognostic research: What, why, and how? *BMJ, 338,* b375. https://doi.org/10.1136/bmj.b375

Neacsiu, A. D., Rizvi, S. L., & Linehan, M. M. (2010). Dialectical behavior therapy skills use as a mediator and outcome of treatment for borderline personality disorder. *Behaviour Research and Therapy, 48*(9), 832–839. https://doi.org/10.1016/j.brat.2010.05.017

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

Ng, F. Y. Y., Townsend, M. L., Miller, C. E., Jewell, M., & Grenyer, B. F. S. (2019). The lived experience of recovery in borderline personality disorder: A qualitative study. *Borderline Personality Disorder and Emotion Dysregulation*, *6*(1), 10. https://doi.org/10.1186/s40479-019-0107-2

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

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

Rajput, D., Wang, W. J., & Chen, C. C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*, *24*(1), 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. (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), 14. https://doi.org/10.1186/s40479-024-00257-0

Robitzsch, A., & Lüdtke, O. (2023). Why full, partial, or approximate measurement invariance are not a prerequisite for meaningful and valid group comparisons. *Structural Equation Modeling*, *30*(6), 859–870. https://doi.org/10.1080/10705511.2023.2191292

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

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Schutjens, V., Nuijen, J., & Magnée, T. (2024). *In de wachtstand: Een literatuurinventarisatie naar de impact van wachttijden in de ggz op patiënten, naasten, huisartsen en POH's-GGZ.* Trimbos-Instituut.

Schwartz, B., Cohen, Z. D., Rubel, J. A., Zimmermann, D., Wittmann, W. W., & Lutz, W. (2021). Personalized treatment selection in routine care: Integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. *Psychotherapy Research*, *31*(1), 33–51. https://doi.org/10.1080/10503307.2020.1769219

Setkowski, K., Palantza, C., van Ballegooijen, W., Gilissen, R., Oud, M., Cristea, I. A., Noma, H., Furukawa, T. A., Arntz, A., van Balkom, A. J. L. M., & Cuijpers, P. (2023). Which psychotherapy is most effective and acceptable in the treatment of adults with a (sub)clinical borderline personality disorder? A systematic review and network meta-analysis. *Psychological Medicine*, *53*(8), 3261–3280. https://doi.org/10.1017/S0033291723000685

Siefert, C. J. (2010). Screening for personality disorders in psychiatric settings: Four recently developed screening measures. In L. Baer & M. A. Blais (Eds.), *Handbook of clinical rating scales and assessment in psychiatry and mental health* (pp. 125–144). Humana Press/Springer Nature.

Southward, M. W., Howard, K. P., & Cheavens, J. S. (2023). Less is more: Decreasing the frequency of maladaptive coping predicts improvements in DBT more consistently than increasing the frequency of adaptive coping. *Behaviour Research and Therapy*, *163,* 104288. https://doi.org/10.1016/j.brat.2023.104288

Steuwe, C., Berg, M., Driessen, M., & Beblo, T. (2023). Therapeutic alliance and dropout in patients with borderline pathology receiving residential dialectical behavior therapy. *BMC Psychiatry*, *23*(1), 605. https://doi.org/10.1186/s12888-023-05061-8

Stoffers-Winterling, J. M., Storebø, O. J., Kongerslev, M. T., Faltinsen, E., Todorovac, A., Sedoc Jørgensen, M., Sales, C. P., Edemann Callesen, H., Pereira Ribeiro, J., Völlm, B. A., Lieb, K., & Simonsen, E. (2022). Psychotherapies for borderline personality disorder: A focused systematic review and meta-analysis. *The British Journal of Psychiatry*, *221*(3), 538–552. https://doi.org/10.1192/bjp.2021.204

Storebø, O. J., Ribeiro, J. P., Kongerslev, M. T., Stoffers-Winterling, J., Sedoc Jørgensen, M., Lieb, K., Bateman, A., Kirubakaran, R., Dérian, N., Karyotaki, E., Cuijpers, P., & Simonsen, E. (2021). Individual participant data systematic reviews with meta-analyses of psychotherapies for borderline personality disorder. *BMJ Open, 11*(6), e047416. https://doi.org/10.1136/bmjopen-2020-047416

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

Swift, J. K., & Greenberg, R. P. (2014). A treatment by disorder meta-analysis of dropout from psychotherapy. *Journal of Psychotherapy Integration*, *24*(3), 193–207. https://doi.org/10.1037/a0037512

Tait, J., Kellett, S., Saxon, D., Deisenhofer, A. K., Lutz, W., Barkham, M., & Delgadillo, J. (2025). Individual treatment selection for patients with post-traumatic stress disorder: External validation of a personalised advantage index. *Psychotherapy Research*, *35*(5), 838–851. https://doi.org/10.1080/10503307.2024.2360449

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2,* 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tusiani-Eng, P., & Yeomans, F. (2018). Borderline personality disorder: Barriers to borderline personality disorder treatment and opportunities for advocacy. *Psychiatric Clinics of North America, 41(4), 695–709*. https://doi.org/10.1016/j.psc.2018.07.006

Twisk, J., De Boer, M., De Vente, W., & Heymans, M. (2013). Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. *Journal of Clinical Epidemiology*, *66*(9), 1022–1028. https://doi.org/10.1016/j.jclinepi.2013.03.017

van Asselt, A. D. I., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., van Dyck, R., Spinhoven, P., van Tilburg, W., Kremers, I. P., Nadort, M., & Severens, J. L. (2008). Out-patient psychotherapy for borderline personality disorder: Cost-effectiveness of schema-focused therapy v. transference-focused psychotherapy. *British Journal of Psychiatry*, *192*(6), 450–457. https://doi.org/10.1192/bjp.bp.106.033597

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

van Bronswijk, S. C., Bruijniks, S. J. E., Lorenzo-Luaces, L., DeRubeis, R. J., Lemmens, L. H. J. M., Peeters, F. P. M. L., & Huibers, M. J. H. (2021). Cross-trial prediction in psychotherapy: External validation of the Personalized Advantage Index using machine learning in two Dutch randomized trials comparing CBT versus IPT for depression. *Psychotherapy Research*, *31*(1), 78–91. https://doi.org/10.1080/10503307.2020.1823029

van de Vijver, F. J. R., & Leung, K. (2021). *Methods and Data Analysis for Cross-Cultural Research* (2nd ed.). Cambridge University Press.

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

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

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

Walton, C. J., Bendit, N., Baker, A. L., Carter, G. L., & Lewin, T. J. (2020). A randomised trial of dialectical behaviour therapy and the conversational model for the treatment of borderline personality disorder with recent suicidal and/or non-suicidal self-injury: An effectiveness study in an Australian public mental health service. *Australian & New Zealand Journal of Psychiatry*, *54*(10), 1020–1034. https://doi.org/10.1177/0004867420931164

Webb, D., & McMurran, M. (2009). A comparison of women who continue and discontinue treatment for borderline personality disorder. *Personality and Mental Health*, *3*(2), 142–149. https://doi.org/10.1002/pmh.69

Werbart, A., Andersson, H., & Sandell, R. (2014). Dropout revisited: Patient- and therapist-initiated discontinuation of psychotherapy as a function of organizational instability. *Psychotherapy Research*, *24*(6), 724–737. https://doi.org/10.1080/10503307.2014.883087

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

Wibbelink, C. J. M., Sellbom, M., Grasman, R. P. P. P., Arntz, A., Sinnaeve, R., & Kamphuis, J. H. (2025). Using the Minnesota Multiphasic Personality Inventory–2 restructured form to predict functioning after treatment for borderline personality disorder: A machine learning approach. *Psychological Assessment, 37*(6–7), 261–272. https://doi.org/10.1037/pas0001385

Widiger, T. A., Hines, A., & Crego, C. (2024). Evidence-based assessment of personality disorder. *Assessment*, *31*(1), 191–198. https://doi.org/10.1177/10731911231176461

Zalewski, M., Walton, C. J., Rizvi, S. L., White, A. W., Gamache Martin, C., O'Brien, J. R., & Dimeff, L. (2021). Lessons learned conducting dialectical behavior therapy via telehealth in the age of COVID-19. *Cognitive and Behavioral Practice*, *28*(4), 573–587. https://doi.org/10.1016/j.cbpra.2021.02.005

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