



## UvA-DARE (Digital Academic Repository)

### Psychometric properties of the 4-week version of the Borderline Personality Disorder Severity Index-5

van Trigt, Shanna; Mendoza-Alvarez, Mariana; van der Zweerde, Tanja; De Picker, Livia; van Straten, Annemieke; Arntz, Arnoud; van Marle, Hein J.F.

**DOI**

[10.1016/j.comppsy.2025.152634](https://doi.org/10.1016/j.comppsy.2025.152634)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

Comprehensive Psychiatry

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

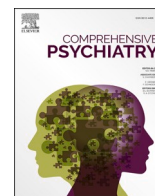
van Trigt, S., Mendoza-Alvarez, M., van der Zweerde, T., De Picker, L., van Straten, A., Arntz, A., & van Marle, H. J. F. (2025). Psychometric properties of the 4-week version of the Borderline Personality Disorder Severity Index-5. *Comprehensive Psychiatry*, 143, Article 152634. <https://doi.org/10.1016/j.comppsy.2025.152634>

**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>)*



## Psychometric properties of the 4-week version of the Borderline Personality Disorder Severity Index-5

Shanna van Trigt<sup>a,b,c,\*</sup>, Mariana Mendoza-Alvarez<sup>d,e</sup>, Tanja van der Zweerde<sup>f</sup>,  
Livia De Picker<sup>d,e</sup>, Annemieke van Straten<sup>b,c</sup>, Arnoud Arntz<sup>g,h</sup>, Hein J.F. van Marle<sup>a,f,i,j</sup>

<sup>a</sup> Amsterdam UMC, location Vrije Universiteit, Psychiatry, Amsterdam, the Netherlands

<sup>b</sup> Amsterdam Public Health, Mental Health program, Amsterdam, the Netherlands

<sup>c</sup> Vrije Universiteit Amsterdam, Clinical, Neuro and Developmental Psychology, Amsterdam, the Netherlands

<sup>d</sup> Collaborative Antwerp Psychiatric Research Institute (CAPRI), University of Antwerp, Antwerp, Belgium

<sup>e</sup> Scientific Initiative of Neuropsychiatric and Psychopharmacological Studies (SINAPS), University Psychiatric Hospital Campus Duffel, Duffel, Belgium

<sup>f</sup> GGZ inGeest Mental Health Care, Amsterdam, the Netherlands

<sup>g</sup> Department of Clinical Psychology, University of Amsterdam, Amsterdam, the Netherlands

<sup>h</sup> Academic Center for Trauma and Personality, Amsterdam, the Netherlands

<sup>i</sup> Amsterdam Neuroscience, Mood, Anxiety, Psychosis, Sleep & Stress program, Amsterdam, the Netherlands

<sup>j</sup> ARQ National Psychotrauma Center, Diemen, the Netherlands

### ARTICLE INFO

#### Keywords:

Borderline personality disorder  
BPDSI  
Psychometric properties  
Reliability  
Validity

### ABSTRACT

This study evaluated the psychometric properties of a new version of the Borderline Personality Disorder Severity Index-5 that assesses BPD symptom severity over the course of 4 weeks instead of the standard 3 months: the BPDSI-5-4wk. Reliability and validity were evaluated in a mixed sample of patients with BPD ( $n = 92$ ), patients with avoidant personality disorder (APD) as clinical control group ( $n = 16$ ), and a non-patient control group ( $n = 20$ ). The study demonstrated very high interrater agreement and test-retest reliability and acceptable to excellent internal consistencies of the BPDSI-5-4wk. Confirmatory factor analysis supported its assumed nine-factor structure. The BPDSI-5-4wk also showed very good construct (i.e. known-group) validity, as well as criterion-related (i.e. concurrent) validity when correlating the BPDSI-5-4wk with the BPDSI-5 (3-month version), the structured clinical interview for DSM personality disorders (SCID-5-P), and several other BPD- and other mental health-related self-report questionnaires. We additionally derived cut-off scores with high sensitivity and specificity for distinguishing BPD from clinical controls (21.02, 85 % sensitivity, 94 % specificity), from non-patient controls (10.50, 98 % sensitivity, 100 % specificity), and from both control groups combined (17.26, 93 % sensitivity, 92 % specificity). A reliable change criterion of 5.88 was established. Preliminary trial data additionally showed the BPDSI-5-4wk's sensitivity to change. The strong reliability and validity of the BPDSI-5-4wk support its value for detailed, dimensional assessment of BPD symptom severity over shorter time frames. This will facilitate frequent treatment response evaluations, rapid indication of follow-up treatment, and better alignment with shorter intervention and follow-up periods in clinical trials.

### 1. Introduction

Borderline personality disorder (BPD) is a severe psychiatric disorder characterized by instability in emotions, interpersonal relationships, and self-image, often accompanied by impulsive and self-harming or suicidal behavior [1]. The prevalence of BPD is estimated at 1–1.5 % in the general population, and up to 10–20 % in clinical populations [2]. Patients often suffer from severe social and occupational dysfunction,

which makes BPD an important public health problem [3].

BPD is a disorder with large variability and instability in symptom presentation and therefore requires reliable and valid instruments to capture symptom severity in both clinical and research settings [4,5]. Currently, the Borderline Personality Disorder Severity Index (BPDSI) is one of the gold standard clinical interviews to assess the severity of BPD symptomatology, providing dimensional scores corresponding to the nine Diagnostic and Statistical Manual of Mental Disorders (DSM) BPD

\* Corresponding author at: Amsterdam UMC, location Vrije Universiteit, Psychiatry, De Boelelaan 1117, Amsterdam, The Netherlands.

E-mail address: [s.vantrigt@amsterdamumc.nl](mailto:s.vantrigt@amsterdamumc.nl) (S. van Trigt).

criteria [5,6]. The BPDSI provides a comprehensive dimensional assessment of each BPD symptom domain with multiple items, which is recommended given the heterogeneity of BPD, to adequately capture changes in specific symptom domains [4]. In the transition to the DSM-5, no changes were made in the diagnostic criteria of BPD and currently the BPDSI-5 is used (with small adaptations with respect to the BPDSI-IV) [5,7]. The BPDSI-IV and – 5 assess BPD symptomatology over the course of the last 3 months. Both from a clinical and research perspective however, there is a need to reliably assess BPD symptoms over a shorter time frame to track acute symptom patterns, initiate and evaluate BPD treatment trajectories and quantify outcomes in clinical research.

A shorter-term BPD severity assessment allows for a more rapid treatment indication of follow-up treatment and more frequent (intermittent) evaluations of treatment response. This is particularly relevant in cases where patients may transition between treatments (e.g., from post-traumatic stress disorder (PTSD) or substance use treatment to BPD-focused interventions). With the current BPDSI-5, treatment indication can be determined after a period of at least 3 months, potentially delaying access to appropriate care. A shorter assessment period during ongoing treatment additionally allows for the timely adjustment or termination of non-effective treatments. Together, this may improve the clinical trajectory of patients, result in a more efficient allocation of care and reduce waiting lists and treatment costs.

Clinical research into BPD treatments likewise benefits from a shorter evaluation period of symptom severity. This is especially true for relatively short psychotherapeutic treatment approaches or medication trials e.g., [8], in which the 3-month evaluation period of the BPDSI-5 is disproportionately long with respect to trial period. Shorter trial and intervention periods have clear benefits over long and expensive clinical trials, which have an increased likelihood of failed trial due to drop-out or non-compliance. In addition, upcoming studies investigating the additive effect of treatment approaches focused on e.g. (childhood)trauma [9] or sleep [10] prior or subsequent to standard BPD treatment, also require shorter-term assessment of BPD symptoms to measure (intermittent) changes due to treatment (steps). Recent methodological recommendations for personality disorder (PD) research emphasized the importance of sufficiently frequent assessments to allow evaluation of complex growth patterns, as PD symptoms often develop or change (e.g. in response to treatment) in nonlinear or discontinuous ways [11].

To address the need for a shorter evaluation period while retaining a comprehensive and dimensional assessment, we aimed to evaluate the psychometric properties of a newly developed version of the BPDSI-5 that assesses BPD symptom severity over the course of the last 4 weeks: the BPDSI-5-4wk. To this end, the reliability of the BPDSI-5-4wk (i.e., interrater agreement, test-retest reliability, and internal consistency), and its validity (i.e., construct validity (using known-group validity), and criterion-related validity (using concurrent validity)) were assessed. Regarding known-group validity we expected that patients with BPD scored significantly higher on all BPDSI-5-4wk items, subscales, and total score than clinical controls (consisting of patients with an avoidant personality disorder (APD)) and non-patient controls. For concurrent validity, significant positive correlations between the BPDSI-5-4wk and several current BPD measures were expected, and negative correlations were expected with measures of quality of life and mental well-being. Finally, we aimed to establish clinical cutoff scores of the BPDSI-5-4wk designating BPD pathology, as well as a reliable change criterion indicating statistically meaningful change.

## 2. Methods

### 2.1. Participants

We included 92 patients diagnosed with BPD, 16 patients diagnosed with APD (and  $\leq 2$  BPD-traits) as clinical control group, and 20 non-patient controls. Patients were recruited at two outpatient PD departments in the Netherlands (GGZ inGeest and Amsterdam Center for

Trauma and Personality), and one university psychiatric hospital offering in- and outpatient dialectical behavior therapy (DBT) treatment programs in Belgium (UPC Duffel). APD was selected as clinical control group given the high incidence of this subtype of PD. Non-patients were recruited from the general population using convenience sampling through the researchers' own networks with efforts to ensure variation in age, sex, and education to enhance representativeness.

Participants met the following inclusion criteria: (1) 18–65 years old, (2) a DSM-5 BPD (BPD group) or APD (clinical control group) diagnosis, confirmed through administration of the Structured Clinical Interview for DSM-5 Personality disorders (SCID-5-P) [12], (3) capable to provide informed consent, (4) mental and physical ability to complete self-report and interview assessments. Exclusion criteria were: (1) more than 2 BPD symptoms for the APD control group and any PD or more than 2 BPD symptoms for the non-patient control group as assessed with the SCID-5-P, (2) diagnosis or treatment for mental health complaints in the past year for the non-patient control group, according to self-report. Patients were not excluded for having other co-morbid psychiatric disorders. Study participation was voluntary and only commenced after obtaining written informed consent. The study was reviewed by the Medical Ethics Committee of the Amsterdam University Medical Center, which declared that the study fell outside the scope of the Medical Research Involving Human Subjects Act (WMO) (date: 9-6-2021, registration number: 2021.0266). The study was registered in the Data Processing Registry of the Amsterdam UMC and approved by the science committee of GGZ inGeest (date: 14-9-2021, registration number: CWO 2021–011). All procedures were in line with institutional guidelines and the Declaration of Helsinki. Participants recruited in the Netherlands received financial compensation for participation. Participants in Belgium were recruited as part of another study and provided informed consent for their data to be used in the current study [13]. They did not receive financial compensation for their participation in the current study.

### 2.2. Measures

#### 2.2.1. Borderline personality disorder severity index-5 (3-month version)

The BPDSI-5 is a semi-structured, clinician-rated interview that measures BPD symptom severity over the last three months [7]. The BPDSI-5 is a slightly adapted version from the BDPSI-IV [5], with DSM-criteria for BPD not changing from DSM-IV to DSM-5. The BPDSI-5 consists of 70 items which can be grouped into nine subscales representing the BPD DSM-5 criteria (fear of abandonment, unstable relationships, self-image disturbance, impulsivity, (para)suicidal behavior, affective instability, emptiness, anger attacks, and dissociation and paranoid ideation). The interviewer rated the frequency in the past three months for each item on an 11-point scale, ranging from 0 (never) to 10 (daily). The items of the subscale self-image are an exception; these were scored on a 5-point Likert scale ranging from 0 (absent) to 4 (e.g., clear and paramount instability of the self-image) and multiplied by 2.5 to obtain an equally weighted score compared to other subscales. Total scores for each subscale were computed by taking the average of the corresponding items and were summed to obtain the total score of all criteria combined, ranging between 0 and 90. The BPDSI-IV demonstrated high interrater reliability, acceptable to high internal consistencies, and very good known-group and concurrent validity [5].

#### 2.2.2. Borderline personality disorder severity index-5–4-week version

The studied BPDSI-5-4wk is an adjusted version of the BPDSI-5 in which the interviewers rated BPD symptoms over the past 4 weeks. The BPDSI-5-4wk implements an 8-point scale ranging from 0 to 10. To maintain the frequency-to-time ratio as applied in scoring of the BPDSI-5, scores 1, 2, and 4 were omitted in the 4-week version. Scoring categories from the BPDSI-5 were adapted to correspond with the 4-week assessment period, with scores of 0 (never), 3 (once), 5 (twice), 6 (three times), 7 (once per week), 8 (multiple times per week, but less than half of the week), 9 (more than half of the week), and 10 (daily).

Total scores for each subscale and all subscales combined were computed identically as for the BPDSI-5. The English version of the BPDSI-5-4wk can be found in Appendix A.

### 2.2.3. Structured clinical interview for DSM-5 personality disorders

The SCID-5-P is a semi-structured interview consisting of 119 items, designed to diagnose PDs according to the DSM-5 criteria [12]. Items can be scored either as absent (0), subthreshold (1) or above threshold (2). To shorten the time of administration, prior to administering, participants filled out the SCID-5 self-report personality questionnaire (SCID-5-SPQ) consisting of 106 yes/no questions as a screening tool. Only the sections that scored above threshold (in addition to the BPD section) were administered during the SCID-5-P interview. The SCID-5-P is a revision of the SCID-II, which in Dutch samples proved to be a reliable and valid instrument, with good test-retest reliability [14] and interrater agreement for each PD (BPD:  $\kappa = 0.91$ , APD:  $\kappa = 0.83$ ) [15]. In participants from Belgium, the SCID-II was used, which is comparable to the SCID-5-P, and only the BPD section of the interview was conducted and scored as either absent or above threshold.

### 2.2.4. Ultra-short borderline personality disorder – Checklist

The ultra-short BPD – checklist (BPD-C) measures symptoms of BPD over the last month [16] and is a shortened version of the 47-item self-report questionnaire BPD-C [17,18]. The ultra-short BPD-C is a 10-item self-report questionnaire, with items corresponding to the DSM-5 BPD diagnostic criteria, except for criterion 9 which was assessed with two items (item 9 and 10, that were averaged). The ultrashort BPD-C is scored on a 5-point Likert scale ranging from 1 (not at all) to 5 (very much), and has good internal consistency, as well as good discriminant and construct validity [16].

### 2.2.5. Mental health quality of life questionnaire

The mental health quality of life questionnaire (MHQoL) assesses self-reported quality of life in people with mental health problems [19]. The questionnaire consists of 7 dimensions: self-image, independence, mood, relationships, daily activities, physical health, and hope, scored on a 4-point scale indicating how well or unwell that dimension is currently rated by the respondent. The MHQoL has high internal consistency, test-retest reliability, and construct and discriminative validity [19].

### 2.2.6. Positive mental health scale

The positive mental health scale (PMH) is a self-report scale for the assessment of positive mental health, and consists of 9 items rated on a Likert scale ranging from 1 (not true) to 4 (true) [20]. The PMH has high internal consistency, good retest-reliability, good convergent and discriminant validity, and is sensitive to therapeutic change [20].

### 2.2.7. Cantril's ladder of life

Cantril's ladder of life (CL) is a 2-item scale assessing life satisfaction, on which respondents rated their own life on a scale ranging from 0 (the worst possible life for you) to 10 (the best possible life for you), for the present moment and the expected situation in 5 years [21]. CL demonstrated adequate reliability and validity [22].

## 2.3. Study procedure

Participants were recruited after clinical intake and were generally not receiving any treatment or were in the beginning of a treatment program at the time of participation. The SCID-5-P was administered to confirm eligibility based on (the absence of) BPD and APD diagnoses. Based on the SCID-5-SPQ, other sections of the SCID-5-P were also administered to establish possible comorbid PD diagnoses (except for the 36 participants from Belgium). In case the interview was already administered within 6 months before participation, relevant data was copied from the medical file to prevent double administration of the

interview, only after participants provided explicit informed consent. If eligible, further assessments were conducted. Four measures assessed (current) severity of BPD symptoms: (1) the BPDSI-5-4wk, (2) the BPDSI-5, (3) BPD section of the SCID-5-P/SCID-II, and (4) the ultra-short BPD-C. The BPDSI-5-4wk and BPDSI-5 were administered by trained interviewers during an appointment on location or through video call. To establish the test-retest reliability, the BPDSI-5-4wk was repeated after one week in 19 patients with BPD and in 20 non-patients by the same interviewer. For the patient subsample, we invited all patients that were recruited in the Netherlands to complete the retest. The final subsample consisted of those who were willing and successfully scheduled for a second assessment within the 1-week window. Participants were asked to fill out the self-report questionnaires online within 1 week after the first BPDSI-5(-4wk) administration. To establish interrater agreement of the BPDSI-5-4wk, part of the interviews (16 in the BPD group, 10 in the APD group, and 11 in the non-patient group) were audio-recorded for double-scoring by raters blind to diagnostic group, for which participants provided separate informed consent.

## 2.4. Analyses

Interrater agreement between two independent raters was analyzed by means of Krippendorff's  $\alpha$ . Test-retest reliability was established by the Intraclass Correlation Coefficient (ICC) between two consecutive assessments of the BPDSI-5-4wk, using an absolute-agreement, two-way, mixed-effect model. The factor structure of the BPDSI-5-4wk was examined using confirmatory factor analysis (CFA) to evaluate the fit of the DSM-5 based nine-factor model. Internal consistency was established with classical reliability indices Cronbach's alpha and McDonald's omega total based on a global exploratory factor approach, providing a general estimate of reliability. Since these analyses do not account for the specific factor structure, we additionally inspected CFA-based reliability estimates omega total and ordinal alpha, which account for the latent factor structure as well as the ordinal nature of the items and offer a more precise and theoretically grounded evaluation of reliability for the BPDSI-5-4wk total scale and subscales. CFA-based reliability indices were estimated only in the full sample, given the limited sample size of subgroups impairing stable estimation of model parameters. Global reliability indices were additionally inspected within the BPD group and, for the 'Relationships' subscale, within participants with and without a partner relationship. Item-rest correlations were additionally inspected to evaluate internal consistency. Known-group validity of the BPDSI-5-4wk subscales and total scores between patients with BPD and patients with APD or non-patients was determined using planned group comparisons. Depending on data characteristics, we conducted robust linear regressions (in case of violation of the homoscedasticity assumption), or robust ANCOVAs (in case of violation of the normal distribution and homoscedasticity assumptions) with and without relevant demographics as covariates. Concurrent validity was assessed with Spearman's (partial) correlations. To control for the significant group difference in educational status, missing data for this variable were handled using multiple imputation ( $m = 5$ ), to allow inclusion as a covariate in the known-group and concurrent validity analyses, with pooled estimates reported. No other variables were imputed, and all available data were used with pairwise inclusion. All tests were interpreted with significance levels of  $\alpha = 0.05$  (except for known-group validity, for which  $\alpha = 0.05/18 = 0.0028$ , to control for multiple testing of the nine symptom subscales compared with both control groups). Analyses were performed using R [23], version 4.3.2.

Cutoff scores were based on receiver operating characteristic (ROC) analyses, with the exact threshold scores derived by 3 methods: Euclidian distance, Youden-index, and Liu's index. The optimal cutoff was indicated by: (i) the minimum of the Euclidean distance expressed as  $\sqrt{[(1-\text{sensitivity})^2 + (1-\text{specificity})^2]}$  [24]; (ii) the maximum of the Youden-index expressed as  $(\text{sensitivity} + \text{specificity} - 1)$  [25]; (iii) and the maximum of the product of sensitivity and specificity [26]. We

expected the three methods to converge. In case of nonconvergence, the mean of the three values would be used. Finally, the reliable change criterion was based on Jacobson and Truax’s method [27], which establishes whether a change in clinical symptoms is significantly greater than what could have occurred due to random measurement error alone. We used the test-retest reliability (ICC value obtained from the full sample) to establish the reliable change criterion, which is preferred over internal consistency as it captures stability over time and better aligns with typical use of the reliable change criterion in clinical settings where repeated measurements are involved [28].

2.5. Power calculations

Power calculations (2-tailed  $\alpha = 0.05$ ; 80 % power) for the correlational analyses to establish concurrent validity showed that 84 participants were needed to detect correlations of  $r > 0.30$ . Therefore, we aimed to include 84 BPD participants. Based on previous psychometric evaluations of the BPDSI [5,6,29], we expected large effect sizes ( $d > 0.80$ ) in the group comparison analyses needed to establish known-group validity. Power calculations showed that  $n = 16$  participants were needed for the clinical and non-patient control groups to detect effect sizes of  $d > 0.80$  in the group comparison analyses. However, to account for the potential loss of degrees of freedom in case covariates needed to be included, we aimed to include 84 patients with BPD, 20 patients with APD, and 20 non-patients, resulting in a total sample size

of  $n = 124$ .

We conducted post-hoc and sensitivity power analyses to assess the ability of our study to detect significant effects. For the correlation analyses ( $\alpha = 0.05$ ), we had 93 % and 84 % power to detect correlations of  $r > 0.30$ , and we could detect correlations of  $r > 0.17$  and  $r > 0.20$  with 80 % power in the full sample ( $n = 128$ ) and BPD group ( $n = 92$ ), respectively. For group comparisons ( $\alpha = 0.0028$ ), we had 82 % and 88 % power to detect effect sizes of  $d > 0.80$ , and we could detect effect sizes of  $d > 0.77$  and  $d > 0.70$  with 80 % power between BPD and APD patients ( $n = 16$ ) and between BPD and non-patients ( $n = 20$ ), respectively.

3. Results

3.1. Sample characteristics and group comparisons

Demographic and clinical characteristics of participants are displayed in Table 1. No significant differences were found between groups for the distribution of sex. The number of comorbid PDs was compared only between the BPD and APD groups, and did not differ significantly. Age and educational status did differ significantly between groups, with post-hoc pairwise comparisons showing significantly lower age in the BPD group than in the APD group, and lower educational status in the BPD group than in the non-patient group. Age and educational status were therefore added as covariates in full-sample correlational and

Table 1  
Demographic and clinical characteristics.

	BPD (n = 92)	APD (n = 16)	NP (n = 20)	Full sample (n = 128)	Group differences $\chi^2$	Post-hoc analysis, BPD vs.	
						APD z	NP z
Age					9.14* <sup>a</sup>	-3.01**	-0.25
Mean (SD)	30.0 (9.6)	37.8 (10.4)	31.1 (12.0)	31.2 (10.3)			
Range	18–60	23–61	20–63	18–63			
Female, n (%)	64 (70 %)	10 (63 %)	14 (70 %)	88 (69 %)	0.33	–	–
Educational status, n (%)					12.62***	-2.10	-3.19**
No education	1 (1.1 %)	0 (0 %)	0 (0 %)	1 (0.8 %)			
Primary education	5 (5.4 %)	0 (0 %)	0 (0 %)	5 (3.9 %)			
Secondary education	33 (35.9 %)	1 (6.3 %)	4 (20.0 %)	38 (29.7 %)			
Post-secondary non-tertiary education	9 (9.8 %)	5 (31.3 %)	0 (0 %)	14 (10.9 %)			
College degree	18 (19.6 %)	5 (31.3 %)	5 (25.0 %)	28 (21.9 %)			
University degree	18 (19.6 %)	5 (31.3 %)	11 (55.0 %)	34 (26.6 %)			
Missing	8 (8.7 %)	0 (0 %)	0 (0 %)	8 (6.3 %)			
BPD symptom severity, M (SD) <sup>b</sup>							
BPDSI-5-4wk total score	30.93 (10.36)	11.90 (6.88)	4.70 (3.12)	24.45 (13.99)			
Range	7.25–57.9	2.31–25.8	0.31–10.3	0.31–57.9			
BPDSI-5 total score	31.90 (10.02)	12.73 (7.15)	4.85 (2.92)	24.84 (14.14)			
Range	8.85–59.2	3.45–25.6	1.00–13.1	1.00–59.2			
Missing, n (%)	8 (8.7 %)	0 (0 %)	0 (0 %)	8 (6.3 %)			
SCID-5-P (# BPD traits)	6.5 (1.3)	0.6 (0.9)	0.1 (0.5)	4.7 (3.0)			
Range	5–9	0–2	0–2	0–9			
Ultra short BPD-C	30.36 (6.95)	19.47 (7.85)	10.90 (2.10)	25.76 (9.96)			
Range	11–44	11–40	9–17	9–44			
Missing, n (%)	7 (7.6 %)	1 (6.3 %)	0 (0 %)	8 (6.3 %)			
(Co-morbid) PDs, n (%)					U <sup>c</sup> 513		
Antisocial PD	1 (1 %)	0 (0 %)	0 (0 %)	1 (1 %)			
Avoidant PD	9 (10 %)	16 (100 %)	0 (0 %)	25 (20 %)			
Dependent PD	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)			
Histrionic PD	1 (1 %)	0 (0 %)	0 (0 %)	1 (1 %)			
Narcissistic PD	1 (1 %)	0 (0 %)	0 (0 %)	1 (1 %)			
Obsessive-Compulsive PD	3 (3 %)	2 (13 %)	0 (0 %)	5 (4 %)			
Paranoid PD	4 (4 %)	0 (0 %)	0 (0 %)	4 (3 %)			
Schizoid PD	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)			
Schizotypal PD	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)			
Missing <sup>d</sup>	34 (37 %)	0 (0 %)	0 (0 %)	34 (27 %)			

Note. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ . APD: avoidant personality disorder control group; BPD-C: Borderline Personality Disorder-Checklist; BPDSI: Borderline Personality Disorder Severity Index; NP: non-patient control group; PD: personality disorder; SCID-5-P: Structured Clinical Interview for DSM-5 Personality disorders. Groups were compared using Kruskal-Wallis tests for age and educational status, Chi-squared test for sex, and Mann-Whitney U test for number of co-morbid PDs. <sup>a</sup> non-parametric test was used due to non-normally distributed data; <sup>b</sup> no group comparisons were performed given the inherent differences between groups; <sup>c</sup> group comparison only between BPD and APD groups; <sup>d</sup> data about co-morbid PDs was not available for participants from Belgium.

between-group analyses.

Table 1 additionally presents descriptive statistics for all BPD severity measures administered in the study, including the BPDSI-5-4wk, BPDSI-5, number of BPD criteria met on the SCID-5-P, and the ultra-short BPD-C. These scores reflect a broad and clinically relevant range of BPD symptom severity within the BPD group.

### 3.2. Interrater agreement

Interrater agreement for the BPDSI-5-4wk total and subscale scores is reported in Table 2. Agreement between raters was near perfect for total and subscale scores in both the full sample and the BPD-only group (Krippendorff's  $\alpha \geq 0.96$ ).

### 3.3. Test-retest reliability

The average time between consecutive assessments was 6.90 days. Test-retest reliability of the BPDSI-5-4wk total and subscale scores is reported in Table 2. For the total score, test-retest reliability was excellent in both groups. Reliability of the subscales ranged from good to excellent in the full sample and from moderate to good in the BPD group.

### 3.4. Factor structure

Prior to conducting the CFA, we examined the distribution of item responses. Three items: 2.4, 5.13, and 8.6 were excluded from the CFA due to insufficient variability, as they contained only two unique response categories preventing reliable estimation of polychoric correlations. CFAs indicated that a nine-factor model, with items loading onto their respective BPDSI subscale, provided a significantly better fit to the data than a one-factor model ( $\chi^2 \text{ diff}(36) = 372.76, p < .001$ ). Fit indices of the nine-factor model indicated acceptable to good model fit (CFI = 0.91, TLI = 0.91, RMSEA = 0.039 [95 % CI: 0.033, 0.045]). Although several items showed weak standardized factor loadings (< 0.30, items 2.8, 4.4, and 5.3), factor loadings were generally moderate to high, supporting the assumed factor structure (Supplementary Table S1).

### 3.5. Internal consistency

Internal consistency of the BPDSI-5-4wk total score based on classical Cronbach's alpha and McDonald's omega estimates was excellent in both the full sample and BPD group (Table 3). Internal consistency of the subscales ranged from questionable to good in the full sample and from poor to acceptable in the BPD group. Additional CFA-based reliability estimates (full sample only) to account for the latent factor structure and the ordinal nature of the items indicated similar or higher reliability, with excellent internal consistency for the total scale and acceptable to good internal consistency for the subscales (Table 3).

Item-rest correlations of the total score ranged from  $r = 0.00$  to  $0.76$ ,

**Table 2**  
Interrater agreement and test-retest reliability of the BPDSI-5-wk.

BPDSI-5-4wk scale	Interrater agreement				Test-retest reliability			
	Full sample (n = 37)		BPD group (n = 16)		Full sample (n = 39)		BPD group (n = 19)	
	$\alpha$	95 % CI	$\alpha$	95 % CI	ICC	95 % CI	ICC	95 % CI
Fear of abandonment	0.98	0.94, 1.00	0.98	0.94, 0.99	0.89	0.81, 0.94	0.75	0.45, 0.89
Relationships	0.99	0.99, 1.00	0.99	0.99, 0.99	0.87	0.75, 0.93	0.83	0.62, 0.93
Identity disturbance	0.96	0.93, 0.99	0.96	0.93, 0.99	0.95	0.90, 0.98	0.88	0.67, 0.95
Impulsivity	0.99	0.98, 1.00	0.99	0.98, 1.00	0.91	0.84, 0.95	0.89	0.73, 0.95
(Para)suicide	0.98	0.96, 1.00	0.98	0.96, 1.00	0.90	0.82, 0.95	0.86	0.67, 0.94
Affective instability	0.99	0.99, 1.00	0.99	0.99, 1.00	0.94	0.89, 0.97	0.80	0.56, 0.92
Emptiness	0.99	0.99, 1.00	0.99	0.99, 1.00	0.87	0.77, 0.93	0.78	0.51, 0.91
Anger attacks	0.99	0.99, 1.00	0.99	0.99, 1.00	0.78	0.61, 0.88	0.73	0.42, 0.89
Dissociation	0.99	0.97, 1.00	0.99	0.97, 0.99	0.87	0.77, 0.93	0.80	0.55, 0.92
Total score	0.99	0.99, 1.00	0.99	0.97, 1.00	0.98	0.96, 0.99	0.94	0.84, 0.98

Note. Interrater agreement is defined by Krippendorff's alpha coefficient ( $\alpha$ ), test-retest reliability by intraclass correlation coefficient (ICC).

**Table 3**  
Internal consistency of the BPDSI-5-4wk total scale and sub-scales.

BPDSI-5-4wk scale	Full sample (n = 128)				BPD group (n = 92)	
	$\omega$	$\alpha$	$\omega$ (CFA)	Ordinal $\alpha$ (CFA)	$\omega$	$\alpha$
Fear of abandonment	0.75	0.74	0.80	0.86	0.68	0.66
Relationships	0.68	0.66	0.77	0.76	0.58	0.55
If in partner relationship (n = 73)	0.77	0.68	–	–	0.60	0.57
If no partner relationship (n = 55)	0.61	0.47	–	–	0.50	0.38
Identity disturbance	0.82	0.81	0.85	0.87	0.74	0.73
Impulsivity	0.67	0.66	0.70	0.73	0.64	0.62
(Para)suicide	0.70	0.69	0.82	0.68	0.66	0.64
Affective instability	0.87	0.87	0.87	0.88	0.77	0.75
Emptiness	0.79	0.78	0.82	0.82	0.69	0.67
Anger attacks	0.73	0.71	0.77	0.84	0.67	0.64
Dissociation	0.79	0.79	0.81	0.87	0.72	0.72
Total scale	0.96	0.94	0.96	0.96	0.93	0.89

Note. Internal consistency was defined by Cronbach's alpha coefficient ( $\alpha$ ) and McDonald's omega total coefficient ( $\omega$ ), and by the confirmatory factor analysis (CFA)-based reliability estimates  $\omega$  and ordinal  $\alpha$  to account for the latent factor structure and ordinal nature of the items.

$M = 0.42 \pm 0.18$ . We inspected items with low item-rest correlations ( $r \leq 0.20$ ) on the total score and their respective subscale. Low item-rest correlations were observed for zero-inflated items (see Supplementary Table S2 for an overview of zero-response percentage for each item), mostly assessing more severe behaviors with low occurrence (e.g., self-harm through biting or inserting objects into the body, suicide attempts, physically attacking others). We decided to keep these items due to their theoretical and clinical importance for capturing key BPD symptoms. Additionally, removing individual items did not substantially improve internal consistency (maximum increase  $\alpha = 0.05$ ), supporting the inclusion of all items for construct coverage.

### 3.6. Known-group validity

BPDSI-5-4wk subscale and total scores of all groups and group comparison statistics are reported in Table 4. Patients with BPD scored significantly higher on the total score and all subscales than both patients with APD and non-patients, irrespective of controlling for the covariates age and educational status. Both the uncontrolled and controlled group comparisons indicate very good known-group validity of the BPDSI-5-4wk, with large effect sizes for all comparisons ( $d \geq 0.87$ ). For the subscale (para)suicide, non-patients scored zero on all items, resulting in zero variance within this group and preventing direct group comparisons. Instead of the comparison between the BPD and non-patient groups, we therefore conducted one-sample *t*-tests to assess

**Table 4**  
BPDSI-5-4wk mean scores per group and group comparisons.

	BPD (n = 92)	APD (n = 16)	NP (n = 20)	BPD vs. APD		BPD vs. NP	
	Mean (SD)	Mean (SD)	Mean (SD)	t	d	t	d
<b>Without covariates</b>							
Fear of abandonment	3.17 (1.75)	0.94 (0.74)	0.34 (0.57)	7.98**	2.16	12.02**	2.97
Relationships	2.95 (1.72)	1.24 (1.14)	0.38 (0.68)	4.83**	1.31	10.52**	2.60
Identity disturbance	4.26 (2.32)	1.64 (1.42)	0.28 (0.42)	6.03**	1.63	13.57**	3.35
Impulsivity	1.69 (1.21)	0.51 (0.65)	0.48 (0.52)	5.87**	1.59	6.10**	1.51
(Para)suicide <sup>a</sup>	0.86 (0.93)	0.04 (0.13)	0.00 (0.00)	7.96**	2.16	8.89**	2.19
Affective instability	7.12 (2.13)	3.71 (2.39)	1.38 (1.28)	5.81**	1.57	16.64**	4.10
Emptiness	6.10 (2.54)	3.19 (2.98)	1.34 (1.59)	3.20*	0.87	10.39**	2.56
Anger attacks	2.32 (1.72)	0.43 (0.79)	0.48 (0.72)	7.21**	1.95	7.07**	1.74
Dissociation	2.50 (1.96)	0.20 (0.29)	0.02 (0.09)	6.16**	1.67	6.84**	1.69
Total score	30.90 (10.40)	11.90 (6.88)	4.70 (3.12)	9.49**	2.57	20.39**	5.03
<b>Controlled for covariates age and educational status</b>							
Fear of abandonment	3.07 (1.73)	0.93 (0.74)	0.33 (0.59)	6.16**	1.67	9.61**	2.37
Relationships	2.89 (1.70)	1.19 (1.15)	0.37 (0.62)	4.56**	1.23	9.01**	2.22
Identity disturbance	4.22 (2.31)	1.57 (1.41)	0.28 (0.45)	5.64**	1.53	11.84**	2.92
Impulsivity	1.42 (1.21)	0.45 (0.66)	0.47 (0.53)	5.51**	1.49	5.70**	1.41
(Para)suicide <sup>a</sup>	0.86 (0.91)	0.04 (0.19)	0.00 (0.14)	6.30**	1.71	9.01**	2.22
Affective instability	7.48 (2.14)	3.66 (2.39)	1.28 (1.27)	5.62**	1.52	13.90**	3.43
Emptiness	6.22 (2.52)	3.14 (2.90)	1.29 (1.63)	3.62*	0.98	9.99**	2.46
Anger attacks	2.34 (1.69)	0.38 (0.77)	0.46 (0.72)	6.48**	1.76	5.73**	1.42
Dissociation	2.18 (1.96)	0.20 (0.28)	0.02 (0.09)	6.53**	1.77	6.85**	1.69
Total score	30.93 (10.36)	11.90 (6.81)	4.70 (3.14)	9.32**	2.52	17.75**	4.38

Note. Adjusted significance levels to correct for multiple testing: \*p < .0028, \*\*p < .001. APD: avoidant personality disorder control group; NP: non-patient control group. Upper half displays means and group comparison statistics without controlling for covariates, lower half displays means and group comparison statistics controlled for the effect of age and educational status. <sup>a</sup>. Due to zero variance within the non-patient group for the subscale (para)suicide, one sample t-tests were conducted for the BPD vs NP comparisons to test if the uncontrolled and controlled BPD group means were significantly different from zero.

whether the BPD group means differed significantly from zero.

3.7. Concurrent validity

Correlation coefficients between the BPDSI-5-4wk total score and other (BPD) measures in both the full sample (also when controlled for age and educational status) and BPD group are displayed in Table 5. The results indicated significant positive associations with the 3-month version of the BPDSI-5, as well as with number of BPD traits according to the SCID-5-P and self-reported BPD symptom severity. All associations with measures of quality of life and mental well-being were significant and negative. All subscales of the BPDSI-5-4wk were significantly and positively associated with their respective subscales in the 3-month version of the BPDSI-5, both in the full sample and the BPD group. Additionally, significant positive associations were found between the BPDSI-5-4wk subscales and the corresponding SCID-5-P criteria in the full sample, and for all subscales except affective instability in the BPD group.

3.8. Clinical norms and reliable change

Cut-off scores derived by the three methods (i.e., Euclidian distance, Youden's index, Liu's index) converged for distinguishing the BPD group from the clinical control group, the non-patient control group, and the two control groups combined. The cut-off score to distinguish patients with BPD from clinical controls was 21.02 (sensitivity = 85 %, specificity = 94 %). The cut-off to differentiate the BPD group from non-patient controls was 10.50 (sensitivity = 98 %, specificity = 100 %), and for the combined control group the cut-off was 17.26 (sensitivity = 93 %, specificity = 92 %).

The reliable change criterion based on the test-retest reliability of the BPDSI-5-4wk (ICC = 0.977) was 5.88. This indicates that a change of 5.88 or more in BPDSI-5-4wk total score reflects a significantly relevant change in clinical symptoms that is unlikely to be due to measurement error alone.

To assess sensitivity to change, we tested a linear mixed model using preliminary (unpublished) randomized controlled trial (RCT) data (n =

**Table 5**  
Spearman (partial) correlations.

	Full sample		BPD group
	ρ	Partial ρ <sup>a</sup>	ρ
<b>Correlations with BPDSI-5-4wk total score</b>			
BPDSI-5 (3-month version)	0.97***	0.97***	0.94***
SCID-5-P (number of BPD traits)	0.74***	0.76***	0.40***
Ultra short BPD-C	0.84***	0.85***	0.70***
MHQoL	-0.69***	-0.70***	-0.48***
PMH	-0.69***	-0.70***	-0.46***
CL	-0.58***	-0.56***	-0.29***
<b>BPDSI-5-4wk subscales with BPDSI-5 subscales</b>			
Fear of abandonment	0.98***	0.97***	0.97***
Relationships	0.93***	0.93***	0.90***
Identity disturbance	0.98***	0.98***	0.97***
Impulsivity	0.93***	0.93***	0.88***
(Para)suicide	0.87***	0.86***	0.82***
Affective instability	0.95***	0.95***	0.93***
Emptiness	0.96***	0.96***	0.95***
Anger attacks	0.91***	0.90***	0.91***
Dissociation	0.96***	0.96***	0.94***
<b>BPDSI-5-4wk subscales with SCID-5-P criteria</b>			
Fear of abandonment	0.69***	0.65***	0.54***
Relationships	0.55***	0.52***	0.23*
Identity disturbance	0.76***	0.73***	0.53***
Impulsivity	0.56***	0.52***	0.33**
(Para)suicide	0.64***	0.56***	0.50***
Affective instability	0.65***	0.64***	0.20
Emptiness	0.60***	0.55***	0.23*
Anger attacks	0.63***	0.58***	0.58***
Dissociation	0.62***	0.58***	0.30**

Note. \*p < .05; \*\*p < .01; \*\*\*p < .001. ρ: Spearman's rho; BPDSI: Borderline personality disorder severity index; SCID-5-P: Structured clinical interview for DSM-5 personality disorders; BPD-C: BPD checklist; MHQoL: Mental health quality of life scale; PMH: Positive mental health scale; CL: Cantril's ladder of life. <sup>a</sup>. Partial correlations are controlled for age and educational status.

63 patients with BPD, see [10] for study protocol). Results showed a significant reduction in BPDSI-5-4wk total scores from baseline to post-treatment at 8 months on average across both treatment arms (Cohen's d = -0.48, p < .001), indicating that the BPDSI-5-4wk is sensitive to

change over time.

#### 4. Discussion

This study aimed to address the clinical and research need for a short-term, clinician-administered measure of BPD symptom severity. We evaluated the psychometric properties of the BPDSI-5-4wk, a clinical interview to assess BPD symptom severity over the course of the last 4 weeks. The BPDSI-5-4wk total score demonstrated very high interrater agreement and test-retest reliability, excellent internal consistency, and very good construct (i.e., known-group) validity and criterion-related (i.e., concurrent) validity.

Reliability and validity of the BPDSI-5-4wk are similar to those observed in previous studies evaluating the BPDSI-IV, which has a 3-month evaluation period [5,29]. Regarding reliability, the interrater agreement was excellent and in line with these previous studies. The test-retest reliability, which had not been previously reported for the BPDSI-IV, was also excellent in the current study. Our CFA results additionally supported the assumed DSM-5-based nine-factor structure of the BPDSI-5-4wk, with acceptable to good global fit indices and generally moderate to high factor loadings. These findings are in line with the previously demonstrated preferred nine-factor structure in the BPDSI-IV [5] and provide evidence for the construct validity of the BPDSI-5-4wk, indicating that the subscales capture distinct, meaningful symptom dimensions. The internal consistency of the total score was additionally similar to those observed previously in both the full sample and the BPD group. Internal consistencies of most subscales were similar as well, with somewhat lower internal consistencies in the current study for relationships, (para)suicide, and anger attacks. These variations may be attributed to the shorter assessment period, with less occurrence of more severe or infrequent behaviors (e.g., self-harm through biting or inserting objects into the body, suicide attempts, physically attacking others), which limits variance and could thereby impact the internal consistency. These high zero-response rate patterns were expected, given the nature of some of the items: the BPDSI-5-4wk is designed for detailed assessment of a wide range of BPD symptoms, including those that occur infrequently but carry significant clinical weight. Based on the theoretical and clinical importance of also capturing these less occurring key symptoms of BPD and a lack of substantial improvement in internal consistency when removing these items, we decided to keep all original items for construct coverage. The internal consistency of the relationships subscale was notably lower in participants without a partner relationship, which had not been examined separately in previous validation studies of the BPDSI-IV. A likely explanation for this finding is the reduced number of items contributing to their score, as only four non-partner relationship items were included. The relatively small number of items combined with a smaller sample size in these subgroup analyses, can contribute to lower or less stable estimates of internal consistency and may have limited the subscale's ability to capture these problems reliably. Additionally, difficulties in non-partner relationships may be more heterogeneous, encompassing a broader and more diverse range of interpersonal difficulties that reduce coherence within this subscale. Therefore, scores on this subscale should be interpreted with caution in individuals without a current partner relationship. In addition, we evaluated the internal consistency based on the CFA-based reliability estimates in the full sample as well. These estimates account for the underlying symptom dimension structure and the ordinal nature of item responses, and indicated excellent internal consistency for the total scale and acceptable to good internal consistency for the subscales. These more advanced reliability estimates and the robust CFA model fit suggest that the BPDSI-5-4wk provides a consistent and coherent representation of both overall and specific dimensions of BPD symptom severity, supporting its use for dimensional assessment of specific BPD symptoms in research settings. However, it is important to note that specific subscale scores may not be suitable for interpretation at the individual level in clinical settings, due to their more moderate

reliability. Furthermore, although the CFA model showed good convergence and fit, our relatively small sample size for conducting factor analyses may limit the stability of our results and these findings should therefore be interpreted with caution.

We additionally found comparable indices of validity to those observed in the previous psychometric evaluations of the BPDSI-IV [5,29]. The total score and all subscales of the BPDSI-5-4wk clearly discriminated patients with BPD from clinical and non-patient groups, with effect sizes similar to that of the previously evaluated 3-month BPDSI-IV. Concurrent validity was also similar, with comparable associations between the BPDSI total score and both the number of DSM-5 BPD traits and BPD checklist score, as well as between BPDSI subscales and their corresponding DSM-5 criteria [5,29]. The non-significant correlation between the BPDSI subscale and SCID-5-P criterion of affective instability in the BPD group can be explained by the high proportion of patients with BPD that fulfilled this SCID-5-P criterion in our sample (88 out of 92), leading to a lack of variability which diminishes the possibility to detect a meaningful correlation. The very high correlations between the BPDSI-5-4wk and corresponding BPDSI-5 subscales further support the construct validity of the BPDSI-5-4wk. Given that both versions assess the same items over different time periods, strong associations were expected and indicated that the 4-week version provides a valid measure of BPD symptom severity over a shorter time frame.

Our analysis of preliminary RCT data additionally indicated that the BPDSI-5-4wk is sensitive to change over time. This result is in line with several longitudinal and interventional studies using the BPDSI-IV, which also demonstrated its sensitivity to change [30–32]. The BPDSI-5-4wk also demonstrated excellent test-retest reliability for the total scale and good to excellent reliability for the subscales. This pattern further supports the BPDSI-5-4wk's stability over time, which is particularly relevant for instruments designed to repeatedly monitor symptom change. However, further evaluation of the BPDSI-5-4wk's sensitivity to change in future clinical trials is needed.

The current psychometric evaluation indicates that the BPDSI-5-4wk is a reliable and valid instrument for assessing BPD symptoms over a relatively short time frame of 4 weeks, offering distinct advantages for both clinical practice and research. In clinical settings, this facilitates timely treatment indication of follow-up treatment (e.g., BPD treatment after initial treatment for PTSD or addiction) and frequent evaluation of treatment response, which could help improve clinical trajectories, reduce treatment waitlists, and potentially lower healthcare costs. In research settings, the BPDSI-5-4wk's capacity to capture symptom changes over shorter periods can support studies with relatively brief intervention and follow-up periods, reducing trial durations and minimizing dropout risk. The BPDSI-5-4wk additionally provides detailed insights into (intermittent) changes in BPD symptoms and their responsiveness to treatment phases and sequential or adjunctive treatments.

Importantly, our study also has several limitations. First, for 36 patients with BPD from the Belgian sample, information on comorbid PDs was unavailable as these sections of the SCID-II were not administered, limiting insight into the clinical characteristics of this subgroup. This affects the specificity of our findings concerning the influence of other co-morbid PDs and could affect the known-group validity, as possible differences in PD comorbidity between the BPD and APD groups may have influenced the observed group differences. Second, we did not assess comorbidities other than PDs, such as depression or anxiety. This prevents us from exploring their potential influence on concurrent validity, for instance on associations between the BPDSI-5-4wk and quality of life-related measures. Third, while we included both outpatients and inpatients, we were restricted to recruitment of treatment-seeking patients, who may differ from individuals with BPD in the general population or forensic settings and potentially limits the generalizability of our results. The generalizability is further limited as our sample consisted predominantly of women, which reflects the sex distribution

observed in clinical populations with BPD. Future studies are needed to evaluate whether instruments such as the BPDSI-5-4wk are equally reliable and valid to measure (changes in) severity of BPD across sex or gender.

Although the BPDSI-5-wk is grounded in the categorical DSM-5 framework, future research could explore how the BPDSI-5-4wk corresponds with more dimensional conceptualizations of personality pathology such as introduced in the DSM-5 Alternative Model for Personality Disorders (AMPD). The BPDSI-5-4wk, while grounded in the categorical DSM-5 model, assesses BPD pathology by rating symptom severity across multiple items per symptom domain, which results in dimensional scores reflecting the extent of symptom expression across core BPD features. This fits within the broader dimensional concept of PDs. Instruments like the BPDSI remain essential within the AMPD framework, which, despite the model's dimensional nature, still recognizes a specific BPD presentation characterized by a unique set of traits that warrant targeted intervention. Moreover, existing BPD treatments that have been developed to address the unique symptomatology of BPD require adequate assessment of symptom severity and rely on instruments like the BPDSI-5-4wk to guide clinical decisions. Investigating how structured symptom-based severity measures like the BPDSI-5-4wk align with and could complement measures of personality functioning and maladaptive traits within the AMPD framework could refine treatment indication and monitoring for patients with BPD symptomatology.

In conclusion, the BPDSI-5-4wk demonstrated favorable psychometric properties, making it a reliable and valid instrument to assess BPD symptoms over the course of the past 4 weeks. By enabling more frequent and intermittent assessments of BPD symptom severity, the BPDSI-5-4wk has the potential to both improve clinical care and advance clinical research in BPD. Future research, including clinical trials implementing the BPDSI-5-4wk, may further show its sensitivity to change and could explore its utility in integrating categorical and dimensional assessment models.

### Funding sources

This work was supported by the Netherlands Brain Foundation (grant number: DR-2019-00319).

### CRedit authorship contribution statement

**Shanna van Trigt:** Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mariana Mendoza-Alvarez:** Writing – review & editing, Investigation. **Tanja van der Zweerde:** Writing – review & editing. **Livia De Picker:** Writing – review & editing. **Annemieke van Straten:** Writing – review & editing, Supervision, Funding acquisition. **Arnoud Arntz:** Writing – review & editing, Resources, Methodology. **Hein J.F. van Marle:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: None.

### Acknowledgements

We would like to thank Eva Fris, Annelies de Kemp, Irini Kanellou, and all students that contributed to recruitment and data collection.

### Appendix A. Supplementary data

Supplementary data to this article and the BPDSI-5-4wk interview

(English) can be found online at <https://doi.org/10.1016/j.comppsych.2025.152634>. The interview is freely available for non-commercial use. For translations, please contact [a.r.arntz@uva.nl](mailto:a.r.arntz@uva.nl).

### References

- [1] American Psychiatric Association. *DSM-5: Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Press; 2013.
- [2] Gunderson JG, Herpertz SC, Skodol AE, Torgersen S, Zanarini MC. Borderline personality disorder. *Nat Rev Dis Primers* 2018;4:18029. <https://doi.org/10.1038/nrdp.2018.29>.
- [3] Juurlink TT, Lamers F, van Marle HJF, Michon H, van Busschbach JT, Beekman AT, et al. Employment in personality disorders and the effectiveness of individual placement and support: outcomes from a secondary data analysis. *J Occup Rehabil* 2020;30(2):255–62. <https://doi.org/10.1007/s10926-019-09868-9>.
- [4] Zanarini MC, Stanley B, Black DW, Markowitz JC, Goodman M, Pilkonis P, et al. Methodological considerations for treatment trials for persons with borderline personality disorder. *An Clin Psychiatry* 2010;22(2):75–83.
- [5] Giesen-Bloo JH, Wachters LM, Schouten E, Arntz A. The borderline personality disorder severity index-IV: psychometric evaluation and dimensional structure. *Pers Individ Dif* 2010;49(2):136–41. <https://doi.org/10.1016/j.paid.2010.03.023>.
- [6] Arntz A, van den Hoorn M, Cornelis J, Verheul R, van den Bosch WM, de Bie AJ. Reliability and validity of the borderline personality disorder severity index. *J Pers Disord* 2003;17(1):45–59. <https://doi.org/10.1521/pedi.17.1.45.24053>.
- [7] Wibbelink CJM, Arntz A, Grasman R, Sinnaeve R, Boog M, Bremer OMC, et al. Towards optimal treatment selection for borderline personality disorder patients (BOOTS): a study protocol for a multicenter randomized clinical trial comparing schema therapy and dialectical behavior therapy. *BMC Psychiatry* 2022;22(1):89. <https://doi.org/10.1186/s12888-021-03670-9>.
- [8] Black DW, Zanarini MC, Romine A, Shaw M, Allen J, Schulz SC. Comparison of low and moderate dosages of extended-release quetiapine in borderline personality disorder: a randomized, double-blind, placebo-controlled trial. *Am J Psychiatry* 2014;171(11):1174–82. <https://doi.org/10.1176/appi.ajp.2014.13101348>.
- [9] van den End A, Dekker J, Beekman ATF, Aarts I, Snoek A, Blankers M, et al. Clinical efficacy and cost-effectiveness of imagery Rescripting only compared to imagery Rescripting and Schema therapy in adult patients with PTSD and comorbid cluster C personality disorder: study Design of a Randomized Controlled Trial. *Front Psych* 2021;12:633614. <https://doi.org/10.3389/fpsy.2021.633614>.
- [10] van Trigt S, van der Zweerde T, van Someren EJW, van Straten A, van Marle HJF. Guided internet-based cognitive behavioral therapy for insomnia in patients with borderline personality disorder: study protocol for a randomized controlled trial. *Internet Interv* 2022;29:100563. <https://doi.org/10.1016/j.invent.2022.100563>.
- [11] Hofmans J, Vergauwe J, De Clercq B. Contemporary methodological considerations for key issues in research on personality disorder development. *Personal Disord* 2023;14(1):83–92. <https://doi.org/10.1037/per0000597>.
- [12] First MB, Williams JBW, Smith Benjamin L, Spitzer RL. *Structured clinical interview for DSM-5® personality disorders (SCID-5-PD)*. APA Publishing; 2016.
- [13] Mendoza Alvarez M, Verbraecken J, Claes L, Vandekerckhove M, De Picker L. Clinical predictors of insomnia in borderline personality disorder: a polysomnographic and subjective examination. *Borderline Personal Disord Emot Dysregul* 2025. <https://doi.org/10.1186/s40479-024-00277-w>.
- [14] Weertman A, Arntz A, Dreessen L, van Velzen C, Vertommen S. Short-interval test-retest interrater reliability of the Dutch version of the structured clinical interview for DSM-IV personality disorders (SCID-II). *J Pers Disord* 2003;17(6):562–7. <https://doi.org/10.1521/pedi.17.6.562.25359>.
- [15] Lobbstaël J, Leurgans M, Arntz A. Inter-rater reliability of the structured clinical interview for DSM-IV Axis I disorders (SCID I) and Axis II disorders (SCID II). *Clin Psychol Psychother* 2011;18(1):75–9. <https://doi.org/10.1002/cpp.693>.
- [16] Wibbelink CJM, Kamphuis JH, Efting M, Sinnaeve R, Boog M, Dek ECP, et al. Efficient screening and outcome assessment for borderline personality disorder: A psychometric evaluation of the uBPDC. 2025.
- [17] Bloo J, Arntz A, Schouten E. The borderline personality disorder checklist: psychometric evaluation and factorial structure in clinical and nonclinical samples. *Roczniki Psychologiczne* 2017;20(2):311–36. <https://doi.org/10.18290/rpsych.2017.20.2.3en>.
- [18] Boog M, Goudriaan AE, Wetering B, Franken IHA, Arntz A. Schema therapy for patients with borderline personality disorder and comorbid alcohol dependence: a multiple-baseline case series design study. *Clin Psychol Psychother* 2022. <https://doi.org/10.1002/cpp.2803>.
- [19] van Krugten FCW, Busschbach JJV, Versteegh MM, Hakkaart-van Roijen L, Brouwer WBF. The mental health quality of life questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Qual Life Res* 2022;31(2):633–43. <https://doi.org/10.1007/s11136-021-02935-w>.
- [20] Lukat J, Margraf J, Lutz R, Der Veld WM, Becker ES. Psychometric properties of the positive mental health scale (PMH-scale). *BMC Psychology* 2016;4(8):1–14. <https://doi.org/10.1186/s40359-016-0111-x>.
- [21] Kilpatrick FP, Cantril H. Self-anchoring scaling - a measure of individuals unique reality worlds. *J Individ Psychol* 1960;16(2):158–73.
- [22] McIntosh CN. Report on the construct validity of the temporal satisfaction with life scale. *Soc Indic Res* 2001;54(1):37–56. <https://doi.org/10.1023/A:1007264829700>.

- [23] R Core Team. R: A language and environment for statistical computing. 4.3.2. ed. Vienna, Austria: R Foundation for Statistical Computing; 2021.
- [24] Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol* 2006;163(7):670–5. <https://doi.org/10.1093/aje/kwj063>.
- [25] Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3(1):32–5. <https://doi.org/10.1002/1097-0142>.
- [26] Liu BC. Quality of life: concept, measure and results. *Am J Econ Soc* 1975;34(1):1–13.
- [27] Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;59(1):12–9. <https://doi.org/10.1037//0022-006x.59.1.12>.
- [28] Blampied NM. Reliable change and the reliable change index: still useful after all these years? *Cognit Behav Therapist* 2022;15. <https://doi.org/10.1017/s1754470x22000484>.
- [29] Kroger C, Vonau M, Kliem S, Roepke S, Kosfelder J, Arntz A. Psychometric properties of the German version of the borderline personality disorder severity index—version IV. *Psychopathology* 2013;46(6):396–403. <https://doi.org/10.1159/000345404>.
- [30] Arntz A, Jacob GA, Lee CW, Brand-de Wilde OM, Fassbinder E, Harper RP, et al. Effectiveness of predominantly group Schema therapy and combined individual and group Schema therapy for borderline personality disorder: a randomized clinical trial. *JAMA Psychiatry* 2022;79(4):287–99. <https://doi.org/10.1001/jamapsychiatry.2022.0010>.
- [31] Soderholm JJ, Socada JL, Ekelund J, Isometsa E. How changes in depression severity and borderline personality disorder intensity are linked - a cohort study of depressed patients with and without borderline personality disorder. *Borderline Personal Disord Emot Dysregul.* 2024;11(1):3. <https://doi.org/10.1186/s40479-024-00247-2>.
- [32] Giesen-Bloo J, van Dyck R, Spinhoven P, van Tilburg W, Dirksen C, van Asselt T, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Arch Gen Psychiatry* 2006;63(6):649–58. <https://doi.org/10.1001/archpsyc.63.6.649>.