Supplementary Material

Table of Contents

**Tables**

S1. Missing data at each assessment timepoint for all study measures

S2. Percentage of treatment completed prior to dropout

S3. Medications and non-study psychosocial interventions

S4. Pearson R correlation matrix of CPT-3 scores across timepoints

**Sections**

1. Analyses to support the assumption that data were missing at random
2. Psychometric properties of clinical and cognitive measures
3. Multilevel model construction
4. Generalized estimating equation construction
5. References

**Table S1. Missing Data at Each Assessment Time Point for all Study Measure**

|  |  |  |  |
| --- | --- | --- | --- |
|  | CPT-3  (number of missing data points) | BIS-11  (number of missing data points) | SASII  (number of missing data points) |
| pre-treatment | 20 | 0 | 1 |
| 3 mo | -- | 29 | 31 |
| 6 mo | 64 | 44 | 46 |
| 9 mo | -- | 56 | 52 |
| 12mo | 86 | 58 | 57 |
| 15 mo | -- | 74 | 67 |
| 18 mo | -- | 80 | 78 |
| 21 mo | -- | 86 | 82 |
| 24 mo | -- | 86 | 91 |

1. **Analyses To Support The Assumption That Data Was Missing At Random.**
2. **No differences in CPT-3 performance between drop outs and treatment completers**

Prior to multilevel modeling, negative binomial regression was used to support the assumption of no significant differences in CPT-3 performance between participants who dropped out of the study, and those who did not. There were no significant differences in CPT-3 performance between participants who dropped out of the study and those who did not, including performance at pre-treatment: commission errors T score (β = 0.004, SE = 0.03, z = 0.12, p = 0.90); d prime T score (β = -0.003, SE = 0.03, z=-0.12, p = 0.91); HRTSD T score (β = 0.02, SE = 0.03, z = -0.64, p = 0.52); HRT-ISI T score (β = 0.02, SE = 0.03, z = -057, p = 0.57); impulsivity raw score (β = 0.003, SE = 0.98, z = 0.003, p = 0.99), performance at the six-month time point: commission errors T score (β = 0.01, SE = 0.04, z = 0.24, p = 0.81); d prime T score (β = 0.03, SE = 0.05, z=0.57, p = 0.57); HRTSD T score (β = -0.03, SE = 0.04, z = -0.79, p = 0.43); HRT-ISI T score (β = -0.03, SE = 0.04, z = -0.79, p = 0.43); impulsivity raw score (β = 0.11, SE = 1.44, z = 0.08, p = 0.94), or performance at the 12-month time point: commission errors T score (β = -0.02, SE = 0.05, z =- 0.43, p = 0.67); d prime T score (β = -0.004, SE = 0.05, z=-0.08, p = 0.94); HRTSD T score (β = 0.02, SE = 0.05, z =0.50, p = 0.62); HRT-ISI T score (β = -0.02, SE = 0.04, z = -0.44, p = 0.66); impulsivity raw score (β = 0.03, SE = 1.61, z = 0.02, p = 0.98).

1. **No significant time by dropout interaction effects for any CPT-3 variable across assessment timepoints**

Further, multilevel modeling was used to confirm that there were no significant time by dropout interaction effects for any CPT-3 variable at either the six- or twelve-month time points: commission errors (six-months: β = 0.43, SE = 1.67, t[313] = 0.25, p = 0.80, twelve-months: β = -0.34, SE = 1.81, t[313]= -0.19, p = 0.85); impulsivity index (six-months: β = 0.06, SE = 0.05, t[313] = 1.24, p = 0.23, twelve-months: β = -0.03, SE = 0.06, t[313] = -0.52, p = 0.60); d prime (six-months: β = 2.14, SE = 1.75, t[313] = 1.22, p = 0.22, twelve-months: β = .81, SE = 1.89, t[313] = 0.43, p = 0.67); HRT-SD (six-months: β = -1.30, SE = 1.60, t[313] = -0.81, p = 0.42, twelve-months: β = 1.87, SE = 1.72, t[313] = 1.09, p = 0.28); HRT-ISI (six-months: β = -1.45, SE = 1.54, t[313] = -0.94, p = 0.35, twelve-months: β = -0.28, SE = 1.88, t[313] = -0.15, p = 0.88).

1. **No significant between-arm differences in missing CPT-3 data at any assessment timepoint**

Finally, T-tests revealed no significant differences in missing CPT-3 data between participants in the six and 12-month treatment arms, at the pre-treatment [t(221.27)= 1.40, p= 0.16], the six month [t(237.62) = 0.58, p = 0.56], or the 12-month assessment time points [t(237.87) = -0.54, p = 0.59].

1. **No significant difference in pre-treatment CPT-3 scores between participants with and without missing self-injury data**

Prior to GEE analyses, T-tests were conducted to support the assumption of no differences in pretreatment neurocognition between participants with and without missing self-injury data. Participants with and without missing self-injury data points did not significantly differ in neurocognition at pre-treatment: impulsivity index: t(198.16)= -0.33 p =0.74), commission errors T score: t(214.6) = -1.002, p = 0.32, d prime T score: t(214.12) = -0.86, p = 0.39, BIS-11: t(237.94) = -1.54, p = 0.12.

1. **No significant differences in missing self-injury data between participants with average v. impaired CPT-3 performance and between participants with pre-treatment impulsivity in the low v. high range.**

Negative binomial regression showed no significant differences in the number of missing self-injury data points between participants with average or elevated CPT-3 T scores at pre-treatment: commission errors T score (β = 0.24, SE = 0.27, z = 0.92, p = 0.36); d prime T score (β = 0.22, SE = 0.28, z=0.81, p = 0.42), or between participants with low and high pre-treatment impulsivity: CPT-3 impulsivity index (β = -0.18, SE = 0.23, z = -0.78, p = 0.44); BIS-11 self-reported impulsivity (β = 0.37, SE = 0.22, z=1.69, p = 0.09).

**Table S2. Percentage of Treatment Completed Prior to Dropout**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Percentage of Treatment Completed Prior to Dropout** | | | | | | | | | | | |
|  | **10% or less** | | **11 to 25%** | | **26% to 50%** | | **51 to 75%** | | **76 to 90%** | | **> 90%** | |
|  | **DBT-12** | **DBT-6** | **DBT-12** | **DBT-6** | **DBT-12** | **DBT-6** | **DBT-12** | **DBT-6** | **DBT-12** | **DBT-6** | **DBT-12** | **DBT-6** |
| n | 6 | 2 | 12 | 2 | 8 | 11 | 8 | 9 | 4 | 3 | 4 | 3 |
| % sample | 5.0 | 1.7 | 10.0 | 1.7 | 6.7 | 9.2 | 6.7 | 7.5 | 3.3 | 2.5 | 3.3 | 2.5 |
| % dropouts | 14.3 | 6.7 | 28.6 | 6.7 | 19.0 | 36.7 | 19.0 | 30.0 | 9.5 | 10.0 | 9.5 | 10.0 |

Percentage of treatment completed prior to dropout was determined by dividing the number of days of treatment completed prior to dropout divided by the total possible days in treatment (DBT-12 = 365, DBT-6=183) Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration).

**Table S3. Estimated Marginal Means of Psychotropic Medications and Non-Study Psychosocial Treatments**

|  |  |  |
| --- | --- | --- |
| Estimated Rate per 3 Months | Psychotropic Medications | Psychosocial Treatments |
| DBT-12 (n = 120)  M [95% CI] | Baseline: 3.23 [2.72, 3.84]  Month 24: 2.19 [1.77, 2.73] | Baseline: 1.41 [1.16, 1.68]  Month 24: 1.02 [0.74, 1.35] |
| DBT-6 (n = 120),  M [95% CI] | Baseline: 2.94 [2.43, 3.57]  Month 24: 2.19 [1.74, 2.78] | Baseline: 1.23 [1.00, 1.44]  Month 24: 1.01 [0.74, 1.28] |

Psychotropic medications and non-study psychosocial treatments were assessed with the Treatment History Interview – 2 (Linehan, 1987). 95% Confidence Intervals (*95% CI*) are bias-corrected bootstrapped, based on 1000 bootstrap samples. Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration). Medications and non-study psychosocial treatments decreased significantly from baseline to 24-month follow up, with no significant differences between arms (McMain et al., 2022).

1. **Psychometric Properties of Clinical and Cognitive Measures**

*The Suicide Attempt Self-Injury Interview*

The SASII has excellent inter-rater reliability (median intraclass correlation = 0.96, range = 0.87 – 0.98) and good validity (83% agreement between SASII and therapist notes on the presence or absence of intentional self-injury and 76% agreement on total overall episodes; Linehan et al., 2006).

*The Barratt Impulsiveness Scale*

The total score has good psychometric properties (Stanford et al., 2009): internal consistency (Cronbach’s alpha = 0.83), test-retest reliability (Spearman’s rho = 0.83), convergent validity with other measures of trait impulsivity, and divergent validity with performance-based measures of impulsivity (Pearson *r* range: -0.02 – 0.06).

*The Conners Continous Performance Test – 3rd Edition (CPT-3)*

CPT-3 outcome variables examined in the current study are associated with adequate to excellent test-retest reliability (intraclass correlation range: 0.66 – 0.85) and split-half reliability (r = 0.92 in a normative sample and r = 0.92 – 0.95 in clinical samples; Conners, 2014).

1. **Multilevel Model Construction**

Prior to multilevel modeling, negative binomial regression was used to support the assumption of no significant differences in CPT-3 performance between participants who dropped out of the study, and those who did not. Further, multilevel modeling was used to confirm that there were no significant time by dropout interaction effects. Finally, T-tests revealed no significant differences in missing CPT-3 data between participants in the six and 12-month treatment arms.

The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to compare and determine the best fitting multilevel model for each CPT-3 variable. A random intercept model was chosen as the best fit for commission errors T-score, d*′* T-score, HRT-SD T-score, and self-reported BIS scores. A random intercept model with a first-order autoregressive correlation (AR1) structure was selected to model the Impulsivity Index. A random slopes model was used to fit the HRT-ISI T-score.

A Bonferroni statistical correction was applied to the results by dividing α = 0.05 by 15 to account for three comparisons within each of the five multilevel models (i.e., effect of time at six-months, effect of time at 12-months, and the effect of treatment length), resulting in a statistical significance threshold of *p* < 0.003. Post-hoc pairwise comparisons were examined for all participants collapsed across treatment arms. A family-wise statistical correction was applied to each post-hoc test to account for three pairwise comparisons across each assessment time point (i.e., pretreatment to six-months; pretreatment to 12-months; and six months to 12-months). Accordingly, for post hoc tests, α = 0.05 was divided by three, resulting in a significance threshold of *p* < 0.017.

Potential effects of comorbid diagnoses on cognitive performance were queried by testing for main and interaction effects among the variables of time, condition, and current comorbid diagnoses (including current major depressive disorder, any current substance use disorder, any current anxiety disorder, and current posttraumatic stress disorder). There were no interaction effects between comorbid diagnoses and the effects of time or condition (*p* > 0.05) and no three-way interaction effects among comorbid diagnoses, time and condition (*p* > 0.05). There were also no main effects of comorbid diagnoses (*p* > 0.05), suggesting that comorbidities were not associated with cognitive performance. As such, comorbidities were not modeled in the final analyses.

1. **Generalized Estimation Equation Construction**

Prior to Generalized Estimation Equation (GEE) analyses, T-tests were conducted to support the assumption of no differences in pretreatment cognitive performance between participants with and without missing self-injury data (see section 1. *Analyses To Support The Assumption That Data Was Missing At Random* above). Additionally, negative binomial regression was used to support the assumption of no differences in missing self-injury data between i) participants with pretreatment CPT-3 performance in the average versus elevated range and ii) participants with pretreatment performance-based or self-reported impulsivity in the low versus high range (see section 1. *Analyses To Support The Assumption That Data Was Missing At Random* above).

GEE is a semi-parametric approach for the analysis of correlated data over multiple time points, and an extension of the general linear model (Burton, Gurrin, and Sly, 1998). Unlike the general linear model, GEE does not make assumptions about the data distribution and permits various specifications of the distribution (Hubbard et al., 2010). This was advantageous in the current analysis because of a non-normal distribution and over-dispersion of the self-harm count data. GEEs in the present analysis were modeled using a Poisson distribution. Independent, exchangeable, and first-order autoregressive (AR1) correlation structures for each GEE model were compared using the quasi-likelihood independence criterion (QIC). An independence correlation structure fit best for pretreatment commission errors and d*′* T-scores and for the Impulsivity Index. An exchangeable correlation structure fit the BIS-11 scores best.

Plots of Cooks Distance for each model were visually inspected to identify outliers with the potential to substantially influence the results (Mendenhall, William, Sincich, and Terry, 1996). Analyses conducted after removing these outliers showed no changes in the significance of the results. As such, no outliers were removed.

**Table S4. Pearson R correlation matrix of CPT-3 scores across timepoints.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Base  d’  T Score | Base Comm  T Score | Base  HRT-SD  T Score | Base  HRT-ISI  T Score | Base  Impulsivity Index  Raw Score | Base  BIS-11  Raw Score | 6-MO  d’  T Score | 6-MO  Comm  T Score | 6-MO  HRT-SD  T Score | 6-MO  HRT-ISI  T Score | 6-MO  Impulsivity  Index  Raw Score | 6-MO  BIS-11  Raw Score | 12-MO  d’  T Score | 12-MO  Comm  T Score | 12-MO  HRT-SD  T Score | 12-MO  HRT-ISI  T Score | 12-MO  Impulsivity  Index  Raw Score | 12-MO  BIS-11  Raw  Score |
| Base  d’  T Score | -- | **r = 0.88**  **p < 0.01** | **r = 0.88**  **p < 0.01** | r = 0.02  p = 0.74 | **r = -0.51**  **p < 0.01** | **r = 0.14**  **p = 0.03** | **r = 0.65**  **p < 0.01** | **r = 0.64**  **p < 0.01** | **r = 0.24**  **p < 0.01** | r = 0.00  p =0.99 | **r = 0.46**  **p <0.01** | r = 0.15  p = 0.05 | **r = 0.58**  **p < 0.01** | **r = 0.61**  **p < 0.01** | r = 0.14  p = 0.11 | r = -0.01  p = 0.91 | **r = 0.48**  **p < 0.01** | r = 0.03  p = 0.68 |
| Base Comm  T Score | -- | -- | **r = 0.14**  **p =0.03** | r = -0.08  p= 0.24 | **r = 0.78**  **p < 0.01** | **r = 0.19**  **p < 0.01** | **r = 0.66**  **p <0.01** | **r = 0.73**  **p < 0.01** | r = 0.15  p = 0.06 | r = -0.04  p = 0.61 | **r = 0.59**  **p < 0.01** | **r = 0.19**  **p = 0.01** | **r = 0.59**  **p < 0.01** | **r = 0.69**  **p < 0.01** | r = 0.01  p = 0.94 | r = -0.09  p = 0.27 | **r = 0.59**  **p < 0.01** | r = 0.11  p = 0.18 |
| Base  HRT-SD  T Score | -- | -- | -- | **r = 0.66**  **p < 0.01** | **r = - 0.26**  **p < 0.01** | **r = 0.18**  **p = 0.01** | **r = 0.26**  **p < 0.01** | **r = 0.16**  **p = 0.04** | **r = 0.64**  **p < 0.01** | **r = 0.47**  **p < 0.01** | r = -0.14  p = 0.07 | **r = 0.18**  **p = 0.01** | **r = 0.23**  **p = 0.01** | r = 0.10  p = 0.22 | **r = 0.67**  **p < 0.01** | **r = 0.51**  **p < 0.01** | r = -0.15  p = 0.08 | **r = 0.19**  **p = 0.01** |
| Base  HRT-ISI  T Score | -- | -- | -- | -- | **r = -0.24**  **p < 0.01** | r = 0.11  p = 0.09 | r = 0.03  p = 0.67 | r = -0.04  p = 0.60 | **r =0.49**  **p < 0.01** | **r = 0.61**  **p < 0.01** | **r = -0.18**  **p = 0.02** | r = 0.10  p = 0.16 | r = 0.04  p = 0.61 | r = -0.05  p = 0.59 | **r = 0.51**  **p < 0.01** | **r = 0.54**  **p < 0.01** | r = -0.17  p = 0.05 | r = 0.10  p = 0.22 |
| Base  Impulsivity Index  Raw Score | -- | -- | -- | -- | -- | r = 0.07  p= 0.28 | **r = 0.45**  **p < 0.01** | **r = 0.52**  **p < 0.01** | r = -0.09  p = 0.26 | **r =-0.16**  **p = 0.04** | **r = 0.69**  **p < 0.01** | **r = 0.17**  **p = 0.03** | **r = 0.38**  **p < 0.01** | **r = 0.50**  **p < 0.01** | **r = -0.21**  **p = 0.01** | **r = -0.23**  **p = 0.01** | **r = 0.62**  **p < 0.01** | r = 0.06  p = 0.48 |
| Base  BIS -11  Raw Score | -- | -- | -- | -- | -- | -- | **r = 0.19**  **p = 0.01** | **r = 0.19**  **p = 0.01** | **r = 0.16**  **p =0.04** | r = 0.04  p = 0.64 | r = -0.02  p = 0.82 | **r = 0.74**  **p < 0.01** | r = 0.16  p = 0.05 | r = 0.12  p = 0.12 | **r = 0.18**  **p = 0.02** | r = 0.03  p = 0.69 | r = 0.01  p = 0.88 | **r = 0.71**  **p < 0.01** |
| 6-MO  d’  T Score | -- | -- | -- | -- | -- | -- | -- | **r = 0.89**  **p < 0.01** | **r = 0.43**  **p < 0.01** | r = 0.08  p = 0.28 | **r = 0.57**  **p < 0.01** | **r = 0.27**  **p = 0.01** | **r = 0.64**  **p < 0.01** | **r = 0.62**  **p < 0.01** | **r = 0.23**  **p < 0.01** | r = 0.07  p = 0.38 | **r = 0.47**  **p < 0.01** | r = 0.15  p = 0.07 |
| 6-MO  Comm  T Score | -- | -- | -- | -- | -- | -- | -- | -- | **r = 0.27**  **p < 0.01** | r = 0.01  p = 0.91 | **r = 0.77**  **p < 0.01** | **r = 0.23**  **p < 0.01** | **r = 0.63**  **p < 0.01** | **r = 0.71**  **p < 0.01** | r = 0.10  p = 0.24 | r = 0.02  p=0 .77 | **r = 0.60**  **p < 0.01** | r = 0.14  p = 0.08 |
| 6-MO  HRT-SD  T Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | **r = 0.63**  **p < 0.01** | **r = -0.17**  **p = 0.02** | **r = 0.23**  **p < 0.01** | **r = 0.28**  **p < 0.01** | r = 0.13  p = 0.13 | **r = 0.73**  **p < 0.01** | **r = 0.48**  **p < 0.01** | **r = -0.19**  **p = 0.02** | **r = 0.21**  **p = 0.01** |
| 6-MO  HRT-ISI  T Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | **r = -0.17**  **p = 0.02** | r = 0.04  p = 0.62 | r = 0.15  p = 0.08 | r = 0.04  p = 0.67 | **r = 0.54**  **p < 0.01** | **r = 0.65**  **p < 0.01** | r = -0.13  p = 0.13 | r =0.05  p = 0.52 |
| 6-MO  Impulsivity Index  Raw Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | r = 0.04  p = 0.57 | **r = 0.42**  **p < 0.01** | **r = 0.56**  **p < 0.01** | **r = 0.40**  **p = 0.01** | r = 0.16  p = 0.20 | **r = 0.60**  **p < 0.01** | r = -0.04  p = 0.62 |
| 6-MO  BIS-11  Raw Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | **r = 0.22**  **p < 0.01** | **r = 0.20**  **p = 0.01** | **r = 0.18**  **p = 0.03** | r = 0.07  p = 0.42 | r = 0.10  p = 0.24 | **r = 0.75**  **p < 0.01** |
| 12-MO  d’  T Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | **r = 0.88**  **p < 0.01** | **r = 0.40**  **p < 0.01** | **r = 0.16**  **p = 0.04** | **r = 0.60**  **p < 0.01** | **r = 0.20**  **p = 0.01** |
| 12-MO  Comm  T Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | r = 0.16  p =0.05 | r = 0.03  p = 0.69 | **r = 0.82**  **p < 0.01** | **r = 0.20**  **p = 0.01** |
| 12-MO  HRT-SD  T Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | **r = 0.59**  **p < 0.01** | **r = -0.23**  **p < 0.01** | **r = 0.24**  **p < 0.01** |
| 12-MO  HRT-ISI  T Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | r = -0.12  p = 0.14 | r = 0.11  p = 0.18 |
| 12-MO  Impulsivity Index  Raw Score | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | r = 0.05  p = 0.57 |

1. Supplementary References
2. Burton, P., Gurrin, L., & Sly, P. (1998). Extending the simple linear regression model to account for correlated responses: An introduction to generalized estimating equations and multi-level mixed modelling. *Stat Med, 17*, 1261–1291.
3. Conners, C.K. (2014). *Conners Continuous Performance Test* (3rd ed.). Toronto, ON: Multi-Health Systems Inc.
4. Hubbard, A. E., Ahern, J., Fleischer, N. L., Van der Laan, M., Lippman, S. A., Jewell, N., … Satariano, W. A. (2010). To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology, 21*, 467–474.
5. Linehan, M.M., Comtois, K.A., Brown, M.Z., Heard, H.L., & Wagner, A. (2006). Suicide attempt self-injury interview (SASII): development, reliability, and validity of a scale to assess suicide attempts and intentional self-injury. *Psychol Assess, 18(*3), 303–12.
6. McMain, S.F., Chapman, A.L., Kuo, J.R., Dixon-Gordon, K.L., Guimond, T.H., Labrish, C., … Streiner, D.L. (2022) The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*, 1-16. doi:10.1159/000525102.
7. Mendenhall, W,, & Sincich, T. (1996). *A Second Course in Statistics: Regression Analysis (5th ed.).* Upper Saddle River, NJ: Prentice-Hall. p. 422.
8. Stanford, M.S., Mathias, C.W., Dougherty, D.M., Lake, S.L., Anderson, N.E., & Patton, J.H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences, 47*, 385-395.
9. Linehan, M.M. & Heard, H.L. (1987). *Treatment history interview (THI-2).* Seattle, WA: University of Washington; Unpublished work (do not use without citation).