The American Psychiatric Association Practice Guideline for the Treatment of Patients with Borderline Personality Disorder

Appendices

Table of Contents

Appendix A. Clinical Questions

The following key questions formed the basis of the systematic review:

1. In patients with borderline personality disorder, what is the efficacy, effectiveness, and risk of harms of various pharmacological and nonpharmacological therapies and different service delivery approaches?
    a. Are there differences in efficacy, effectiveness, or risk of harms regarding different subgroups based on age, gender, race/ethnicity, or genotypes?
2. In patients with borderline personality disorder, what is the comparative efficacy, effectiveness, and risk of harms of various pharmacological and nonpharmacological therapies and different service delivery approaches?
    a. Are there any differences in efficacy, effectiveness, or risk of harms regarding different subgroups based on age, gender, race/ethnicity, or genotypes?

Figure A—1 presents the analytic framework for our key questions.

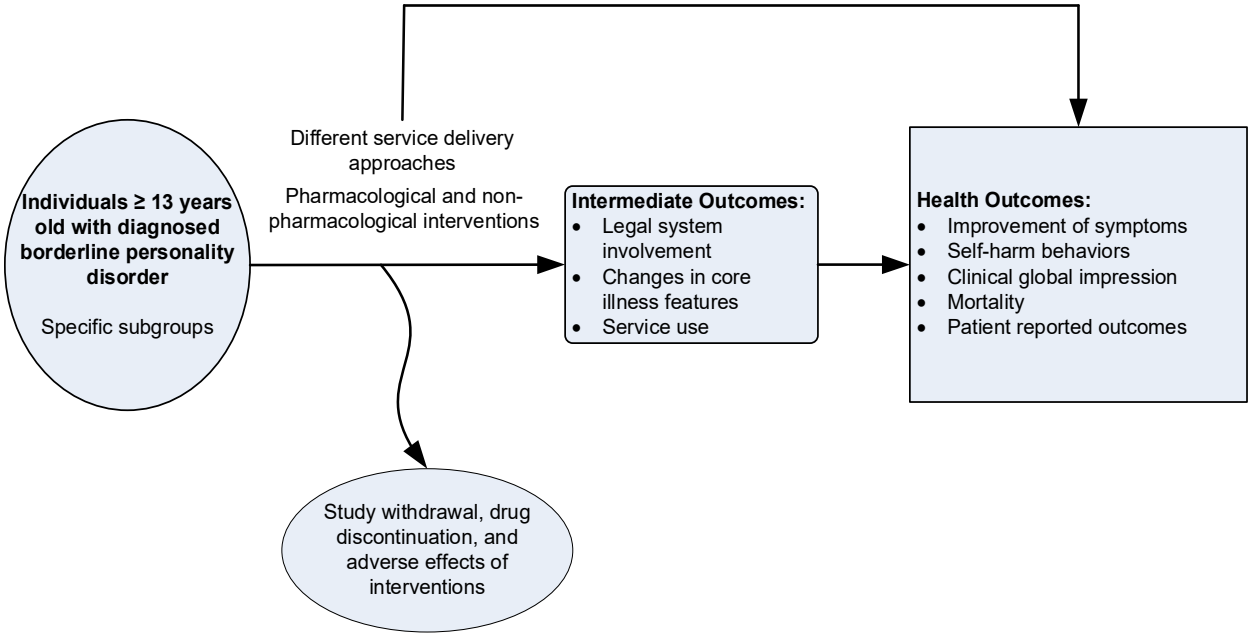


Figure A—1. Analytic framework.

Appendix B. Search Strategies, Study Selection, and Search Results

The methods for this systematic review follow the Agency for Healthcare Quality and Research (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (available at http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm) and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist (Moher et al. 2015). The final protocol of this review was registered on PROSPERO (Registration #: CRD42020194098). All methods and analyses were determined a priori.

Literature Searches

We built on the original search strategy on interventions for borderline personality disorder (BPD) conducted on June 7, 2018, by Doctor Evidence but made searches more specific. To ensure optimal recall, we ensured that the revised search strategy still detected all studies that met inclusion criteria of the original search. We searched MEDLINE, EMBASE, the Cochrane Library, and PsycINFO from January 1, 2018, to June 15, 2020. We ran overlapping update searches of MEDLINE and PsycINFO in April and September 2021. Our search strategies used a variety of terms, medical subject headings (MeSH), and major headings, and were limited to English language and human-only studies.

To minimize retrieval bias, we manually searched reference lists of landmark studies and background articles on this topic for relevant citations that electronic searches might have missed.

Doctor Evidence Original Search Strategy

Search Date: June 7, 2018

Table B—1. PubMed search strategy for borderline personality disorder.

Search ID#	Query	Results
#1	("Borderline Personality Disorder"[Mesh]) OR (borderline [tiab] AND personality [tiab])	8962
#2	("animals"[MeSH Terms] OR animal [tiab] OR animals [tiab] OR rat [tiab] OR rats [tiab] OR mouse [tiab] OR mice [tiab] OR rodent [tiab] OR rodents [tiab]) NOT ("humans"[MeSH Terms] OR humans [tiab] OR human [tiab])	4419530
#3	#1 NOT #2	8957
	Limit to English	7983

Table B—2. EMBASE search strategy for borderline personality disorder.

Search	Query	Results
#1	exp *borderline state/ or (borderline and personality).ti. or (borderline and personality).ab.	11073
#2	limit #1 to (article or article in press or conference paper)	7571
#3	#2 not ((exp animal/ or nonhuman/) not exp human/)	7564

#4	#2 not ((animal or animals or rat or rats or mouse or mice or rodent or rodents) not (humans or human)).ti,ab.	7548
#5	#3 or #4	7569
#6	limit #5 to yr="1883 - 2002"	2765
#7	limit #5 to yr="2002 - Current"	4929
#8	remove duplicates from #6	2740
#9	remove duplicates from #7	4739
#10	#8 or #9	7337
#11	limit #10 to english language	6356

Table B—3. Cochrane Library search strategy for borderline personality disorder.

Search	Query	Results
#1	MeSH descriptor: [Borderline Personality Disorder] explode all trees	390
#2	borderline and personality:ti,ab,kw (Word variations have been searched)	684
#3	#1 or #2	684
#4	#3 not (pubmed or embase):an	145 in trials 6 in Cochrane reviews 9 in other reviews

Table B—4. PsycINFO search strategy for borderline personality disorder.

Search	Query	Limiters/Expanders	Results
S1	MM "Borderline Personality Disorder"		5,220
S2	DE "Borderline Personality Disorder"		7,857
S3	MA "borderline personality disorder"		4,192
S4	TI "borderline personality" OR AB "borderline personality" OR SU "borderline personality" OR KW "borderline personality"		11,400
S5	S1 OR S2 OR S3 OR S4		11,400
S6	(MM "Animals" OR DE "Animals" OR DE "Vertebrates" OR DE "Amphibia" OR DE "Birds" OR DE "Fishes" OR DE "Mammals" OR DE "Pigs" OR DE "Reptiles" OR DE "Rats" OR DE "Rodents" OR DE "Mice")		329,022
S7	TI "animals" OR TI "animal" OR TI "mouse" OR TI "mice" OR TI "rodent" OR TI "rodents" OR TI "rat" OR TI "rats" OR SU "animals" OR SU		426,155

Search	Query	Limiters/Expanders	Results
	"animal" OR SU "mouse" OR SU "mice" OR SU "rodent" OR SU "rodents" OR SU "rat" OR SU "rats" OR KW "animals" OR KW "animal" OR KW "mouse" OR KW "mice" OR KW "rodent" OR KW "rodents" OR KW "rat" OR KW "rats" OR AB "animals" OR AB "animal" OR AB "mouse" OR AB "mice" OR AB "rodent" OR AB "rodents" OR AB "rat" OR AB "rats"		
S8		Limiters - Population Group: Animal	385,743
S9	S6 OR S7 OR S8		459,805
S10		Limiters - Population Group: Human	3,780,890
S11	TI "humans" OR TI "human" OR AB "humans" OR AB "human" OR SU "humans" OR SU "human" OR KW "humans" OR KW "human"		1,585,426
S12	S10 OR S11		3,888,530
S13	S9 NOT S12		310,376
S14	S5 NOT S13		11,398
S15		Limiters - Publication Type: All Journals	3,518,961
S16	S14 AND S15		9,386
S17	LA English		4,207,720
S18	S16 AND S17		8,116

RTI Updated Search Strategy

Search Date: June 15, 2020

Table B—5. PubMed search strategy for borderline personality disorder.

Search	Query	Results
#1	"Borderline Personality Disorder"[Mesh] OR "Borderline Disorder"[ti] OR "Borderline Personality Disorder"[tiab] OR "borderline-patient"[ti] OR "borderline patient"[ti] OR "borderline-patients"[ti] OR "borderline patients"[ti]	8,693
#2	#1 AND ("2018/01/01"[Date - Publication] : "3000"[Date - Publication])	1,202
#3	#2 AND English[lang]	1,161

Table B—6. EMBASE search strategy for borderline personality disorder.

Search	Query	Results
#1	"Borderline Personality Disorder"[Mesh] OR "Borderline Disorder"[ti] OR "Borderline Personality Disorder"[tiab] OR "borderline-patient"[ti] OR "borderline patient"[ti] OR "borderline-patients"[ti] OR "borderline patients"[ti]	8,693
#2	#1 AND ("2018/01/01"[Date - Publication] : "3000"[Date - Publication])	1,202
#3	#2 AND English[lang]	1,161

Table B—7. Cochrane Library search strategy for borderline personality disorder.

Search	Query	Results
#1	"Borderline Personality Disorder"[Mesh] OR "Borderline Disorder"[ti] OR "Borderline Personality Disorder"[tiab] OR "borderline-patient"[ti] OR "borderline patient"[ti] OR "borderline-patients"[ti] OR "borderline patients"[ti]	8,693
#2	#1 AND ("2018/01/01"[Date - Publication] : "3000"[Date - Publication])	1,202
#3	#2 AND English[lang]	1,161

Table B—8. PsycINFO (via ProQuest) search strategy for borderline personality disorder.

Search	Query	Results
S1	if("Borderline Personality Disorder") OR mjsub("Borderline Personality Disorder") OR mainsubject("Borderline Personality Disorder") OR ti("Borderline Personality Disorder" OR "Borderline Disorder" OR "borderline-patient" OR "borderline patient" OR "borderline-patients" OR "borderline patients") OR ab("Borderline Personality Disorder") *Additional limits - Date: After January 01 2018; Language: English*	986

Search Date: April 6, 2021

Table B—9. PubMed search strategy for borderline personality disorder.

Search	Query	Results
#1	"Borderline Personality Disorder"[Mesh] OR "Borderline Disorder*"[ti] OR "Borderline Personality Disorder*"[tiab] OR "borderline patient"[ti] OR "borderline patients"[ti]	9,260
#2	#1 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])	9,258
#3	(#2) AND (("2020"[Date - Publication] : "3000"[Date - Publication])) Filters: English	744

Table B—10. PsycINFO (via ProQuest) search strategy for borderline personality disorder.

Search	Query	Results
S1	DE "Borderline Personality Disorder"	8,991
S2	borderline W1 (disorder# OR patient#)	13,511
S3	S1 OR S2	13,511
S4	S3 (Limiters – Publication Year 2020 – 2021; Language: English)	510

Search Date: September 24, 2021

Table B—11. PubMed search strategy for borderline personality disorder.

Search	Query	Results
#1	"Borderline Personality Disorder"[Mesh] OR "Borderline Disorder*"[ti] OR "Borderline Personality Disorder*"[tiab] OR "borderline patient"[ti] OR "borderline patients"[ti]	9,488
#2	#1 NOT ("Animals"[Mesh] NOT "Humans"[Mesh])	9,486
#3	(#2) AND (("2020"[Date - Publication] : "3000"[Date - Publication])) Filters: English	949

Table B—12. PsycINFO (via ProQuest) search strategy for borderline personality disorder.

Search	Query	Results
S1	DE "Borderline Personality Disorder"	9,216
S2	borderline W1 (disorder# OR patient#)	13,784
S3	S1 OR S2	13,784
S4	S3 (Limiters – Publication Year 2020 – 2021; Language: English)	749

Criteria for Inclusion/Exclusion of Studies in the Review

The criteria for inclusion and exclusion of studies are designed to identify research that can answer the key questions. The criteria are based on the population, intervention/exposure, comparator, outcomes, time frames, country and clinical settings, and study design (PICOTS).

Table B—13. Inclusion and exclusion criteria.

Criteria	Include	Exclude

Participants/population	• Age ≥13
	• Diagnosed with BPD as defined by DSM-IV, DSM-IV-TR, DSM-5 (Section II or Section III), or ICD-10
	• For mixed population studies, BPD must account for ≥75% of the total population
	• Subgroups of interest
	  ○ Co-occurring mental disorder
	  ○ Age
	  ○ Gender
	  ○ Race/ ethnicity
	  ○ Genotypes (related to treatment selection, treatment response or adverse effects)
		• Age <13
		• Individuals with borderline traits without a specific diagnosis
		• Diagnosed with BPD as defined by DSM-III-R
		• Studies in which the primary research focus is a different diagnosis with co-occurring BPD in a subset (<75% of the total population)

Intervention(s)/exposure(s)	• Yoga
	• Exercise
	• Peer-support interventions
	• Psychosocial support
	• Safety planning
	• Service delivery approaches:
	  ○ Stepped-care
	  ○ Collaborative care
	  ○ Measurement-based care
	  ○ Treatment setting comparisons
	  ○ Face-to-face sessions
	  ○ Group sessions
	  ○ Online programs
	  ○ Therapeutic community
	  ○ Video
	• Progressive Muscle Relaxation
	• Somatic therapies:
	  ○ Electroconvulsive therapy (ECT)
	  ○ Repetitive transcranial magnetic stimulation (rTMS)
	  ○ Transcranial alternating current stimulation (tACS)
		• Complementary/alternative treatments not listed for inclusion
		• Somatic therapies:
		  ○ Bioenergetic analysis
		  ○ Body psychotherapy
		  ○ Core energetics
		  ○ Hakomi
		  ○ Somatic experiencing
		• Pharmacotherapies:
		  ○ Acetazolamide
		  ○ Ethosuximide
		  ○ Felbamate
		  ○ Fosphenytoin
		  ○ Lacosamide
		  ○ Methsuximide
		  ○ Pentobarbital
		  ○ Perampanel
		  ○ Primidone
		  ○ Rufinamide
		  ○ Droperidol
		  ○ Nalmefene
		  ○ Butabarbital
		  ○ Secobarbital

Criteria	Include	Exclude
	○ Transcranial direct current stimulation (tDCS) ○ Transcranial magnetic stimulation (TMS) • Pharmacotherapies: ○ Anticonvulsant "mood stabilizers": – Carbamazepine – Divalproex sodium – Gabapentin – Lamotrigine – Levetiracetam – Oxcarbazepine – Phenytoin – Pregabalin – Tiagabine – Topiramate – Valproate – Valproic acid – Vigabatrin – Zonisamide ○ Antidepressants: – Amitriptyline – Amoxapine – Bupropion – Citalopram – Clomipramine – Desipramine – Desvenlafaxine – Doxepin – Duloxetine – Escitalopram – Fluoxetine – Fluvoxamine – Impiramine – Isocarboxazid – Maprotiline – Mirtazapine – Milnacipran – Nefazodone – Nortriptyline – Paroxetine – Phenelzine – Protriptyline – Sertraline	

Criteria	Include	Exclude
	– Selegiline – Tranylcypromine – Trazodone – Trimipramine – Venlafaxine – Vilazodone – Vortioxetine o Antipsychotics: – Aripiprazole – Asenapine – Chlorpromazine – Clozapine – Fluphenazine – Haloperidol – Iloperidone – Loxapine – Lurasidone – Olanzapine – Paliperidone – Perphenazine – Pimozide – Prochlorperazine – Quetiapine – Risperidone – Thioridazine – Thiothixene – Trifluoperazine – Ziprasidone o Benzodiazepines: – Alprazolam – Clobazam – Clonazepam – Clorazepate – Chlordiazepoxide diazepam – Estazolam – Flurazepam – Lorazepam – Midazolam – Oxazepam – Quazepam – Temazepam – Triazolam o Opioid agonists and antagonists: – Buprenorphine	

Criteria	Include	Exclude
	– Naloxone – Naltrexone o Sedative-hypnotic medications: – Eszopiclone – Ramelteon – Suvorexant – Tasimelteon – Zaleplon – Zolpidem – Melatonin o Other pharmacotherapies: – Clonidine – Lithium – Prazocin • Psychotherapies: o Acceptance and Commitment Therapy (ACT) o Client-Centered Therapy o Cognitive Analytic Therapy (CAT) o Cognitive Behavioral Therapy (CBT) o Cognitive Rehabilitation o Cognitive Therapy (CT) o Comprehensive Validation Therapy o Dialectical Behavior Therapy (DBT) o Dual-Focused Schema Therapy o Dynamic Deconstructive Psychotherapy (DDP) o Emotion Regulation Group Intervention o Emotion Regulation Training (ERT) o Good Psychiatric Management (GPM) o Group Analytic Psychotherapy o Humanistic and Integrative Psychotherapy o Individual Psychotherapy o Interpersonal Group Psychotherapy o Interpersonal Psychotherapy (IPP) o Interpersonal Therapy (IPT) o Manual-Assisted Cognitive Therapy (MACT) o Mentalization-Based Therapy (MBT) o Mindfulness-Based Cognitive Therapy (MBCT)	

Criteria	Include	Exclude
	○ Motive-Oriented Therapeutic Relationship (MOTR) ○ Nidotherapy ○ Problem-Solving Therapy ○ Psychoanalytic Therapy (Psychoanalysis) ○ Psychodynamic Interpersonal Therapy (PIT) ○ Psychodynamic Therapy ○ Psychodynamic/Psychoanalytic Psychotherapy ○ Psychoeducation ○ Psychotherapy Focused on Psychic Representation ○ Rogerian Supportive Therapy ○ Schema-Focused Cognitive Therapy ○ Schema-Focused Therapy ○ Schema-Focused Psychotherapy (SFP) ○ Sequential Brief Adlerian Psychodynamic Psychotherapy ○ Supervised Team Management ○ Supportive Therapy ○ System-Based Psychotherapy ○ Systemic Therapy ○ Systems Training for Emotional Predictability and Problem Solving (STEPPS) ○ Transference-Focused Psychotherapy (TFP)	
Comparator(s)/control	• Interventions listed above for inclusion • Placebo • Treatment as usual • Wait-list control • Community treatment by experts • General psychiatric management • Standard group treatment • Standard psychiatric care • Structured clinical management	• Interventions listed as excluded above for interventions/exposures
Outcomes	**Pre-specified outcomes and outcome measures** A. BPD symptoms/diagnostic criteria 1. Frantic efforts to avoid real or imaginary abandonment	• Outcomes not listed, imaging markers, physiological markers, and biomarkers • Outcomes that were not pre-specified, e.g., during

Criteria	Include	Exclude
	2. Pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation a. Inventory of Interpersonal Problems (IIP) b. Distorted self-image 3. Identity disturbances: markedly and persistent unstable self-image or sense of self a. Distorted self-image 4. Impulsivity a. Impulsivity b. Impulsive/behavioral c. Risk taking behaviors d. Lack of restraint e. Barratt Impulsiveness Scale (BIS-11) f. Multi-Impulsivity Scale (MIS) 5. Recurrent suicidal behavior, gestures or threats; or self-mutilating behavior a. Nonsuicidal self-injury b. Suicide attempts c. Suicide d. Suicidal ideation e. Self-destructive behavior f. Beck Scale for Suicide Ideation (BSS) g. Self-Harm Behavior Survey h. Suicidal Behaviors Questionnaire (SBQ) and SBQ-R i. Parasuicide History Interview (PHI) j. Borderline Personality Disorder Severity Index (BPDSI) Parasuicidality Subscale k. Columbia Suicide Severity Rating Scale (C-SSRS) l. Deliberate Self-Harm Inventory (DSHI)	post-hoc, exploratory analyses

Criteria	Include	Exclude
	m. Self-Injurious Thoughts and Behaviors Interview-Self-Report 6. Affective instability, due to a marked reactivity of mood a. Irritability b. Mood swings c. Difficulties in Emotion Regulation Scale (DERS) d. Affective dysregulation 7. Chronic feelings of emptiness 8. Inappropriate intense anger or difficulty controlling anger a. Aggression b. Anger c. Hostility d. Aggressive behavior e. Antisocial behavior f. Spielberger State-Trait Anger Expression Inventory (STAXI) g. Spielberger State-Trait Anger Scale (STAS) h. Acting Out Scale (AOS) i. Aggression Questionnaire (AQ) j. Anger, Irritability, and Assault Questionnaire (AIAQ) k. Overt Aggression Scale (OAS) l. Buss Durkee Hostility Inventory (BDHI) 9. Transient, stress-related paranoid ideation, or severe dissociative symptoms a. Dissociation B. Scales for BPD 1. Borderline Personality Disorder Severity Index (BPDSI) 2. Zanarini Rating Scale (ZAN-BPD) C. Other symptoms commonly found in individuals with BPD, but not part of the diagnostic criteria 1. Depression and Anxiety	

Criteria	Include	Exclude
	a. Spielberger State-Trait Anxiety Inventory (STAI) b. Symptom Checklist-90 (SCL-90) c. Beck Anxiety Inventory (BAI) d. Beck Depression Inventory (BDI) e. Beck Hopelessness Scale (BHS) f. Hamilton Rating Scale for Anxiety (Ham-A) g. Hamilton Rating Scale for Depression (Ham-D) h. Hospital Anxiety and Depression Scale (HADS) i. Montgomery-Åsberg Depression Rating Scale (MADRS) j. Patient Health Questionnaire (PHQ-9) k. Brief Symptom Inventory (BSI) l. Generalized Anxiety Disorder 7-item scale (GAD-7) m. Patient Health Questionnaire-Adolescent n. Patient Health Questionnaire: Somatic, Anxiety, and Depressive Symptoms D. Functioning Scales 1. Global Adjustment Scale 2. Global Assessment of Functioning (GAF) 3. QOL 4. Global Social Adjustment (GSA) 5. Global Severity Index (GSI) 6. Number of years with employment 7. Social Adjustment Scale (SAS) 8. Social and Occupational Functioning Assessment Scale 9. Social Functioning Questionnaire (SFQ) 10. Social History Interview (SHI) 11. Social Problem-Solving Inventory	

Criteria	Include	Exclude
	12. World Health Organization - Disability Assessment Schedule (WHO-DAS) E. Adverse Events (AEs) 1. Rate of any AEs 2. Overall serious treatment-related adverse event rate 3. Specific serious treatment-related adverse events 4. Study withdrawal due to AE 5. Study withdrawal for any reason	
Timing	Treatment duration ≥8 weeks	Treatment duration <8 weeks
Setting/context	Very high Human Development Index (HDI) Countries*	All other countries
Study design	• RCTs phase 2 | 3 | 4 • Nonrandomized clinical trials (N≥50): o Phase 1 | 2 | 3 | 4 • Observational studies, comparative (N≥50) o Cross-sectional o Prospective cohort o Retrospective cohort o Nonconcurrent cohort o Case-control Pooled analyses of controlled studies	• Single-arm dose-finding trials • Observational, noncomparative • Case reports/series • Prognostic course/factor studies • Modeling studies • Pre-clinical • Narrative reviews • Systematic reviews/meta-analyses (will be used for hand searches)

Note. *Very High HDI Countries: Andorra, Argentina, Australia, Austria, Bahamas, Bahrain, Barbados, Belarus, Belgium, Brunei Darussalam, Bulgaria, Canada, Chile, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong, China (SAR), Hungary, Iceland, Ireland, Israel, Italy, Japan, Kazakhstan, Korea (Republic of), Kuwait, Latvia, Liechtenstein, Lithuania, Luxembourg, Malaysia, Malta, Montenegro, Netherlands, New Zealand, Norway, Oman, Poland, Portugal, Qatar, Romania, Russian Federation, Saudi Arabia, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, Taiwan**, United Arab Emirates, United Kingdom, United States, Uruguay.

** The United Nations does not recognize Taiwan (i.e., Republic of China) as a sovereign state and does not include it in the HDI report. However, Taiwan's government calculated its HDI to be 0.885, based on 2014 data and using the same methodology as the United Nations. This HDI value would place Taiwan among countries in the "very high" human development category and will be included in this report.

Abbreviations: BPD, borderline personality disorder; KQ, key question; N, sample size; NA, not applicable; RCT, randomized controlled trial.

Literature Review, Data Abstraction, and Data Management

To ensure accuracy, two reviewers independently reviewed all titles, abstracts, and full-text articles. We used Distiller SR, an online tool to conduct systematic reviews, to screen the literature (DistillerSR, Evidence Partners, Ottawa, Canada)**.** We resolved discrepancies by consensus or by involving a third, senior reviewer.

All results at both title/abstract and full-text review stages were tracked in an EndNote® bibliographic database (Thomson Reuters, New York, NY). Appendix C presents the list of studies excluded (with reasons) at the full-text level.

We designed, pilot tested, and used a structured data abstraction form in DistillerSR to ensure consistency of data abstraction. We abstracted data into categories that included (but were not limited to) the following: study design, eligibility criteria, intervention, methods of outcome assessment, population characteristics, sample size, attrition, results, and adverse event incidence. A second team member verified abstracted study data for accuracy and completeness.

Assessment of Risk of Bias of Individual Studies

To assess the risk of bias of studies, we used the Risk Of Bias In Non-randomised Studies of Interventions (ROBINS-I) (Sterne et al. 2016) for nonrandomized controlled studies and for randomized controlled trials (RCTs), we used the Cochrane Risk of Bias 2 tool. Two independent reviewers assessed the risk of bias at the study level and also considered rating bias at an outcome level if methodological limitations might affect different outcomes in a different way (e.g., lack of blinding might increase the risk of bias for quality of life but not for overall mortality). We assigned a “high risk of bias” rating to studies that had very serious limitations in design or conduct which might invalidate findings regarding all or individual outcomes. Disagreements between the two reviewers were resolved by discussion and consensus or by consulting a third member of the team. Risk of bias diagrams were generated using the Risk-of-bias VISualization (robvis) tool (McGuinness and Higgins 2020; see Appendix E).

Data Synthesis

We summarized all included studies in narrative form and in summary tables that tabulate the important features of the study populations, design, intervention, outcomes, setting, country, and results. If we found three or more similar studies addressing an outcome of interest, we considered quantitative analysis (i.e., meta-analysis) if studies were similar (in population, interventions, comparators, and outcomes). For all analyses, we used random-effects models (restricted maximum likelihood random effects) to estimate pooled effects. To determine whether quantitative analyses are appropriate, we assessed the clinical and methodological heterogeneity of the studies under consideration following established guidance (Gartlehner et al. 2012). If we conducted meta-analyses, we assessed statistical heterogeneity in effects between studies by calculating the chi-squared statistic and the I^2 statistic (the proportion of variation in study estimates attributable to heterogeneity). We examined potential sources of heterogeneity using sensitivity analyses. When quantitative analyses were not appropriate (e.g., due to heterogeneity, insufficient numbers of similar studies, or insufficiency or variation in outcome reporting), we synthesized the data qualitatively.

Grading the Certainty of Evidence for Major Comparisons and Outcomes

We graded the certainty of evidence of relevant outcomes based on current GRADE (Grading of Recommendations Assessment, Development and Evaluation) guidance (Balshem et al. 2011). Developed to grade the overall certainty of a body of evidence, this approach incorporates five key domains: (1) risk of bias, (2) inconsistency, (3) indirectness, (4) imprecision of the evidence, and (5) reporting bias. It also considers other optional domains that may be relevant for some scenarios. These included plausible confounding that would decrease the observed effect and strength of association (i.e., magnitude of effect) or factors that would increase the strength of association (i.e., dose-response effect). Two reviewers assessed each domain for each selected outcome and resolved differences by consensus discussion. We documented all decisions regarding up- or down-grading the certainty of evidence to ensure transparency. We used GradePro to develop summary of findings tables for the guideline panel.

Table B—14 describes the grades of certainty of evidence, which reflect the certainty of the body of evidence regarding a specific outcome.

Table B—14. Definitions of the grades of certainty of evidence.

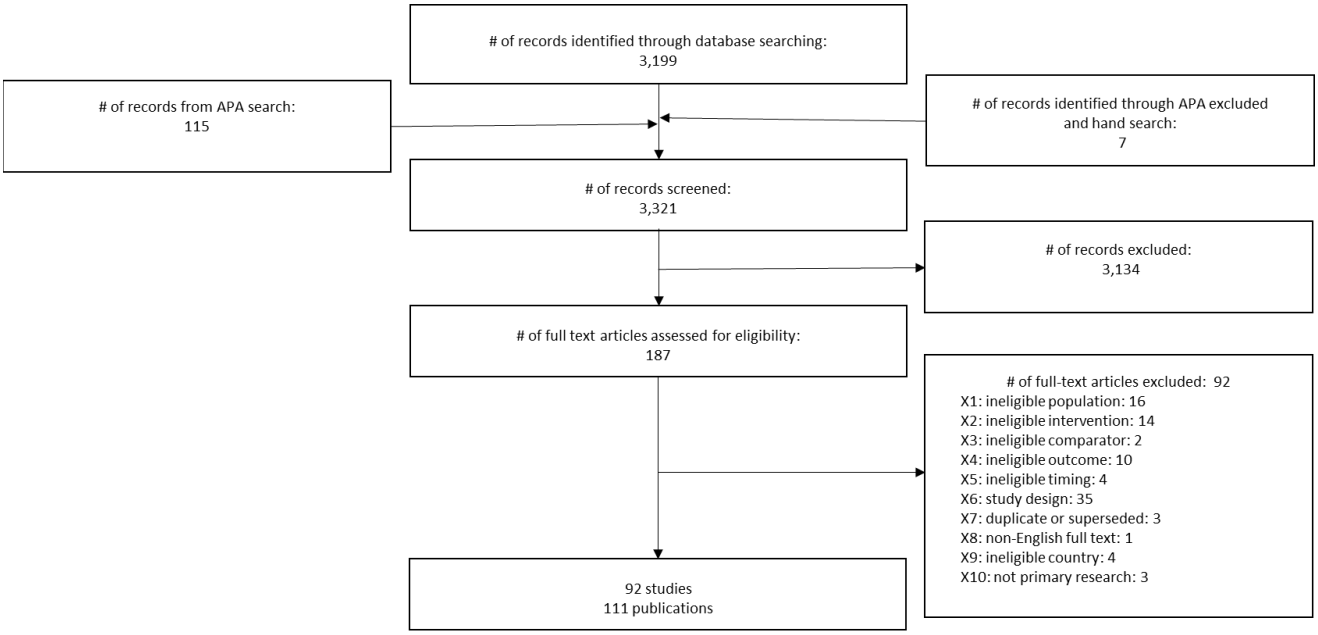
Grade	Definition
High	We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable (i.e., another study would not change the conclusions).
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Very Low	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. The body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Source. Adapted from Balshem et al. 2011.

Results of Literature Search and Literature Screening

We screened 3,321 titles and abstracts from our literature searches. This represents 3,206 records from database and hand searches plus 115 studies previously included by a comparable search conducted by Doctor Evidence of which we excluded 32 references. Overall, we identified 92 studies reported in 111 publications that met inclusion criteria (Figure B—1).

Figure B—1. PRISMA flow chart.



Abbreviations. APA, American Psychiatric Association; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Appendix C. Review of Research Evidence Supporting Guideline Statements

Assessment and Determination of Treatment Plan

Statement 1 – Initial Assessment

APA *recommends* **(1C)** that the initial assessment of a patient with possible borderline personality disorder include the reason the individual is presenting for evaluation; the patient's goals and preferences for treatment; a review of psychiatric symptoms, including core features of personality disorders and common co-occurring disorders; a psychiatric treatment history; an assessment of physical health; an assessment of psychosocial and cultural factors; a mental status examination; and an assessment of risk of suicide, self-injury, and aggressive behaviors, as outlined in APA's Practice Guidelines for the Psychiatric Evaluation of Adults (3rd edition).

Evidence for this statement comes from general principles of assessment and clinical care in psychiatric practice. Expert opinion suggests that conducting such assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details, see Guideline I, "Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History," Guideline III, "Assessment of Suicide Risk," Guideline IV, "Assessment of Risk for Aggressive Behaviors," Guideline V, "Assessment of Cultural Factors," and Guideline VI, "Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016a). A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any studies related to this recommendation in the context of BPD treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Assessment of a Patient with Possible BPD

On the basis of the limitations of the evidence for assessment of a patients with possible BPD, no grading of the body of research evidence is possible.

Statement 2 – Quantitative Measures

APA *suggests* **(2C)** that the initial psychiatric evaluation of a patient with possible borderline personality disorder include a quantitative measure to identify and determine the severity of symptoms and impairments of functioning that may be a focus of treatment.

Evidence for this statement comes from general principles of assessment and clinical care in psychiatric practice. Consequently, the strength of research evidence is rated as low. Expert opinion suggests that conducting quantitative assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and longitudinal assessment of patient symptoms and treatment effects. This recommendation is also consistent with Guideline VII, "Quantitative Assessment," as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016a).

Grading of the Overall Supporting Body of Research Evidence for Use of Quantitative Measures

On the basis of the limitations of the evidence for use of quantitative measures, no grading of the body of research evidence is possible.

Statement 3 – Treatment Planning

APA *recommends* **(1C)** that a patient with borderline personality disorder have a documented, comprehensive, and person-centered treatment plan.

Evidence for this statement comes from general principles of assessment and clinical care in psychiatric practice. For additional details, see the American Psychiatric Association Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016a). A detailed systematic review to support this statement was outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any studies that directly related to this recommendation in the context of BPD treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Evidence-Based Treatment Planning

On the basis of the limitations of the evidence for evidence-based treatment planning, no grading of the body of research evidence is possible.

Statement 4 – Discussion of Diagnosis and Treatment

APA *recommends* **(1C)** that a patient with borderline personality disorder be engaged in a collaborative discussion about their diagnosis and treatment, which includes psychoeducation related to borderline personality disorder.

In terms of collaborative discussion about diagnosis and treatment, evidence for this statement comes from general principles of clinical care in psychiatric practice. Psychoeducation is also generally accepted as an important element of psychiatric care. In addition, several studies have examined effects of psychoeducation in individuals with BPD, but these did not find a significant effect of psychoeducation, per se.

Psychoeducation Versus Wait-List

Two RCTs (N=50 and N=80), rated as having a moderate risk of bias, assessed the effectiveness of psychoeducation compared with a wait-list control over 12 weeks (Zanarini and Frankenburg 2008; Zanarini et al. 2018). Psychoeducation consisted of an internet-based program detailing the latest information on BPD in one study (Zanarini et al. 2018) and a single workshop in the other (Zanarini and Frankenburg 2008). Participants received psychoeducation in addition to treatment as usual (TAU). Participants in the control group were on a wait-list for psychoeducation and continued with TAU only.

All participants were women, and the majority were white. The mean age was 21 (Zanarini et al. 2018) and 19 years (Zanarini and Frankenburg 2008). Only one study reported the severity of BPD at baseline (Zanarini et al. 2018). Participants were mildly ill at baseline with mean Zanarini Rating Scale for BPD (ZAN-BPD) scores ranging from 10.13 to 12.13 (Zanarini et al. 2018).

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—1 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

Both studies assessed the severity of BPD on the ZAN-BPD (Zanarini and Frankenburg 2008; Zanarini et al. 2018) and reported nonsignificant differences between the psychoeducation and the wait-list groups. In addition, one RCT reported similar treatment effects between groups on the Borderline Evaluation of Severity Over Time (BEST) scale (Zanarini et al. 2018). This RCT reported significantly better scores for the psychoeducation group after 12 months of follow-up (Zanarini et al. 2018). The investigators, however, tested 10 outcome measures and did not adjust for multiple comparisons.

Severity of Symptoms Associated With Borderline Personality Disorder

The larger of the two RCTs (N=80; Zanarini et al. 2018) employing internet-based psychoeducation reported no significant differences between intervention and wait-list groups for anxiety and depressive symptoms. Participants in the psychoeducation group, however, achieved significantly better scores on the Social Adjustment Scale than participants in the wait-list group. As mentioned above, however, this study tested 10 outcome measures and did not adjust for multiple testing.

Global Impression and Functioning

One study (N=80; Zanarini et al. 2018) reported similar effects and no significant differences on the Sheehan Disability Scale after 12 weeks and 12 months.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

None of the studies reported on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events (Zanarini and Frankenburg 2008; Zanarini et al. 2018).

Table C—1. Certainty-of-evidence ratings of outcomes comparing psychoeducation with wait-list control.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with wait-list	Difference in effects with psychoeducation
Severity of BPD assessed with: ZAN-BPD follow-up: mean 12 weeks	130 (2 RCTs) (Zanarini and Frankenburg 2008; Zanarini et al. 2018)	⨁⨁◯◯ LOW[a] for similar effects	-	The mean score at endpoint was 9.16	mean 1.33 lower (ns)
Anxiety assessed with: CUXOS follow-up: mean 12 weeks	80 (1 RCT) (Zanarini et al. 2018)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 40.11	mean 4.96 lower (ns)
Depression assessed with: CUDOS follow-up: mean 12 weeks	80 (1 RCT) (Zanarini et al. 2018)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 26.89	mean 6.11 lower (ns)
Functioning assessed with: SDS follow-up: mean 12 weeks	80 (1 RCT) (Zanarini et al. 2018)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 9.76	mean 2.18 higher (ns)

Note. [a] Studies do not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. BPD, borderline personality disorder; CI, confidence interval; CUDOS, Clinically Useful Depression Outcome Scale; CUXOS, Clinically Useful Anxiety Outcome Scale; GRADE Grading of Recommendations Assessment, Development, and Evaluation; ns, not significant; SDS, Sheehan Disability Scale; RCT, randomized controlled trial; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Grading of the Overall Supporting Body of Research Evidence for Psychoeducation in Patients With BPD

- Magnitude of effect: None noted. In the two studies that specifically assessed psychoeducation in BPD, no differences were noted as compared to a wait list control condition.

- Risk of bias: Moderate. Both studies of psychoeducation in BPD were rated as having a moderate risk of bias.

- Applicability: In both studies, participants were female, with a mean age of 19 to 22 years. Race was predominantly white in both studies with some other races and ethnicities represented in one study. One study used in-person psychoeducation, whereas the other study used internet-based psychoeducation, which is less common. One of the studies also excluded individuals who were currently receiving psychiatric treatment, which would also be atypical. Thus, the applicability of these studies to typical treatment of individuals with BPD appears limited.

- Directness: Direct. Measured outcomes include BPD symptom severity and functioning.

- Consistency: Inconsistent. The internet-based psychoeducation study showed better outcomes with psychoeducation at 12 weeks on social adjustment and at 12 months on BPD severity, whereas the other study showed no differences with psychoeducation.

- Precision: Imprecise. The studies did not meet the optimal information size (i.e., number of participants in a meta-analysis).

- Dose-response relationship: Not applicable. Dose-response was not studied.

- Confounding factors (including likely direction of effect): Not identified.

- Publication bias: Not identified.

- Overall strength of research evidence: Low. Only two studies are available that assessed BPD severity and for other outcomes including functioning, only one study was available. Both studies were relatively small and were of moderate risk of bias. The strength of evidence was also downgraded for imprecision and there was inconsistency in the findings of the two studies.

Psychosocial Interventions

Statement 5 – Psychotherapy

APA *recommends* **(1B)** that a patient with borderline personality disorder be treated with a structured approach to psychotherapy that has support in the literature and targets the core features of the disorder.

Evidence in the treatment of adults with BPD comes from the systematic review conducted by RTI. The data from clinical trials include comparisons with wait-list control and TAU conditions as well as head-to-head comparisons of specific psychotherapies. For the vast majority of treatments, there were only one or two studies of each comparison, which makes it challenging to draw robust conclusions. Notably, in the vast majority of studies that used TAU or an active comparator treatment, all treatment arms

showed improvement with psychotherapy even when differences between the treatment groups did not show statistically significant differences. This consistency as well as the superiority of many of the psychotherapies to TAU led the writing group to assess the overall strength of research evidence as moderate for psychotherapy in BPD.

For adolescents with BPD, the evidence for psychotherapeutic interventions is more limited but generally consistent with the benefits of treatment found in adults. Two studies in adolescents met the inclusion criteria for this review (Chanen et al. 2008; Santisteban et al. 2015) and are discussed in further detail below and in Appendix D. Other studies in adolescents did not meet inclusion criteria, primarily because they included patients with borderline traits as well as patients who fulfilled criteria for a diagnosis of BPD. A systematic review of studies in adolescents concluded that additional rigorous trials are needed because current studies have small samples, high attrition rates, inconsistent findings, and high risks of bias (Jørgensen et al. 2021).

Interpersonal Psychotherapy Versus Wait-List Plus Clinical Management

One RCT (Bozzatello and Bellino 2020) evaluated the efficacy of interpersonal psychotherapy compared with wait-list plus clinical management. The study included 43 participants in Italy who were assessed at 10 months. This study was rated as having a moderate risk of bias. The trial was funded by the Italian government.

The majority of the study participants were female; race was not reported. The overall mean age of participants was 35 years of age. The study excluded patients receiving psychiatric services or who had existing schizophrenia, bipolar disorder, mental impairment, or drug or alcohol dependence.

The intervention group received 22 sessions in the first 20 weeks and 20 sessions in the last 20 weeks. Each session lasted 50 minutes. TAU consisted of case management provided by hospital and primary and community care services.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—2 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

After 10 months of treatment, the study reported significantly greater improvements on the Borderline Personality Disorder Severity Index for participants in the interpersonal psychotherapy group compared with the wait-list plus clinical management group (Bozzatello and Bellino 2020).

Severity of Symptoms Associated With Borderline Personality Disorder

After 10 months of treatment, the study reported significantly greater improvements on the Barratt Impulsiveness Scale, version 11, but not on the Self-Harm Inventory, for participants in the interpersonal psychotherapy group compared with the wait-list plus clinical management group (Bozzatello and Bellino 2020).

Global Impression and Functioning

After 10 months of treatment, the study reported significantly greater improvements on the Clinical Global Impression Scale, Severity item and the Social Occupational Functioning Assessment Scale for participants in the interpersonal psychotherapy group compared with the wait-list plus clinical management group (Bozzatello and Bellino 2020).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—2. Certainty-of-evidence ratings of outcomes comparing interpersonal psychotherapy with wait-list plus clinical management.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with wait-list plus clinical management	Difference in effects with interpersonal psychotherapy
Severity of BPD assessed with: BPDSI follow-up: mean 10 months	43 (1 RCT) (Bozzatello and Bellino 2020)	⨁⨁◯◯ LOW[a] for greater effects with IPT	-	The mean score at endpoint was 36.1	mean 8.4 lower (p=0.01)
Severity of BPD symptoms assessed with: BIS-11 and SHI follow-up: mean 10 months	43 (1 RCT) (Bozzatello and Bellino 2020)	⨁◯◯◯ VERY LOW[b] for similar effects	-	The mean score at endpoint on BIS-11 was **64.8,** and on SHI was **6.91**	mean 12.6 lower on BIS-11 (p=0.03) and **2.8 higher** on SHI (p=0.27)
Functioning assessed with: CGI-S and SOFAS follow-up: mean 10 months	43 (1 RCT) (Bozzatello and Bellino 2020)	⨁⨁◯◯ LOW[a] for greater effects with IPT	-	The mean score at endpoint on CG1-S was 3.1 and on SOFAS was 57.1	mean 1.0 lower on CGI-S (p=0.009) and 11.1 higher on SOFAS (p=0.02)

Note. [a] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision; inconsistent direction of effect on measures of severity of BPD symptoms; downgraded 1 step for inconsistency
Abbreviations. CI, confidence interval; BIS-11, Barratt Impulsiveness Scale, version 11; BPDSI, Borderline Personality Disorder Severity Index; CGI-S, Clinical Global Impression-Severity; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; IPT, interpersonal psychotherapy; No., number; RCT, randomized controlled trial; SOFAS, Social Occupational Functioning Assessment Scale; SHI, Self-Harm Inventory.

Acceptance and Commitment Therapy Versus Treatment as Usual

One RCT (Morton et al. 2012) evaluated the efficacy of acceptance and commitment therapy (ACT) in addition to TAU compared with TAU alone. The Australian study included 41 participants who were followed for a duration of 13 weeks. The study was rated as having a moderate risk of bias because of high attrition. The trial did not report funding.

Almost all of the study participants were female. The mean age of the ACT group was 36 years while the mean age of the TAU group was 34 years. The study excluded participants with psychotic symptoms (besides "reactive psychotic symptoms" associated with BPD [not specified further]), with intellectual disability, with cognitive impairment, or who were a significant risk to other participants.

ACT was delivered as weekly group sessions that included performing mindfulness exercises, doing emotions skills training, focusing on awareness of one's values, and identifying choice points for action. TAU consisted of case management provided by public mental health services in Australia.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—3 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

After 13 weeks of treatment, the study reported significantly greater improvements on the BEST scale for participants in the ACT group compared with the TAU group (Morton et al. 2012).

Severity of Symptoms Associated With Borderline Personality Disorder

After 13 weeks, participants who received ACT in addition to TAU had significantly greater improvements than participants treated with TAU only on the Beck Hopelessness Scale, the Difficulties in Emotion Regulation Scale, and the subscale for anxiety of the Depression Anxiety Stress Scale. Changes on the subscales for depression and stress of the Depression Anxiety Stress Scale were also greater for the ACT group but did not achieve statistical significance (Morton et al. 2012).

Global Impression and Functioning

The study did not assess measures of global impression or functioning (Morton et al. 2012).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—3. Certainty-of-evidence ratings of outcomes comparing acceptance and commitment therapy with treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effect with ACT
Severity of BPD assessed with: BEST follow-up: mean 13 weeks	41 (1 RCT) (Morton et al. 2012)	⨁⨁◯◯ LOW[a] for greater effect with ACT	-	The mean score at endpoint was **47.4**	mean **17.2 lower** (p=0.028)
Anxiety assessed with: DASS follow-up: mean 12 days	41 (1 RCT) (Morton et al. 2012)	⨁⨁◯◯ LOW[a] for greater effect with ACT	-	The mean score at endpoint was **26.3**	mean **11.6 lower** (p=0.025)
Depression assessed with: DASS follow-up: mean 13 weeks	41 (1 RCT) (Morton et al. 2012)	⨁⨁◯◯ LOW[a] for greater effect with ACT	-	The mean score at endpoint was **31.0**	mean **15 lower** (ns)
Difficulties in Emotion Regulation assessed with: DERS follow-up: mean 13 weeks	41 (1 RCT) (Morton et al. 2012)	⨁⨁◯◯ LOW[a] for greater effect with ACT	-	The mean score at endpoint was **140.0**	mean **35.3 lower** (p=0.008)
Hopelessness assessed with: BHS follow-up: mean 13 weeks	41 (1 RCT) (Morton et al. 2012)	⨁⨁◯◯ LOW[a] for greater effect with ACT	-	The mean score at endpoint was **16.4**	mean **8.9 lower** (p=0.006)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[a] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. ACT, acceptance and commitment therapy; BEST, Borderline Evaluation of Severity Over Time; BHS, Beck Hopelessness Scale; BPD, borderline personality disorder; CI, confidence interval; DASS, Depression Anxiety Stress Scale; DERS, Difficulties in Emotion Regulation Scale; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; No., number; ns, not significant; RCT, randomized controlled trial; TAU, treatment as usual.

Manual-Assisted Cognitive Therapy Versus Treatment as Usual

One U.S. RCT (Weinberg et al. 2006) evaluated the efficacy of manual-assisted cognitive therapy (MACT), compared with TAU. Overall, the study provided data on 30 participants. The study was rated as having a moderate risk of bias. Follow-up duration was 6 months after treatment. The study was supported by a Young Investigator Award from the Borderline Personality Disorder Research Foundation. The majority of the study participants were female and white and had a mean age of 28 years. The study did not report on baseline severity. The study excluded participants with psychotic disorders, substance abuse disorder, or risk of suicide.

MACT was administered as an adjunctive intervention to TAU and comprised six sessions, over six to eight weeks, incorporating elements of dialectical behavior therapy (DBT), cognitive-behavioral therapy (CBT), and bibliotherapy, modified to focus on deliberate self-harm. Each session was structured around a chapter of a booklet, covering functional analysis of episodes of parasuicide (defined as deliberate self-harm or suicide attempts), emotion regulation strategies, problem-solving strategies, management of negative thinking, management of substance use, and relapse prevention strategies. TAU consisted of standard care.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—4 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

The study did not report on the severity of BPD.

Severity of Symptoms Associated With Borderline Personality Disorder

The study (Weinberg et al. 2006) reported significant reductions in the frequency and severity of deliberate self-harm for participants in the MACT group when compared with TAU after six months of treatment. The authors recorded the use of the Parasuicide History Interview to identify the frequency or severity of deliberate self-harm but did not specify the range of the scale for assessing severity.

Global Impression and Functioning

The study did not report on global impression or functioning.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—4. Certainty-of-evidence ratings of outcomes comparing manual-assisted cognitive therapy with treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effects with MACT
Deliberate Self-harm assessed with: Deliberate self-harm frequency (scale NR) follow-up: mean 6 months	30 (1 RCT) (Weinberg et al. 2006)	⨁⨁◯◯ LOW[a] for greater effects with MACT	-	The mean at endpoint for frequency was 6.69	mean 4.71 lower (p<0.001)
Deliberate Self-harm assessed with: Deliberate self-harm severity (scale NR) follow-up: mean 6 months	30 (1 RCT) (Weinberg et al. 2006)	⨁⨁◯◯ LOW[a] for greater effects with MACT	-	The mean severity score at endpoint was 1.01	mean 0.5 lower (p<0.001)

Note. [a] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. CI, confidence interval; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; MACT, manual-assisted cognitive therapy; No., number; NR, not reported; RCT, randomized controlled trial; TAU, treatment as usual.

Cognitive-Behavioral Therapy Versus Treatment as Usual

The BOSCOT (Borderline Personality Disorder Study of Cognitive Therapy) RCT (Davidson et al. 2006) evaluated the efficacy of CBT in addition to TAU compared with TAU only. The study included 106 participants in the United Kingdom who were followed for a duration of 24 months. The study was rated as having a moderate risk of bias. The trial was funded by a public foundation.

The majority of the study participants were female, and all of them were white (Davidson et al. 2006). The overall mean age of participants was 32 years of age. The study excluded patients receiving psychiatric services or who had existing schizophrenia or bipolar disorder, mental impairment, or drug or alcohol dependence.

The intervention group received an average of 27 sessions of CBT over 12 months in addition to TAU (Davidson et al. 2006). Each session lasted one hour. TAU consisted of case management provided by hospital and primary and community care services.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—5 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

The study did not report on the severity of BPD.

Severity of Symptoms Associated With Borderline Personality Disorder

The proportion of participants in the study (Davidson et al. 2006) who engaged in suicidal acts (defined as acts that were deliberate, life threatening, and resulting in or requiring medical intervention) was not significantly different between treatment groups after 24 months of follow-up. The number of mean suicidal acts per person had not reached significant differences at 12 months but was significantly lower for participants in the CBT group than the TAU group after 24 months. Improvements on the State-Trait Anxiety Inventory were significantly greater for participants in the CBT group compared with those treated with TAU only after 24 months but not after 12 months. No significant differences between treatment groups could be detected on the Beck Depression Inventory or for the number of hospitalizations after 12 months.

Global Impression and Functioning

No significant differences between treatment groups were detected for the Social Functioning Questionnaire and the European Quality of Life–5 Dimension instrument after 12 months (Davidson et al. 2006).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—5. Certainty-of-evidence ratings of outcomes comparing cognitive behavioral therapy with treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effects with CBT
Anxiety assessed with: STAI follow-up: mean 24 months	102 (1 RCT) (Davidson et al. 2006)	⨁⨁◯◯ LOW[b] for greater effect with CBT	-	The mean score at endpoint was 50.9	mean 7.96 lower (0 to 0)
Depression assessed with: BDI follow-up: mean 24 months	102 (1 RCT) (Davidson et al. 2006)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 28.8	mean 2.3 lower (0 to 0)
Proportion of Participants with Suicidal Acts follow-up: mean 24 months	102 (1 RCT) (Davidson et al. 2006)	⨁⨁◯◯ LOW[a] for similar risks	**OR 0.78** (0.30 to 1.98)	531 per 1,000	62 fewer per 1,000 (277 fewer to 161 more)
Mean Number of Suicidal Acts follow-up: mean 24 months	102 (1 RCT) (Davidson et al. 2006)	⨁⨁◯◯ LOW[a] for greater effect with CBT	-	The mean number at endpoint was 1.73	mean 0.91 lower (1.67 lower to 0.15 lower)
Quality of Life assessed with: EuroQuol-5D follow-up: mean 24 months	102 (1 RCT) (Davidson et al. 2006)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 0.66	mean 0.02 lower (0 to 0)
Social Functioning assessed with: SFQ follow-up: mean 24 months	102 (1 RCT) (Davidson et al. 2006)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 12.3	mean 0.7 lower (0 to 0)

Note. [a] Few events; downgraded 2 steps for imprecision.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. BDI, Beck Depression Inventory; CBT, cognitive-behavioral therapy; CI, confidence interval; EuroQuol-5D, European Quality of Life–5 Dimension; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; No., number; OR, odds ratio; RCT, randomized controlled trial; SFQ, Social Functioning Questionnaire; STAI, State-Trait Anxiety Inventory; TAU, treatment as usual.

Dialectical Behavior Therapy Versus Treatment as Usual

Six studies, four RCTs (Carter et al. 2010; Feigenbaum et al. 2012; McMain et al. 2017; Verheul et al. 2003), a nonrandomized trial (Bohus et al. 2004), and a retrospective cohort study (Gregory and Sachdeva 2016), evaluated the efficacy of DBT compared with TAU. Overall, these studies provided data on 483 participants. Three studies were rated as having a high risk of bias, two as moderate risk of bias and one as low risk of bias. Reasons for ratings of high risk of bias were lack of intention-to-treat analysis and high attrition. Follow-up durations ranged from 3 to 12 months. One trial was funded by a health insurance company; the other studies were publicly funded or did not report source of funding.

The majority of study participants were female, and mean ages ranged from 25 to 35 years. Only one study, in which the majority of participants were white, reported on race or ethnicity. Likewise, only one study reported the severity of BPD at baseline (Gregory and Sachdeva 2016). In this retrospective cohort study, participants were moderately ill at baseline, with BEST scores of 45 to 49. Studies excluded patients with psychiatric comorbidities such as schizophrenia, major depressive disorder (MDD), alcohol or substance use disorder, and bipolar disorder.

DBT combines weekly individual psychotherapy sessions, weekly skills training groups, and weekly supervision and consultation meetings for the therapists. One study assessed brief DBT with skills training only over 20 weeks (McMain et al. 2017). All studies enrolled outpatients, except a study from Germany, which conducted DBT as an inpatient treatment (Bohus et al. 2004).

TAU consisted of a range of individualized service provisions and professional mental health care. All except one study (Gregory and Sachdeva 2016) employed a wait-list design where participants of the TAU groups were offered DBT at the end of the study.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—6 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

In the study by McMain and colleagues (2017) (N=84), rated as having a moderate risk of bias, participants receiving brief DBT achieved significantly greater reductions on the Borderline Symptom List-23 compared with participants in the TAU group at the end of the intervention (20 weeks) but not at the 32-week follow-up. A retrospective cohort study (N= 41; Gregory and Sachdeva 2016) also reported no significant differences on the BEST scale between participants treated with DBT and TAU after 12 months.

Severity of Symptoms Associated With Borderline Personality Disorder

All six studies reported on changes in symptoms associated with BPD (Bohus et al. 2004; Carter et al. 2010; Feigenbaum et al. 2012; Gregory and Sachdeva 2016; McMain et al. 2017; Verheul et al. 2003). The two RCTs (N=84 and N=58), rated as having a moderate risk of bias, reported fewer suicide attempts in participants assigned to the DBT group than in participants receiving TAU (McMain et al. 2017; Verheul et al. 2003). By contrast, two studies (one RCT [Feigenbaum et al. 2012] and one cohort study [Gregory and Sachdeva 2016]), rated as having a high risk of bias, reported no significant differences in suicide attempts between treatment groups.

All studies reported on self-harm, defined variously as deliberate self-harm, self-injury, and self-mutilation. The majority of trials also showed greater reductions in self-harm in the DBT group than in the TAU group. In two trials (total N of 108), the difference in self-mutilating behaviors reached statistical significance (Bohus et al. 2004; Verheul et al. 2003).

Two studies, rated as having a high risk of bias, reported no significant differences in dissociative experiences between DBT and TAU (Bohus et al. 2004; Feigenbaum et al. 2012). One study reported on improvements of aggression (Feigenbaum et al. 2012) and impulsiveness (McMain et al. 2017), respectively; neither reported significant differences.

Studies reported mixed results regarding differences in efficacy between DBT and TAU to improve the severity of anger (Bohus et al. 2004; Feigenbaum et al. 2012; McMain et al. 2017) and depressive symptoms (Bohus et al. 2004; Feigenbaum et al. 2012; McMain et al. 2017).

Global Impression and Functioning

Significantly more participants in the brief DBT group than in the TAU group achieved clinically relevant improvements on the Symptom Checklist-90-Revised at 32 weeks (McMain et al. 2017). Likewise, Bohus and colleagues (2004) reported greater improvements on the Global Severity Index and the Global Assessment of Functioning Scale after four months of treatment with DBT than TAU.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

None of the studies reported on the incidence of adverse events and serious adverse events. The retrospective cohort study (Gregory and Sachdeva 2016) found no differences in withdrawals due to adverse events between participants treated with DBT and TAU (0% vs. 0%).

Table C—6. Certainty-of-evidence ratings of outcomes comparing dialectical behavior therapy with treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effect with DBT
Severity of BPD assessed with: BSC-23 follow-up: mean 32 weeks	125 (1 RCT, 1 observational study) (Gregory and Sachdeva 2016; McMain et al. 2017)	⨁⨁◯◯ LOW[a] for similar effects	-	The mean score at endpoint was 45.99*	mean 4.91 points higher (ns)
Anger, Depression assessed with: various scales follow-up: 3 to 12 months	227 (1 RCT, 1 nRCT, 1 observational study) (Bohus et al. 2004; Feigenbaum et al. 2012; Gregory and Sachdeva 2016; McMain et al. 2017)	⨁◯◯◯ VERY LOW[a,c,d,e] for similar effects	-	Inconsistent effects with TAU	inconsistent
Dissociative Experiences assessed with: DES follow-up: 3 to 12 months	102 (1 nRCT, 1 RCT) (Bohus et al. 2004; Feigenbaum et al. 2012)	⨁◯◯◯ VERY LOW[a,b] for similar effects		The mean score at endpoint was 83.3	mean 0.1 higher (ns)
Impulsiveness assessed with: BIS follow-up: mean 32 weeks	84 (1 RCT) (McMain et al. 2017)	⨁⨁◯◯ LOW[b] for similar effects	-	The mean score at endpoint was 55.16	mean 1.84 points lower (ns)
Self-harm assessed with: DHSI, self-injury, self-mutilation follow-up: mean 3 to 12 months	367 (4 RCTs, 1 nRCT, 1 observational study) (Bohus et al. 2004; Carter et al. 2010; Feigenbaum et al. 2012; Gregory and Sachdeva 2016; McMain et al. 2017; Verheul et al. 2003)	⨁⨁◯◯ LOW[c,d] for greater effect with DBT	not estimable	The mean score for DHSI at endpoint was 1.14*	mean 0.34 points lower (ns)

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effect with DBT
Suicidal and Nonsuicidal Self-injuries assessed with LSASI follow-up: mean 32 weeks	184 (3 RCTs) (Feigenbaum et al. 2012; McMain et al. 2017; Verheul et al. 2003)	⨁⨁◯◯ LOW[a] for greater effect with DBT	-	The mean score at endpoint was 2.56*	
General Psychopathology Achieving clinically relevant improvement on SCL-90-R follow-up: mean 32 weeks	134 (2 RCTs) (Bohus et al. 2004; McMain et al. 2017)	⨁⨁◯◯ LOW[a] for greater effect with DBT	**OR 3.44** (NR)	**184 per 1,000***	
Functioning assessed with: GAF follow-up: mean 4 months	50 (1 RCT) (Bohus et al. 2004)	⨁◯◯◯ VERY LOW[a,b] for greater effect with DBT	-	The mean score at endpoint was 49.4	
Withdrawal due to Adverse Events	41 (1 observational study) (Gregory and Sachdeva 2016)	⨁◯◯◯ VERY LOW[a,b] for similar risks	**RR 1** (-- to --)	0 per 1,000	

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

* Data based on McMain et al. 2017.

[a] Studies do not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

[c] Studies report inconsistent results regarding differences in treatment effects; downgraded 1 step for inconsistency.

[d] Studies do not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

[e] Two of three studies are high risk of bias.

Abbreviations. BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; BSC-23, Borderline Symptom Checklist-23; CI, confidence interval; DBT, dialectical behavior therapy; DES, Dissociative Experiences Scale; DHSI, Deliberate Self-Harm Inventory; GAF, Global Assessment of Functioning; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; LSASI, Lifetime Suicide Attempt Self-Injury Interview; No., number; NR, not reported; nRCT, nonrandomized clinical trial; ns, not significant; OR, odds ratio; RCT, randomized controlled trial, RR: risk ratio; SCL-90, Symptom Checklist-90; TAU, treatment as usual.

Dialectical Behavior Therapy Versus Mentalization-Based Treatment

One nonrandomized clinical trial (Barnicot and Crawford 2019) conducted in the United Kingdom, rated as having a high risk of bias, compared DBT with mentalization-based treatment (MBT) in 90 patients with BPD. The majority of participants were female (72%) with a mean age of 31 years. More than one-third (36%) were Black or belonged to a minority ethnic group. Mean baseline BPD severity ranged from 40.7 to 44.8 points on the BEST scale. Reasons for high risk of bias included selection bias and confounding.

Treatment duration for both DBT and MBT was 12 months, and the study was funded through the United Kingdom's National Institute for Health (Barnicot and Crawford 2019). DBT included weekly individual therapy and group skills training, telephone skills coaching, and team consultation. MBT included weekly or fortnightly individual therapy and weekly group therapy along with a short-term, 10-week group program offering psychoeducation and support aimed at helping patients get a better understanding of their problems and suggestions for better ways of dealing with them.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—7 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

At the end of the 12-month treatment phase, there was no significant difference in severity of BPD between DBT and MBT as measured by the BEST scale (Barnicot and Crawford 2019). There was significant improvement from baseline in both groups.

Severity of Symptoms Associated With Borderline Personality Disorder

At the end of the 12-month treatment phase, there was no significant difference between DBT and MBT in the number of self-harm incidents over the previous three months or in the number of dissociative symptoms and emotional dysregulation (Barnicot and Crawford 2019). Significant improvement from baseline in the severity of symptoms specific to BPD occurred in both groups.

Global Impression and Functioning

The study did not look at global impression or functioning at follow-up.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on treatment-related adverse events, including withdrawal due to adverse events.

Table C—7. Certainty-of-evidence ratings for mentalization-based treatment compared with dialectical behavior therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with DBT	Difference in effect with MBT
Severity of BPD assessed with: BEST follow-up: 12 months	90 (1 nRCT) (Barnicot and Crawford 2019)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **35.0** points	mean **0.8** points higher (ns)
Dissociative Experiences assessed with: DES follow-up: 12 months	90 (1 nRCT) (Barnicot and Crawford 2019)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **30.6** points	mean **4** points lower (ns)
Emotional Dysregulation assessed with: DERS follow-up: 12 months	90 (1 nRCT) (Barnicot and Crawford 2019)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **103.1** points	mean **5.6** points higher (ns)
Self-harm Incidents assessed with: SASII follow-up: 12 months	90 (1 nRCT) (Barnicot and Crawford 2019)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The median no. at endpoint was **2.0**	mean **10.5** more (ns)

Note. [a] High risk for bias in selection of participants into the study and high risk for confounding; downgraded 1 step for risk of bias.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. BEST, Borderline Evaluation of Severity Over Time; BPD, borderline personality disorder; CI, confidence interval; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; DES, Dissociative Experiences Scale; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; MBT, mentalization-based treatment; No., number; nRCT, non-randomized controlled trial; ns, nonsignificant; SASII, Suicide Attempt Self-Injury Interview.

Dialectical Behavior Therapy Versus General Psychiatric Management for Borderline Personality Disorder

One Canadian RCT (McMain et al. 2012; described in 3 publications), rated as having a high risk of bias, compared DBT with well-specified general psychiatric management in 180 patients with BPD. The majority of participants were female (86%) with a mean age of 30 years. Race and ethnicity were not reported. Mean baseline BPD severity ranged from 14.9 to 15.5 points on the ZAN-BPD. Reasons for high risk of bias included high attrition (38%) at 12 months.

Treatment duration was 12 months, and the study was funded through the Canadian Institutes for Health Research (McMain et al. 2012). DBT included weekly individual therapy and group skills training, weekly telephone coaching with explicit focus on self-harm and suicidal behavior, and weekly therapist team consultation. General manualized psychiatric management consisted of weekly individual therapy that was expanded away from focusing on self-harm and suicidal behaviors and included medication management. Generalized psychiatric therapy also included mandated therapist supervision weekly meetings.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—8 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

At the end of the 12-month treatment phase and again at the 36-month follow-up, there was no significant difference in severity of BPD on the ZAN-BPD among patients receiving DBT and those receiving general psychiatric management (McMain et al. 2012). There was significant improvement from baseline in both groups.

Severity of Symptoms Associated With Borderline Personality Disorder

With respect to symptoms specific to BPD, after 12 months of treatment and at the 36-month follow-up, there were no significant differences between DBT and general psychiatric management across multiple measures of symptom severity including the number of suicidal episodes and the number of nonsuicidal self-injuries as measured on the Suicide Attempt Self-Injury Interview and improvement on the Inventory of Interpersonal Problems scale (McMain et al. 2012). With respect to depression, there was no significant difference between groups in Beck Depression Inventory scores at the end of the 12-month treatment phase. However, at 36 months (24-month post-treatment), mean Beck Depression Inventory scores were significantly lower among patients in the general psychiatric management group than in the DBT group.

Global Impression and Functioning

The study reported no significant differences between treatment groups on the Symptom Checklist-90-Revised and the Inventory of Interpersonal Problems (McMain et al. 2012).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on treatment-related adverse events including withdrawal due to adverse events.

Table C—8. Certainty-of-evidence ratings for dialectical behavior therapy versus general psychiatric management for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with general psychiatric management	Difference in effect with DBT
Severity of BPD assessed with: ZAN-BPD follow-up: 36 months	180 (1 RCT) (McMain et al. 2012)	⊕⊕◯◯ LOW[a,b] for similar effects	-	The mean score at endpoint was 6.66 points	mean 1.63 points higher (ns)
Depression assessed with: BDI follow-up: 36 months	180 (1 RCT) (McMain et al. 2012)	⊕⊕◯◯ LOW[a,b] for greater effect with general psychiatric management	-	The mean score at endpoint was 18.05 points	mean 6.40 points higher (p=0.004)
Interpersonal Functioning assessed with: IIP follow-up: 36 months	180 (1 RCT) (McMain et al. 2012)	⊕⊕◯◯ LOW[a,b] for similar effects	-	The mean score at endpoint was 84.36 points	mean 10.12 points higher (ns)
Nonsuicidal Self-injuries assessed with SASII follow-up: 36 months	180 (1 RCT) (McMain et al. 2012)	⊕⊕◯◯ LOW[a,b] for similar effects	-	The mean no. at endpoint was 1.09	mean 1.09 more (ns)
Suicidal Episodes assessed with: SASII follow-up: 36 months	180 (1 RCT) (McMain et al. 2012)	⊕⊕◯◯ LOW[a,b] for similar effects	-	The mean no. at endpoint was 0.29	mean 0.26 more (ns)
Symptom distress assessed with: SCL-90-R total score follow-up: 36 months	180 (1 RCT) (McMain et al. 2012)	⊕⊕◯◯ LOW[a,b] for similar effects	-	The mean score at endpoint was 1.03 points	mean 0.23 points higher (ns)

Note. [a] High risk of bias due to attrition; downgraded 1 step for risk of bias.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

Abbreviations. BDI, Beck Depression Inventory; BPD, borderline personality disorder; CI, confidence interval; DBT, dialectical behavior therapy; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; IIP, Inventory of Interpersonal Problems; No., number; ns, nonsignificant; RCT, randomized controlled trial; SASII, Suicide Attempt Self-Injury Interview; SCL-90-R, Symptom Checklist-90-Revised; ZAN-BPD, Zanarini Rating Scale for BPD.

Dialectical Behavior Therapy Versus Systems Training for Emotional Predictability and Problem-Solving

One nonrandomized clinical trial (Guillén Botella et al. 2021) conducted in Spain, rated as having a high risk of bias, compared DBT with Systems Training for Emotional Predictability and Problem-Solving (STEPPS) in 72 patients with BPD. The overwhelming majority of participants were female (94%), and all were white with a mean age of 32 years. Mean baseline BPD severity ranged from 35.8 to 38.6 points on the Borderline Symptom List-23. The study was rated as having a high risk of bias due to high attrition (32%).

Treatment duration was six months (Guillén Botella et al. 2021). DBT included weekly individual therapy and group skills training, telephone skills coaching, and team consultation. STEPPS included group therapy, a reinforcement team, telephone consultations with relatives, consultations with other professionals, and weekly clinician meetings. The study funding source was not reported.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—9 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

At the end of a six-month treatment phase, compared with STEPPS, DBT resulted in a greater improvement in BPD symptom severity with significantly lower scores on the Borderline Symptom List-23 scale (Guillén Botella et al. 2021). Both DBT and STEPPS resulted in a significant improvement in BPD severity from baseline.

Severity of symptoms Associated With Borderline Personality Disorder

Following six months of treatment, there was no significant difference between DBT and STEPPS in suicide risk, depression, anxiety, dissociation experiences, and resilience scores (Guillén Botella et al. 2021). Severity of symptoms decreased across both groups.

Global Impression and Functioning

Following 6 months of treatment, there was no significant difference between STEPPS and DBT on quality-of-life scores (Guillén Botella et al. 2021).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report treatment-related adverse events including withdrawal due to adverse events.

Table C—9. Certainty-of-evidence ratings for dialectical behavior therapy versus systems training for emotional predictability and problem-solving.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with DBT	Difference in effect with STEPPS
Severity of BPD assessed with: BSL-23 follow-up: 6 months	72 (1 nRCT) (Guillén Botella et al. 2021)	⊕◯◯◯ VERY LOW[a,b] for greater effect with DBT	-	The mean score at endpoint was **23.56** points	mean **5.73** points higher (p=0.03)
Anxiety follow-up: 6 months	72 (1 nRCT) (Guillén Botella et al. 2021)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **8.40** points	mean **0.71** points higher (ns)
Depression assessed with: BDI follow-up: 6 months	72 (1 nRCT) (Guillén Botella et al. 2021)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **28.03** points	mean **6.7** points lower (ns)
Dissociation Experiences assessed with: DES-II follow-up: 6 months	72 (1 nRCT) (Guillén Botella et al. 2021)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **20.81** points	mean **2.8** points lower (ns)
Suicide Risk assessed with: SRS follow-up: 6 months	72 (1 nRCT) (Guillén Botella et al. 2021)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **7.0** points	mean **1.56** points higher (ns)
Quality of Life assessed with: QoL follow-up: 6 months	72 (1 nRCT) (Guillén Botella et al. 2021)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **6.31** points	mean **1.16** points lower (ns)

Note. [a] High risk of bias due to high attrition and moderate for confounding; downgraded 1 step for risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. BDI, Beck Depression Inventory; BPD, borderline personality disorder; BSL-23, Borderline Symptom List-23; CI, confidence interval; DBT, dialectical behavior therapy; DES-II, Dissociative Experiences Scale-II; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; nRCT, non-randomized controlled trial; No., number; ns: nonsignificant; QoL, Quality of Life Index; SRS, Suicide Risk Scale; STEPPS, Systems Training for Emotional Predictability and Problem-Solving.

Dialectical Behavior Therapy Versus Dynamic Deconstructive Psychotherapy

One three-armed retrospective cohort study (Gregory and Sachdeva 2016; reported in 2 publications) conducted in the United States, rated as having a high risk of bias, compared DBT with dynamic deconstructive psychotherapy (DDP) and TAU in 68 patients with BPD. The majority of participants were female (81%) and white (88%) with a mean age of 31 years. Mean baseline BPD severity ranged from 45.5 to 49.2 points on the BEST scale. Reasons for the rating of high risk of bias included high attrition (53%) and confounding.

Treatment duration for both DBT (N=25) and DDP (N=27) was 12 months (Gregory and Sachdeva 2016). DBT included weekly individual therapy, weekly group sessions, and telephone skills coaching. DDP included weekly individual sessions that combined elements of translational neuroscience, object relations theory, and deconstructionist philosophy. The study was supported by the American Psychoanalytic Association.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—10 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

At the end of a 12-month treatment phase, participants receiving DDP achieved significantly greater reductions on the BEST scale compared with participants receiving DBT (Gregory and Sachdeva 2016). Both DBT and DDP resulted in a significant improvement in BPD severity from baseline.

Severity of Symptoms Associated With Borderline Personality Disorder

Following 12 months of treatment, reductions in self-harm, as measured on the Suicidal Behaviors Questionnaire, and improvements in depression scores on the Beck Depression Inventory were significantly greater among patients receiving DDP than for those receiving DBT (Gregory and Sachdeva 2016). There was no difference at 12 months between DDP and DBT in reported suicide attempts.

Global Impression and Functioning

At 12 months, DDP resulted in significant greater improvement in disability with significantly lower scores on the Sheehan Disability Scale compared with DBT (Gregory and Sachdeva 2016).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report treatment-related adverse events including withdrawal due to adverse events.

Table C—10. Certainty-of-evidence ratings for dialectical behavior therapy compared with dynamic deconstructive psychotherapy for borderline personality disorder.

Outcomes	No. of participants (studies) follow-up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with DBT	Difference in effect with DDP
Severity of BPD assessed with: BEST follow-up: 12 months	52 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effect with DDP	-	The mean score at endpoint was 41.8 **points**	mean 8.8 points lower (p=0.04)
Depression assessed with: BDI follow-up: 12 months	52 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effect with DDP	-	The mean score at endpoint was 27.6 points	mean 10.5 points lower (p=0.009)
Disability assessed with: SDS follow-up: 12 months	52 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effect with DDP	-	The mean score at endpoint was 6.1 **points**	mean 2.3 points lower (p=0.049)
Self-harm assessed with: SBQ follow-up: 12 months	52 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effect with DDP	-	The mean no. at endpoint was 2.4	mean 1.1 fewer (p=0.02)
Suicide attempts assessed with: SBQ follow-up: 12 months	52 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean no. at endpoint was 1.3	mean 0.74 fewer (ns)

Note. [a] High risk of bias due to confounding and attrition; downgraded 1 step due to risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BPD, borderline personality disorder; CI, confidence interval; DBT, dialectical behavior therapy; DDP, dynamic deconstructive psychotherapy; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; ns, nonsignificant; SBQ, Suicidal Behaviors Questionnaire; SDS, Sheehan Disability Scale.

Dialectical Behavior Therapy Versus Transference-Focused Psychotherapy Versus Supportive Therapy

One three-armed RCT (Clarkin et al. 2007) rated as having a high risk of bias and conducted in the United States compared DBT with transference-focused psychotherapy (TFP) and supportive therapy in 90 patients with BPD and reported results for patients for whom they had at least three data points (N=62). The majority of participants were female (92%), white (68%), and with a mean age of 31 years. Mean baseline BPD severity was not reported. We rated the study as having high risk of bias due to the randomization process and high attrition (31%). Treatment duration was 12 months. DBT included weekly individual therapy, weekly group sessions, and telephone skills coaching. TFP included two individual weekly sessions focused primarily on the dominant affect-laden themes that emerge in the patient-therapist relationship. Supportive treatment included one weekly session supplemented with additional sessions as needed. The study was supported by the Borderline Personality Disorder Research Foundation.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—11, Table C—12, and Table C—13 present certainty-of-evidence ratings for the different comparisons.

Severity of Borderline Personality Disorder

The study did not report on severity of BPD.

Severity of Symptoms Associated With Borderline Personality Disorder

Following 12 months of treatment, there was a reduction in suicidal behavior (compared with baseline) among patients receiving DBT and TFP but not among those receiving supportive therapy (Clarkin et al. 2007). However, there was no significant difference between DBT, TFP, and supportive therapy. There was also no significant difference between treatment groups on the Beck Depression Inventory.

Global Impression and Functioning

Following 12 months of treatment, patients exhibited no significant differences between DBT, TFP, and supportive therapy on the Global Assessment of Functioning scale or the Brief Symptom Inventory for anxiety (data not provided) (Clarkin et al. 2007).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report treatment-related adverse events including withdrawal due to adverse events.

Table C—11. Certainty-of-evidence ratings for dialectical behavior therapy compared with transference-focused psychotherapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TFP	Difference in effect with DBT
Anxiety assessed with: BSI follow-up: 12 months	40 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Depression assessed with: BDI follow-up: 12 months	40 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Suicidal Behaviors assessed with: OAS-M follow-up: 12 months	40 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Global Functioning assessed with: GAF follow-up: 12 months	40 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)

Note. [a] High risk of bias due to improper randomization and high attrition; downgraded 1 step for risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. BDI, Beck Depression Inventory; BSI, Brief Symptom Inventory; CI, confidence interval; DBT, dialectical behavior therapy; GAF, Global Assessment of Functioning; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; NR, not reported; ns, nonsignificant; OAS-M, Overt Aggression Scale-Modified; RCT, randomized controlled trial; transference-focused psychotherapy.

Table C—12. Certainty-of-evidence ratings for dialectical behavior therapy compared with supportive therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with supportive therapy	Difference in effect with DBT
Anxiety assessed with: BSI follow-up: 12 months	39 (1 RCT) (Clarkin et al. 2007)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Depression assessed with: BDI follow-up: 12 months	39 (1 RCT) (Clarkin et al. 2007)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Global Functioning assessed with: GAF follow-up: 12 months	39 (1 RCT) (Clarkin et al. 2007)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Suicidal Behaviors assessed with: OAS-M follow-up: 12 months	39 (1 RCT) (Clarkin et al. 2007)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)

Note. [a] High risk of bias due to improper randomization and high attrition; downgraded 1 step for risk of bias.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. BDI, Beck Depression Inventory; BSI, Brief Symptom Inventory; CI, confidence interval; DBT, dialectical behavior therapy; GAF, Global Assessment of Functioning; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; NR, not reported; ns, nonsignificant; OAS-M, Overt Aggression Scale-Modified; RCT, randomized controlled trial.

Table C—13. Certainty-of-evidence ratings for transference-focused psychotherapy compared with supportive therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TFP	Difference in effect with Supportive Therapy
Anxiety assessed with: BSI follow-up: 12 months	45 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Depression assessed with: BDI follow-up: 12 months	45 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Suicidal Behaviors assessed with: OAS-M follow-up: 12 months	45 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)
Global Functioning assessed with: GAF follow-up: 12 months	45 (1 RCT) (Clarkin et al. 2007)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was NR	NR (ns)

Note. [a] High risk of bias due to improper randomization and high attrition; downgraded 1 step for risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. BDI, Beck Depression Inventory; BSI, Brief Symptom Inventory; CI, confidence interval; GAF, Global Assessment of Functioning; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; NR, not reported; ns, nonsignificant; OAS-M, Overt Aggression Scale-Modified; RCT, randomized controlled trial; transference-focused psychotherapy.

Dialectical Behavior Therapy Components Versus Other Components of Dialectical Behavior Therapy

DBT is a multifaceted cognitive-behavioral treatment approach that includes individual therapy, group skills training, telephone coaching, and a consultation team meeting for therapists. Three studies (1 nonrandomized clinical trial, 1 RCT, 1 prospective cohort study; Andión et al. 2012; Linehan et al. 2015; Lyng et al. 2020) assessed the comparative value of individual therapy components of DBT. Together, these studies provided data on 238 participants. One study (Andión et al. 2012) compared the individual therapy component of DBT with combined individual and group therapy. Another (Lyng et al. 2020) compared the stand-alone group skills component with six months of the full four-component DBT program. A third three-armed study (Linehan et al. 2015) compared 12 months of standard DBT (i.e., the full 4-component program) with stand-alone group skills training and individual therapy with an activities group. All three studies were rated as having a high risk of bias. Reasons for ratings of high risk of bias included high overall attrition or high differential attrition, bias due to deviations from the intended intervention, and bias due to confounding (Andión et al. 2012; Linehan et al. 2015; Lyng et al. 2020).

The majority of participants were female with a mean age across studies ranging from 26 to 33 years. Race was reported in just one of three studies in which more than 70% of participants were white (Linehan et al. 2015). Two studies were conducted in Europe (Andión et al. 2012; Lyng et al. 2020) and one in the United States (Linehan et al. 2015). Just one study provided baseline information on BPD severity, reporting a mean score on the Borderline Symptom List-23 of 2.7 points (Lyng et al. 2020) Treatment durations ranged from six months (Lyng et al. 2020) to one year (Andión et al. 2012; Linehan et al. 2015). One study followed patients through 18 months (6 months after the end of the intervention) (Andión et al. 2012), and another study followed patients through two years (12 months following the end of treatment) (Linehan et al. 2015). Studies were generally funded by public funds with no commercial funding.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—14, Table C—15, and Table C—16 present certainty-of-evidence ratings for different comparisons.

Severity of Borderline Personality Disorder

One prospective cohort study (Lyng et al. 2020), rated as having a high risk of bias, assessed improvements in the severity of BPD. The study, which included 88 participants, reported no clinical improvements in Borderline Symptom List-23 scores among patients receiving six months of stand-alone DBT skills training or six months of the full four-component DBT program and no significant difference between the groups. There were several serious limitations to the study including that high-risk patients (defined as a suicide attempt and/or deliberate self-harm that had required treatment by a physician in the previous 6 months) were excluded from the DBT skills training group but not from the full DBT group.

Severity of Symptoms Associated With Borderline Personality Disorder

Three studies (Andión et al. 2012; Linehan et al. 2015; Lyng et al. 2020) investigating individual components of DBT assessed changes in the severity of symptoms associated with BPD and all reported no significant differences between groups regarding reduction in suicide attempts and improvements in self-harm acts and suicidal ideation. One study (Linehan et al. 2015) found a significant improvement in Hamilton Rating Scale for Depression scores at the end of one-year treatment among participants

receiving standard DBT and the group skills component of DBT versus those receiving only the individual therapy component of DBT (p=0.02). There were no differences in anxiety scores at the end of the one-year treatment phase (Linehan et al. 2015).

Global Impression and Functioning

The study by Lyng and colleagues (N=88), rated as having a high risk of bias, comparing six months of stand-alone DBT skills training with six months of the full four-component DBT program reported no significant difference between groups on the Global Severity Index of the Symptom Checklist-90-Revised (Lyng et al. 2020).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

No studies reported on treatment-related adverse events including withdrawal due to adverse events.

Table C—14. Certainty-of-evidence ratings for dialectical behavior therapy group skills training compared with standard dialectical behavior therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with standard DBT	Difference in effect with DBT group skills training
Severity of BPD assessed with: BSL-23 follow-up: mean 6 months	88 (1 observational study) (Lyng et al. 2020)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **2.56** points	mean **0.51** Points lower (ns)
Self-harm Acts (NSSI) assessed with: SASII follow-up: mean 2 years	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[b,c] for similar effects	-	The mean no. at endpoint was **7.9**	mean **1.5** more (ns)
Suicidal Ideation assessed with: SBQ and BSS follow-up: 6 months to 2 years	154 (1 RCT, 1 observational study) (Linehan et al. 2015; Lyng et al. 2020)	⊕◯◯◯ VERY LOW[b,c] for similar effects	-	Not estimable (different scales)	mean **4.1 to** mean **7.7** points lower (ns)
Suicide Attempts assessed with: SASII follow-up: mean 2 years	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[b,c] for similar effects	-	The mean no. at endpoint was **2.0**	mean **0.5** fewer (ns)
General Psychopathology assessed with: SCL-90 follow-up: mean 6 months	88 (1 observational study) (Lyng et al. 2020)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **2.09**	mean **0.32** Points lower (ns)

Note. [a] High risk of bias due to attrition, confounding, and selection bias; downgraded 2 steps for risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[c] High risk of bias due to deviations from intended intervention and attrition; downgraded 1 step for risk of bias.
Abbreviations. BPD, borderline personality disorder; BSL-23, Borderline Symptom List-23; BSS, Beck Scale for Suicide Ideation; CI, confidence interval; DBT, dialectical behavior therapy; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; ns, nonsignificant; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; SASII, Suicide Attempt Self-Injury Interview; SBQ, Suicidal Behaviors Questionnaire; SCL-90, Symptom Checklist-90.

Table C—15. Certainty-of-evidence ratings for individual dialectical behavior therapy compared with standard dialectical behavior therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with standard DBT	Difference in effect with individual DBT therapy
Anxiety assessed with: Ham-A follow-up: end of 1-year treatment	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **17.2**	mean **7.1** points higher (ns)
Depression assessed with: Ham-D follow-up: end of 1-year treatment	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[a,b] for greater effect with DBT	-	The mean score at endpoint was **12.3**	mean **5.9** points higher (p=0.03)
Self-harm Acts (NSSI) assessed with: SASII follow-up: 2 years	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean no. at endpoint was **7.9**	mean **8.1** more (ns)
Suicidal Ideation assessed with: SBQ follow-up: 2 years	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **28.9** points	mean **3.4** Points lower (ns)
Suicide Attempts assessed with: SASII follow-up: mean 2 years	66 (1 RCT) (Linehan et al. 2015)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean no. at endpoint was **2.0**	mean **1.6** more (ns)

Note. [a] High risk of bias due to deviations from intended intervention and attrition; downgraded 1 step due to risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. CI, confidence interval; DBT, dialectical behavior therapy; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; no., number; ns, nonsignificant; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; SASII, Suicide Attempt Self-Injury Interview; SBQ, Suicidal Behaviors Questionnaire.

Table C—16. Certainty-of-evidence ratings for combined individual plus group dialectical behavior therapy compared with individual dialectical behavior therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with individual DBT therapy	Difference in effect with Combined individual plus group therapy DBT
Self-harm Behaviors assessed with: NR follow-up: 18 months	51 (1 nRCT) (Andión et al. 2012)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean no. at endpoint was **22**	mean **13** fewer (ns)
Suicide Attempts assessed with: NR follow-up: 18 months	51 (1 nRCT) (Andión et al. 2012)	⨁◯◯◯ VERY LOW[a,b] for similar effects	-	The mean no. at endpoint was **14**	mean **8** fewer (ns)

Note. [a] High risk of bias due to deviations from intended intervention; downgraded 1 step due to risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
Abbreviations. CI, confidence interval; DBT, dialectical behavior therapy; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; NR, not reported; ns, nonsignificant; nRCT, non-randomized controlled trial.

Dialectical Behavior Therapy Versus Community Therapy by Experts

One RCT (N=111) (Linehan et al. 2006), rated as having a high risk of bias, compared DBT with community therapy offered by nonbehavioral psychotherapy experts over one year. All participants were female with a mean age of 29 years who had at least two suicide attempts; the majority were white (87%). The severity of BPD at baseline was not reported.

We rated the study as having a high risk of bias because of lack of intention-to-treat analysis. The follow-up duration was two years, and the study was funded by the National Institute of Mental Health (Linehan et al. 2006).

The intervention group received standard DBT for one year, including weekly individual psychotherapy sessions, weekly group skills training, and telephone consultation as needed (Linehan et al. 2006). Community treatment by experts involved selected psychotherapists who were matched with therapists administering DBT by controlling for sex, availability, expertise, allegiance, training, and experience).

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—17 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

The study did not report on severity of BPD.

Severity of Symptoms Associated With Borderline Personality Disorder

At the end of the treatment period (12 months) and after the 2-year follow-up, participants in the DBT group had significantly fewer suicide attempt and emergency department visits or hospital admissions because of suicidal ideation and behavior (Linehan et al. 2006).

No significant differences between treatment groups were apparent for self-harm and depressive symptoms (Linehan et al. 2006).

Global Impression and Functioning

The study did not report on global impression and functioning.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—17. Certainty-of-evidence ratings of outcomes comparing dialectical behavior therapy with community therapy by experts.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with community therapy by experts	Difference in effect with DBT
Suicide Attempts follow-up: mean 2 years	101 (1 RCT)(Linehan et al. 2006)	⨁⨁◯◯ LOW[a,b] for greater effects with DBT	**HR 2.66** (2.40 to 18.07)	469 per 1,000	345 more per 1,000 (312 more to 531 more)
Self-harm assessed with: mean number of events follow-up: mean 2 years	101 (1 RCT) (Linehan et al. 2006)	⨁⨁◯◯ LOW[a,b] for similar effects	-	The mean number at endpoint was 3.0	mean 0 lower (ns)
Depression assessed with: Ham-D	101 (1 RCT) (Linehan et al. 2006)	⨁⨁◯◯ LOW[a,c] for similar effects	-	The mean score at endpoint was 14.4	mean 1.8 lower (ns)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[a] Lack of intention-to-treat analysis: downgraded 1 step for risk of bias.

[b] Overall few events; downgraded 1 step for imprecision.

[c] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

Abbreviations. CI, confidence interval; DBT, dialectical behavior therapy; GRADE Grading of Recommendations Assessment, Development, and Evaluation; Ham-D, Hamilton Rating Scale for Depression; HR, hazard ratio; ns, nonsignificant; RCT, randomized controlled trial.

Dynamic Deconstructive Psychotherapy Versus Treatment as Usual

A retrospective cohort study (Gregory and Sachdeva 2016) evaluated the efficacy of DDP compared with TAU. The study provided data on 44 participants. The study was rated as having a high risk of bias because of confounding and attrition. The follow-up duration was 12 months, and the study was funded by the American Psychoanalytic Association. The majority of the study participants were female and white, and the mean age was 28 years. This study reported baseline BEST scores ranging from 46 to 49. The study excluded patients with schizophrenia, intellectual disabilities, or dementia.

DDP involved weekly individual sessions over a 12-month period and combined elements of translational neuroscience, object relations theory, and deconstruction philosophy (Gregory and Sachdeva 2016). TAU consisted of unstructured psychotherapy.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—18 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

At the one-year follow-up, the study found that participants in the DDP group had significant improvements in the differences on the BEST scale when compared with the TAU group (Gregory and Sachdeva 2016).

Severity of Symptoms Associated With Borderline Personality Disorder

The study reported no significant differences in the mean number of self-injuries or suicide attempts but did report significant improvements in mean scores on the Beck Depression Inventory for participants in the DDP group when compared with TAU (Gregory and Sachdeva 2016).

Global Impression and Functioning

The study reported significant improvements in mean scores on the Sheehan Disability Scale for participants in the DDP group when compared with TAU after 12 months (Gregory and Sachdeva 2016).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—18. Certainty-of-evidence ratings of outcomes comparing dynamic deconstructive psychotherapy with treatment as usual.

Outcomes	No. of participants (studies) Follow-up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effects with DDP
Severity of BPD assessed with: BEST follow-up: mean 1 year	44 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effects with DDP	-	The mean severity score at endpoint was 42.9	Mean 9.9 lower (p=0.006)
Depression assessed with: BDI follow-up: mean 1 year	44 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effects with DDP	-	The mean depression score at endpoint was 29.6	Mean 12.5 lower (p<0.001)
Self-injuries follow-up: mean 1 year	44 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for similar effect	-	The mean number of self-injuries at endpoint was 1.8	Mean 0.5 lower (ns)
Suicide Attempts follow-up: mean 1 year	44 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for similar effect	-	The mean number of suicides at endpoint attempts was 1.5	Mean 0.94 lower (ns)
Functioning assessed with: SDS follow-up: mean 1 year	44 (1 observational study) (Gregory and Sachdeva 2016)	⊕◯◯◯ VERY LOW[a,b] for greater effects with DDP	-	The mean functioning score was 7.0	Mean 3.2 lower (p<0.001)

Note. [a] Not controlled for confounding; downgraded 2 steps for risk of bias.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BPD, borderline personality disorder; CI, confidence interval; DDP, dynamic deconstructive psychotherapy: GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; SDS, Sheehan Disability Scale; TAU, treatment as usual.

Mentalization-Based Treatment Versus Treatment as Usual

One RCT (Beck et al. 2020) evaluated the efficacy of MBT compared with TAU alone. The Danish study included 112 participants who were followed for a duration of 12 months. The study was rated as having a high risk of bias because of high attrition. The trial reported no commercial funding.

Almost all of the study participants were female with the exception of one person (Beck et al. 2020). The mean age was 16 years. The study excluded participants with comorbid diagnosis of pervasive developmental disorder, learning disability, anorexia, current psychosis, schizophrenia or schizotypal personality disorder, antisocial personality disorder, any other mental disorder other than BPD considered the primary diagnosis, current (past 2 months) substance use disorder (but not substance abuse), and current psychiatric inpatient treatment.

MBT, delivered over 12 months, consisted of three introductory sessions, 37 weekly group sessions (90-minutes each), five individual case formulation sessions, and six sessions for caregivers (Beck et al. 2020). TAU consisted of at least 12 individual supportive sessions, one per month, comprising psychoeducation, counseling, and crisis management and sessions as needed.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—19 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

After 12 months of treatment, the study reported no significant differences between groups on the Borderline Personality Features Scale for Children, the Borderline Personality Features Scale for Parents, or the ZAN-BPD (Beck et al. 2020).

Severity of Symptoms Associated With Borderline Personality Disorder

After 12 months of treatment, the study reported no significant differences between groups on self-harm (measured by the Risk-Taking and Self-Harm Inventory for Adolescents) or depression (measured by the Beck Depression Inventory for Youth) (Beck et al. 2020).

Global Impression and Functioning

After 12 months of treatment, the study reported no significant differences between groups on the Children's Global Assessment Scale (Beck et al. 2020).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study reported no adverse events in either arm.

Table C—19. Certainty-of-evidence ratings of outcomes comparing mentalization-based treatment with treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effects with MBT (95% CI)
Severity of BPD assessed with: BPFS-C, BPFS-P, ZAN-BPD follow-up: mean 1 years	112 (1 RCT) (Beck et al. 2020)	⨁⨁◯◯ LOW[a,b] for similar effects	-	The mean score at endpoint for BPFS-C was 71.3, for BPFS-P was 68.7, for ZAN-BPD was 8.0	Mean for BPFS-C was 0 (ns), for BPFS-P was 0.1 lower (-7.0 to 7.3), for ZAN-BPD was 0.6 lower (95% CI, -4.0 to 2.8)
BPD Symptoms assessed with: BDI-Y, RTSHIA follow-up: mean 1 year	112 (1 RCT) (Beck et al. 2020)	⨁⨁◯◯ LOW[a,b] for similar effects	-	The mean score at endpoint for BDI-Y was 64.3, for RTSHIA was 39.0	Mean for BDI-Y was 0.7 lower (-6.5 to 5.1), for RTSHIA was 1.4 lower (-7.1 to 4.3)
Functioning assessed with: CGAS follow-up: mean 1 year	112 (1 RCT) (Beck et al. 2020)	⨁⨁◯◯ LOW[a,b] for similar effects	-	The mean score at endpoint was 46.7	Mean was 0.5 higher (-5.8 to 6.7)

Note. [a] High attrition; downgraded 1 step for risk of bias.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

Abbreviations. BDI-Y, Beck Depression Inventory-Youth; BPD, borderline personality disorder; BPFS-C, Borderline Personality Features Scale for Children; BPFS-P, Borderline Personality Features Scale for Parents; CGAS, Children's Global Assessment Scale; CI, confidence interval; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; MBT, mentalization-based treatment; no., number; RCT, randomized controlled trial; RTSHIA, Risk-Taking and Self-Harm Inventory for adolescents; TAU, treatment as usual; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Mentalization-Based Treatment Versus Supportive Therapy

Three RCTs, described in four articles, compared MBT with supportive therapy (Bateman and Fonagy 2009; Bateman et al. 2021; Carlyle et al. 2020; Jørgensen et al. 2013). Together, these studies provided data on 317 participants. Supportive therapy was not identical across the studies, but all included group sessions that focused on supportive techniques such as problem-solving. Two studies were rated as having a moderate risk of bias (Bateman and Fonagy 2009; Carlyle et al. 2020), and the other as a high risk of bias (Jørgensen et al. 2013). Reasons for ratings of high risk of bias included high attrition and deviations from the intended intervention.

The majority of participants were female, and the mean age across the three studies was 31 years. Race was reported in two studies in which the majority of participants were white (Bateman and Fonagy 2009; Carlyle et al. 2020). Two studies were conducted in Europe (Bateman and Fonagy 2009; Jørgensen et al. 2013) and one in New Zealand (Carlyle et al. 2020). No study reported severity of BPD at baseline; however, one study reported global severity of symptoms at baseline that ranged from 1.7 to 2.0 points on the Symptom Checklist-90-Global Severity Index scale (Jørgensen et al. 2013). Treatment durations ranged from 18 months (Bateman and Fonagy 2009; Carlyle et al. 2020) to 24 months (Jørgensen et al. 2013). No study had commercial funding; one was funded through a foundation grant (Bateman and Fonagy 2009).

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—20 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

No study reported on the severity of BPD.

Severity of Symptoms Associated With Borderline Personality Disorder

All three studies assessed symptoms associated with BPD and reported mixed findings (Bateman and Fonagy 2009; Bateman et al. 2021; Carlyle et al. 2020; Jørgensen et al. 2013). Following 18 months of treatment, one study (N=134), rated as having a moderate risk of bias, reported a significant reduction in suicide attempts, hospitalizations, and life-threatening self-harm in the previous six-month period, along with improvements in interpersonal functioning and depression among patients receiving MBT compared with supportive therapy and case management (Bateman and Fonagy 2009). A six-year follow-up of 97 participants reported that, compared with the supportive treatment and case management group, significantly more of the MBT group who had achieved the primary recovery criteria (i.e., free of self-harm, suicide attempts, and inpatient hospital stays) remained well over a six-year follow-up period (Bateman et al. 2021).

In contrast, a similar study, rated as having a moderate risk of bias, attempting to replicate findings by Bateman and colleagues, found no significant differences between groups in incidents of severe self-harm and suicide attempts in the previous six months (Carlyle et al. 2020). Similarly, a study (N=111), rated as having a high risk of bias, reported no differences between groups in terms of interpersonal functioning, depression, and anxiety (Jørgensen et al. 2013).

Global Impression and Functioning

With the exception of one outcome for which there was agreement, studies reported mixed findings in terms of global impression and functioning (Bateman and Fonagy 2009; Bateman et al. 2021; Carlyle et al. 2020; Jørgensen et al. 2013). One study (N=134), rated as having a moderate risk of bias, reported significant improvements in Global Severity Index (using the Symptom Checklist-90 Global Severity Index) among patients receiving MBT compared with supportive therapy and case management (Bateman and Fonagy 2009). In contrast, a study (N=111), rated as having a high risk of bias, reported no differences between groups on the Symptom Checklist-90 Global Severity Index (Jørgensen et al. 2013). Both studies reported significant improvement in independently-rated global assessment functioning among patients receiving MBT compared with patients receiving supportive therapy (Bateman and Fonagy 2009; Jørgensen et al. 2013).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

No study reported on treatment-related adverse events including withdrawal due to adverse events.

Table C—20. Certainty-of-evidence ratings for mentalization-based treatment compared with supportive therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with supportive therapy	Difference in effect with MBT
Anxiety assessed with: BAI follow-up: 24 months	85 (1 RCT) (Jørgensen et al. 2013)	⊕◯◯◯ VERY LOW[a,e] for similar effects	-	The mean score at endpoint was 15.6 points	mean 2.1 points lower (ns)
Depression assessed with: BDI follow-up: 18 to 24 months	219 (2 RCTs) (Bateman and Fonagy 2009; Jørgensen et al. 2013)	⊕◯◯◯ VERY LOW[b,c,d] for inconsistent effects	-	The mean score at endpoint was 18.68[f] points	Inconsistent findings
General Psychopathology assessed with: SCL-90-GSI follow-up: 18 to 24 months	219 (2 RCTs) (Bateman and Fonagy 2009; Jørgensen et al. 2013)	⊕◯◯◯ VERY LOW[b,c,d] for inconsistent effects	-	The mean score at endpoint was 1.55[f] points	Inconsistent findings
Global Functioning assessed with: GAF follow-up: 18 to 24 months	219 (2 RCTs) (Bateman and Fonagy 2009; Jørgensen et al. 2013)	⊕⊕◯◯ LOW[b,d] for greater effect with MBT	-	The mean score at endpoint was 53.2[f] points	mean 7.7 points higher[f] (p<0.001)
Interpersonal Functioning assessed with: IIP follow-up: 18 to 24 months	219 (2 RCTs) (Bateman and Fonagy 2009; Jørgensen et al. 2013)	⊕◯◯◯ VERY LOW[b,c,d] for inconsistent effects	-	The mean score at endpoint was 1.65[f] points	Inconsistent findings
Severe Self-harm Incidents assessed with: SCL-90-R follow-up: 18 months	206 (2 RCTs) (Bateman and Fonagy 2009; Carlyle et al. 2020)	⊕⊕◯◯ LOW[c,d] for inconsistent effects	-	The mean no. at endpoint was 1.66[f]	Inconsistent findings

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with supportive therapy	Difference in effect with MBT
Suicide Attempts assessed with: SCL-90-R follow-up: 18 months	206 (2 RCTs) (Bateman and Fonagy 2009; Carlyle et al. 2020)	⨁⨁◯◯ LOW[c,d] for inconsistent effects	-	The mean no. at endpoint was 0.32 [f]	Inconsistent findings

Note. [a] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[b] One of 2 studies was high risk of bias due to attrition and deviations from intended intervention; downgraded 1 step for risk of bias.
[c] Two studies reported opposite direction of outcome; downgraded 1 step for inconsistency.
[d] Studies do not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.
[e] High risk of bias due to attrition and deviations from intended intervention; downgraded 1 step for risk of bias.
[f] Value is for the study rated at a moderate risk of bias (Bateman and Fonagy 2009).
***Abbreviations*.** BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; CI, confidence interval; GAF, Global Assessment of Functioning; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; GSI, Global Severity Index; IIP, Inventory of Interpersonal Problems; MBT, mentalization-based treatment; no., number; ns, nonsignificant; RCT, randomized controlled trial; SCL-90-R, Symptom Checklist-90-Revised; SCL-90-GSI, Symptom Checklist-90-Global Severity Index.

Mentalization-Based Treatment Versus Specialized Psychotherapy

Two studies, one RCT (Laurenssen et al. 2018), rated as having a moderate risk of bias, and one observational study with a nonconcurrent control group (Bales et al. 2015), rated as having a high risk of bias, compared day-hospital MBT with another specialized psychotherapy. Together, these studies provided data on 299 participants. Day-hospital MBT differed from other MBT in terms of intensity; it involved daily group psychotherapy and weekly individual therapy along with art and writing therapy. The specialized psychotherapy comparator groups consisted of a variety of treatments, settings, and durations that were explicitly not limited to supportive therapy. Reasons for the high risk of bias rating included confounding and measurement of outcomes.

The majority of participants were female, and the mean age across the two studies was 32 years (Bales et al. 2015; Laurenssen et al. 2018). Race was not reported in either study, both of which were conducted in the Netherlands. One study reported BPD severity ranging from 32.8 to 34.3 points at baseline using the BPD Severity Index (Laurenssen et al. 2018). Treatment duration was 18 months in both studies with one study following patients through 36 months (Bales et al. 2015). Neither study had commercial funding.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—21 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

One study (N=95), rated as having a moderate risk of bias, examined improvements in the severity of BPD as a primary outcome of interest and found no significant difference between 18 months of day-hospital MBT and 18 months of specialized psychotherapy in BPD Severity Index total scores and Personality Assessment Inventory–Borderline Features Scale scores (Laurenssen et al. 2018). There was significant improvement from baseline in both groups.

Severity of Symptoms Associated With Borderline Personality Disorder

One study, rated as having a moderate risk of bias, reported no significant difference between day-hospital MBT and specialized psychotherapy on the Inventory of Interpersonal Problems (Laurenssen et al. 2018). There was significant improvement from baseline in both groups.

Global Impression and Functioning

Both studies (Bales et al. 2015; Laurenssen et al. 2018) examined global symptom severity using the Global Severity Index of the Brief Symptom Inventory and found mixed results. As with the other outcomes, the study by Laurenssen and colleagues (2018) (N=95), rated as having a moderate risk of bias, reported no significant difference in the severity of symptoms among patients receiving day-hospital MBT and those receiving specialized psychotherapy. In contrast, at the end of 18 months of treatment and again at the 36-month follow-up, a study (Bales et al. 2015) (N=204), rated as having a high risk of bias, reported significant improvements in symptom severity (measured using the Brief Symptom Inventory-Global Severity Index) among patients receiving day-hospital MBT compared with those receiving specialized psychotherapy.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

One study (Laurenssen et al. 2018), rated as having a moderate risk of bias, reported no serious adverse events among patients receiving either day-hospital MBT or other specialized psychotherapy.

Table C—21. Certainty-of-evidence ratings for mentalization-based treatment compared with specialized psychotherapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with specialized psychotherapy	Difference in effect with day-hospital MBT
Severity of BPD assessed with: BPDSI follow-up: 18 months	95 (1 RCT) (Laurenssen et al. 2018)	⨁⨁◯◯ LOW[a] for similar effects	-	The mean score at endpoint was **21.39** points	mean **0.76** points lower (ns)
General Psychopathology assessed with: GSI of BSI follow-up: 18 to 36 months	299 (1 RCT, 1 observational study) (Bales et al. 2015; Laurenssen et al. 2018)	⨁◯◯◯ VERY LOW[b,c,d] for inconsistent effects	-	The mean score at endpoint was **1.04**[e] points	**Inconsistent findings**
Interpersonal Functioning assessed with: IIP follow-up: 18 months	95 (1 RCT) (Laurenssen et al. 2018)	⨁⨁◯◯ LOW[a] for similar effects	-	The mean score at endpoint was **NR**	**NR** (ns)

Note. [a] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[b] One of 2 studies was rated at a high risk of bias due to confounding and high risk for bias in the measurement of outcomes; downgraded 1 step for risk of bias.
[c] Two studies reported opposite direction of outcome; downgraded 1 step for inconsistency.
[d] Studies do not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.
[e] Value is for the study rated at a high risk of bias (Bales et al. 2015). Data NR for the study rated at a moderate risk of bias (Laurenssen et al. 2018).
Abbreviations. BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; BSI, Brief Symptom Inventory; CI, confidence interval; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; GSI, Global Severity Index; IIP, Inventory of Interpersonal Problems; MBT, mentalization-based treatment; no., number; NR, not reported; ns, nonsignificant; RCT, randomized controlled trial.

Systems Training for Emotional Predictability and Problem Solving Versus Treatment as Usual

Two RCTs (Blum et al. 2008; Bos et al. 2010) and one prospective cohort study (González-González et al. 2021) evaluated the efficacy of STEPPS compared with TAU. The studies provided data on 362 participants. One RCT (Bos et al. 2010) was rated as having a moderate risk of bias because of differential attrition, and one RCT (Blum et al. 2008) and one cohort study (González-González et al. 2021) were rated as having a high risk of bias for high overall attrition. Additionally, the cohort study had risks of bias from selection and confounding. The timing of the initial follow-up ranged from 20 to 24 weeks for the two RCTs. Both reported one-year outcomes. For one RCT, the primary endpoint was at 20 weeks (Blum et al. 2008); for the other, the primary endpoint was at one year (Bos et al. 2010). For the cohort study (González-González et al. 2021), the primary endpoint was at two years. Both RCTs were funded; neither had pharmaceutical industry support. The cohort study did not have specific funding. The majority of the study participants were female, and the mean age was 32 in the RCTs and 34 in the cohort study. Only one of the studies reported ethnicity; 94 percent of participants were white (Blum et al. 2008). One RCT reported mean baseline BEST scores ranging from 39 to 40 (Blum et al. 2008). The cohort study reported mean baseline BEST scores ranging from 50 to 52. Studies excluded patients with psychotic or primary neurological disorders, who were cognitively impaired, or who had participated in STEPPS previously.

STEPPS involved 18 or 20 weekly therapy sessions; components included psychoeducation about BPD, emotion management skills training, and behavior management skills training. TAU consisted of usual care such as individual psychotherapy, medication, and case management.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—22 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

All three studies reported that STEPPS was associated with significant improvements in BPD-specific symptoms (ZAN-BPD and Borderline Personality Disorder checklist-40) at the primary endpoint (20 weeks, 1 year, and 2 years, respectively) as compared to TAU. One RCT, rated as having a high risk of bias, however, reported no differences on the BEST scale for participants in the STEPPS group compared with the TAU group at 20 weeks or between 20 weeks and one year (Blum et al. 2008).

Severity of Symptoms Associated With Borderline Personality Disorder

One RCT (Blum et al. 2008), rated as having a high risk of bias, reported significant improvement in impulsiveness (measured by the Barratt Impulsiveness Scale) and depression (measured by the Beck Depression Inventory) for participants in the STEPPS group when compared with TAU at 20 weeks. The same RCT reported no significant differences in suicide attempts or self-harm acts at one year.

Global Impression and Functioning

Both RCTs (Blum et al. 2008; Bos et al. 2010) reported on global impression and functioning using four scales: global impression using Symptom Checklist-90 (at 20 weeks and 1 year in 1 RCT and at 24 weeks in another) and Clinical Global Impressions (at 20 weeks and 1 year in 1 RCT), quality of life using the World Health Organization Quality of Life scale (at 1 year in 1 RCT), and functioning using the Social Adjustment Scale and Global Assessment Scales (at 20 weeks and 1 year in 1 RCT). Together, these findings suggest benefits in global impression and functioning for the STEPPS group compared with TAU.

Regarding global impressions at 20-24 weeks, using Symptom Checklist-90, both RCTs reported significant improvement for the STEPPS when compared with TAU (Blum et al. 2008; Bos et al. 2010). One RCT (Blum et al. 2008), rated as having a high risk of bias, also reported significant improvement for the STEPPS group when compared with TAU at 20 weeks in Clinical Global Impressions severity and improvement ratings. The same study, rated as having a high risk of bias, reported no significant differences between 20 weeks and one year in Symptom Checklist-90 or Clinical Global Impressions severity or improvement ratings.

Regarding quality of life, one RCT (Bos et al. 2010), rated as having a moderate risk of bias, reported significant improvement for the STEPPS group when compared with TAU at one year.

Regarding functioning, one RCT (Blum et al. 2008), rated as having a high risk of bias, reported significant differences favoring the STEPPS group at 20 weeks and no significant differences between 20 weeks and one year in functioning (measured by the Global Assessment Scale). However, the same study reported no significant differences in social adjustment (measured by the Social Adjustment Scale at 20 weeks and between 20 weeks and 1 year).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The studies did not report on adverse events or withdrawals due to adverse events.

Table C—22. Certainty-of-evidence ratings of outcomes comparing systems training for emotional predictability and problem solving with treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effects with STEPPS
Severity of BPD assessed with: ZAN-BPD, BPD-40, BEST follow-up: mean 20 weeks to 2 years	240 (2 RCTs, 1 prospective cohort) (Blum et al. 2008; Bos et al. 2010; González-González et al. 2021)	⊕⊕⊕◯ MODERATE for greater effects with STEPPS[a]	-	The mean score at primary endpoint on ZAN-BPD was 13.4; on BPD-40 was **88.6**; on BEST at primary endpoint was 34.1 in the trial, **28.8** in the cohort	Mean 3.6 lower on ZAN-BPD; 10.4 lower on BPD-40 (p= 0.001); 2.3 lower on BEST (ns) in the trial, **17.7 lower** in the cohort (p<0.0)
Depression assessed with: BDI follow-up: mean 20 weeks	124 (1 RCT) (Blum et al. 2008)	⊕⊕◯◯ LOW[a,b] for greater effect with STEPPS	-	The mean score at primary endpoint was 25.8	Mean 3.8 higher (p=0.03)
Impulsiveness assessed with: BIS follow-up: mean 20 weeks	124 (1 RCT) (Blum et al. 2008)	⊕⊕◯◯ LOW[a,b] for greater effect with STEPPS	-	The mean score at primary endpoint was 76.8	Mean 4.1 lower (p=0.004)
Self-harm Attempts follow-up: mean 1 year	124 (1 RCT) (Blum et al. 2008)	⊕⊕◯◯ LOW[a,b] for similar effects	not estimable	Not reported	(ns)
Suicide Attempts follow-up: mean 1 year	124 (1 RCT) (Blum et al. 2008)	⊕⊕◯◯ LOW[a,b] for similar effects,	not estimable	Not reported	(ns)
General Psychopathology assessed with: CGI-S, CGI-I, SCL-90 follow-up: 20 weeks to 1 year	203 (2 RCTs) (Blum et al. 2008; Bos et al. 2010)	⊕⊕⊕◯ MODERATE for greater effects with STEPPS[a]	-	Varied by study and measure	p≤0.03

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with TAU	Difference in effects with STEPPS
Quality of Life assessed with: WHOQOL follow-up: mean 1 year	79 (1 RCT) (Bos et al. 2010)	⨁⨁⨁◯ MODERATE[b] for greater effect with STEPPS	-	The mean score at primary endpoint was 11.3	Mean 1.3 higher (0 to 0)
Functioning assessed with: GAS, SAS follow-up: mean 20 weeks	124 (1 RCT) (Blum et al. 2008)	⨁⨁◯◯ LOW[a,b] for greater effect with STEPPS	-	The mean score at primary endpoint on GAS was 43.5 and on SAS was 26.3	Mean 7 higher on GAS (ns) and **1.7 lower** on SAS (ns)

Note. [a] High overall attrition; downgraded 1 step for risk of bias.

[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

Abbreviations. BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; BPD-40, Borderline Personality Disorder checklist-40; CGI-I, Clinical Global Impressions-Improvement; CGI-S, Clinical Global Impressions-Severity; CI, confidence interval; GAS, Global Assessment Scale; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; ns, not significant; RCT, randomized controlled trial; SAS, Social Assessment Scale; SCL-90, Symptom Checklist-90; STEPPS, Systems Training for Emotional Predictability and Problem Solving; TAU, treatment as usual; WHOQOL, World Health Organization Quality of Life; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Transference-Focused Psychotherapy Versus Treatment as Usual

One RCT (Doering et al. 2010) conducted in Austria and Germany evaluated the efficacy of TFP compared with TAU. The study provided data on 104 participants. The study was rated as having a high risk of bias because of high differential attrition from follow-up. Follow-up duration was 12 months. The trial was funded by the Austrian National Bank.

All of the study participants were female and had a mean age of 28 years (Doering et al. 2010). The ethnicity of the participants was not reported. Authors noted that the study included participants with less severe BPD, with higher Global Assessment of Functioning scores, fewer comorbid Axis I and II disorders, and fewer self-harming acts than other treatment studies of BPD because patients with more severe symptoms would receive inpatient treatment in Austria and Germany. Studies excluded patients with schizophrenia; bipolar I and II disorder with a major depressive, manic, or hypomanic episode during the previous six months; substance use disorder in the last six months; or organic pathology or intellectual disability.

TFP is a modified psychodynamic therapy and consists of two 50-minute sessions delivered every week by experienced clinical psychologists or medical doctors, along with medications as needed for one year of treatment (Doering et al. 2010). TAU consisted of individualized standard care from community psychiatrists.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—23 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

TFP was superior to TAU in last-observation-carried-forward analyses for the number of DSM-IV diagnostic criteria on average for BPD and proportion having fewer than five DSM-IV borderline criteria after one year (Doering et al. 2010).

Severity of Symptoms Associated With Borderline Personality Disorder

The study (Doering et al. 2010) reported a significantly lower proportion of participants with suicide attempts for TFP than TAU for last-observation-carried-forward analyses and marginally significant for number of suicide attempts. However, completers analyses controlling for dose response for number of psychotherapy sessions (48.5 sessions, on average, for TFP vs. 18.6 for community psychotherapists) found no significant differences in either measure. The study reported no significant differences in depression (Beck Depression Inventory) or state and trait anxiety (State–Trait Anxiety Inventory).

Global Impression and Functioning

TFP was significantly superior to TAU for Global Assessment of Functioning scores but not for the Brief Symptom Inventory (Doering et al. 2010).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report on the incidence of adverse events, serious adverse events, or withdrawal due to adverse events.

Table C—23. Certainty-of-evidence ratings of outcomes comparing transference-focused psychotherapy with treatment by treatment as usual.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effects with treatment by community psychotherapists	Difference in effects with TFP
Severity of BPD Symptoms assessed with: Proportion meeting fewer than 5 DSM-IV diagnostic criteria follow-up: mean 1 year	104 (1 RCT) (Doering et al. 2010)	⊕⊕◯◯ LOW[a,b] for greater effect with TFP	**RR 2.23** (1.07 to 4.65)	154 per 1,000	189 more per 1,000 (11 more to 562 more)
Anxiety assessed with: STAI follow-up: mean 1 year	104 (1 RCT) (Doering et al. 2010)	⊕◯◯◯ VERY LOW[a,d] for similar effect	-	The mean score at endpoint for state was 50.47 and trait anxiety was 55.49	Mean score for state was 2.30 higher and trait anxiety was 0.43 lower (ns)
Depression assessed with: BDI follow-up: mean 1 year	104 (1 RCT) (Doering et al. 2010)	⊕◯◯◯ VERY LOW[a,d] for similar effect	-	The mean score at endpoint was 20.02	Mean 1.65 higher (ns)
Suicide Attempts assessed with: Proportion with any suicide attempts follow-up: mean 1 year	104 (1 RCT) (Doering et al. 2010)	⊕◯◯◯ VERY LOW[a,c] for similar effect	**RR 0.63** (0.27 to 1.51)*	135 per 1,000	50 fewer per 1,000 (98 fewer to 69 more)
General Psychopathology assessed with: BSI follow-up: mean 1 year	104 (1 RCT) (Doering et al. 2010)	⊕◯◯◯ VERY LOW[a,d] for similar effect	-	The mean score at endpoint was 1.27	MD 0.06 higher (ns)
Functioning assessed with: GAF follow-up: mean 1 year	104 (1 RCT) (Doering et al. 2010)	⊕⊕◯◯ LOW[a,b] for greater effect with TFP	-	The mean score at endpoint was 56.06	Mean 2.6 higher (p=0.001)

Note. * Calculated based on data at follow-up.

[a] High overall and differential attrition; downgraded 1 step for risk of bias.

[b] Few events or study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for precision.

[c] Few events; significant LOCF results, adjustment for dose in completers analyses no longer significant; downgraded 2 steps for precision.

[d] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); results likely had wide Cis, p not significant; downgraded 2 steps for precision.

***Abbreviations*.** BDI, Beck Depression Inventory; BPD, borderline personality disorder; BSI, Brief Symptom Inventory; CI, confidence interval; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders*; GAF, Global Assessment of Functioning; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; LOCF, last observation carried forward; MD, mean difference; no., number; ns, not significant; RCT, randomized controlled trial; RR, risk ratio; STAI, State–Trait Anxiety Inventory; TFP, transference-focused psychotherapy.

Transference-Focused Psychotherapy Versus Schema-Focused Therapy

One RCT (described in 2 publications; Giesen-Bloo et al. 2006; Spinhoven et al. 2007), rated as having a high risk of bias and conducted in the Netherlands, compared TFP with schema-focused therapy (SFT) in 88 patients with BPD. The majority of participants were female (93%) with a mean age of 31 years. Race and ethnicity were not reported. Mean baseline BPD severity ranged from 33.5 to 34.4 points on the BPD Severity Index. Reasons for a rating of high risk of bias included high attrition (39%) and measurement of outcomes.

Treatment duration was three years (Giesen-Bloo et al. 2006). Both TFP and SFT included two 50-minute sessions per week. The TFP focused on the patient-therapist relationship, while the SFT involved integrated cognitive therapy focused on four schema modes. The study was funded by a grant from the Dutch Health Care Insurance Board.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—24 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

At the end of a three-year treatment phase, participants receiving SFT exhibited significant greater clinical improvement on the BPD Severity Index than patients receiving TFP. Reliable clinical improvement (defined as improvement of at least 11.7 points at the last assessment) favored SFT over TFP (RR=2.33 [95% CI, 1.24 to 4.37]) (Giesen-Bloo et al. 2006).

Severity of Symptoms Associated With Borderline Personality Disorder

The study did not report on symptoms associated with BPD.

Global Impression and Functioning

After three years of treatment, there was no significant difference between TFP and SFT in quality of life measures (Giesen-Bloo et al. 2006). There was significant improvement in quality-of-life scores from baseline in both groups.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report treatment-related adverse events including withdrawal due to adverse events.

Table C—24. Certainty-of-evidence ratings for transference-focused psychotherapy compared with schema-focused therapy for borderline personality disorder.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with transference-focused psychotherapy	Difference in effect with SFT
Severity of BPD assessed with: BPDSI follow-up: 3 years	88 (1 RCT) (Giesen-Bloo et al. 2006)	⊕◯◯◯ VERY LOW[a,b] for greater effect with SFT	-	The mean score at endpoint was **21.87** points	mean **5.63 points lower** (p=0.005)
Quality of Life assessed with: EuroQol follow-up: 3 years	88 (1 RCT) (Giesen-Bloo et al. 2006)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **67.5** points	mean **3.0 points lower** (ns)
Quality of Life assessed with: WHOQOL follow-up: 3 years	88 (1 RCT) (Giesen-Bloo et al. 2006)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was **11.09** points	mean **0.5 points higher** (ns)

Note. [a] High risk of bias due to high attrition and moderate risk of bias related to measurement of outcomes; downgraded 1 step for risk of bias.
[b] Study does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
***Abbreviations*.** BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; CI, confidence interval; EuroQol, European Quality of Life scale; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; no., number; ns, nonsignificant; RCT, randomized controlled trial; SFT, schema-focused therapy; TFP, transference-focused psychotherapy; WHOQOL, World Health Organization Quality of Life Scale.

Psychotherapy for Special Populations

Detailed information on main study characteristics and treatment effects is presented in Appendix D for nine studies that compared various psychotherapies within special populations. Overall, there is no evidence to support one psychotherapy over another for any of the special populations identified.

Comprehensive Validation Therapy Plus 12-Step Versus Dialectical Behavior Therapy for BPD and Substance Use Disorder

One RCT (N=24; Linehan et al. 2002), rated as having a moderate risk of bias, conducted in the United States compared comprehensive validation therapy plus 12-step (a manualized approach that provided the major acceptance-based strategies used in DBT in combination with participation in 12-step programs) with DBT for the treatment of comorbid BPD and substance use disorder. At a 16-month follow-up, there was no significant difference between comprehensive validation therapy plus 12-step and DBT in percentage opiate-positive urine specimens, Brief Symptom Inventory scores, and scores on the Global Adjustment Scale, although the percentage opiate-positive urine specimens decreased and rating scale scores improved in both groups. In addition, the incidence of parasuicidal behavior, measured using the Parasuicide History Interview, did not differ between groups and was low throughout the treatment period.

Mentalization-Based Treatment Plus Combination Substance Use Disorder Treatment Versus Substance Use Disorder Treatment Alone for BPD and Substance Use Disorder

One feasibility RCT (N=46; Philips et al. 2018) conducted in Sweden, rated as having a high risk of bias, compared MBT plus combination substance use disorder treatment with substance use disorder treatment alone for the treatment of BPD and substance use disorder. The MBT included a combination of individual therapy and group therapy over 18 months. At 18 months there was no significant difference between groups on any outcome measured, including borderline symptom severity, suicide attempts, self-harm, inventory of interpersonal problems, reflective functioning, and global functioning.

Dynamic Deconstructive Psychotherapy Versus TAU in the Community for BPD and Alcohol Use Disorder

One RCT (N=30; Gregory et al. 2008) conducted in the United States, rated as having a high risk of bias, compared DDP with TAU for the treatment of comorbid BPD and alcohol use disorder. DDP involved weekly individual therapy focused on fostering verbalization of affects and elaboration of recent interpersonal experiences into simple narratives. Participants were encouraged but not required to attend some form of group therapy. Most TAU participants received a combination of individual psychotherapy and medication management. At 12 months, there was no significant difference between DDP and TAU in parasuicide behavior (measured using the adapted three-month version of the Lifetime Parasuicide Count), alcohol misuse, and dissociation. DDP led to significant improvements in depression and in core symptoms of BPD as measured by the BEST scale.

Dialectical Behavior Therapy Plus Dialectical Behavior Therapy-Prolonged Exposure Versus Dialectical Behavior Therapy Alone for BPD and Posttraumatic Stress Disorder

One RCT described in two publications (N=26; Harned et al. 2014, 2018), rated as having a high risk of bias and conducted in the United States, compared DBT plus DBT-prolonged exposure with standard

DBT for the treatment of comorbid BPD and posttraumatic stress disorder. This pilot study did not conduct a between-group statistical analysis on the primary outcomes related to intentional self-harm. Preliminary findings suggested that DBT plus prolonged exposure may improve global social adjustment, health-related quality of life, and achievement of good global functioning, but not interpersonal problems or quality of life.

Cognitive-Behavioral Therapy Versus Dialectical Behavior Therapy for BPD and Eating Disorders

One nonrandomized clinical trial (N=118; Navarro-Haro et al. 2021), rated as having a moderate risk of bias, compared CBT (described as TAU) with DBT for the treatment of comorbid BPD and eating disorders and found no significant differences between groups in the primary outcome of suicide attempts in the previous six months. Depression scores on the Beck Depression Inventory-II were significantly better among patients receiving DBT than cognitive-behavioral therapy. At a 6-year follow-up of 69 participants, there were no significant differences between participants who had received DBT and those who had received CBT for depression, emotional regulation, and resilience.

Specialist Supportive Clinical Management Versus Modified Mentalization-Based Treatment for BPD and Eating Disorders

One RCT (N=68; Robinson et al. 2016) conducted in the United Kingdom, rated as having a high risk of bias, compared specialist supportive clinical management with modified MBT for the treatment of comorbid BPD and eating disorders and found no significant difference between groups on the ZAN-BPD.

Cognitive Therapy Plus Fluoxetine Versus Interpersonal Therapy Plus Fluoxetine for BPD and Major Depressive Disorder

One RCT (N=32; Bellino et al. 2007), conducted in Italy and rated as having a moderate risk of bias, compared cognitive therapy plus fluoxetine with interpersonal therapy plus fluoxetine for the treatment of comorbid BPD and MDD and at the 24-weeks follow-up found no differences between groups in symptoms of depression, anxiety, or global functioning scales.

Individual Drug Counselling Versus Integrative Borderline Personality Disorder-Oriented Adolescent Family Therapy for BPD and Substance Use Disorder Among Adolescents

One RCT (N=40; Santisteban et al. 2015), conducted in the United States and rated as having a high risk of bias, compared individual drug counseling with integrative BPD-oriented adolescent family therapy for the treatment of comorbid BPD and substance use disorder. Individual drug counseling consisted of two sessions per week of individual manualized drug counseling with a monthly family meeting with caregivers. Goals of the treatment included identifying signs and symptoms of addiction and triggers to use, increasing motivation to achieve and sustain abstinence, and developing more effective problem-solving strategies. Integrative BPD-oriented adolescent family therapy consisted of two sessions per week that included family therapy, individual therapy, and skills-building interventions that targeted factors that directly contribute to adolescent drug abuse and other self-harm behaviors such as emotion dysregulation and impulsivity, failure to establish life goals and ineffective life skills, unstable family attachment, and maladaptive family interactions. At the 12-month follow-up, there was no significant difference between individual drug counseling and integrative BPD-oriented adolescent family therapy

on BPD behavior as measured on the borderline personality scale from the Millon Adolescent Clinical Inventory and no significant difference in substance use.

Manualized Good Clinical Care Versus Cognitive Analytic Therapy for Adolescents With BPD

One RCT (N=86; Chanen et al. 2008), rated as having a moderate risk of bias and conducted in Australia, compared manualized good clinical practice with cognitive analytic therapy (which uses integrative psychotherapy) for adolescents with BPD. At 24 months, there were no significant differences between groups across a range of outcomes including BPD severity, parasuicidal behaviors, and functioning.

Grading of the Overall Supporting Body of Research Evidence for Benefits of Psychotherapy in BPD

- Magnitude of effect: Low. When studies showed differences between treatments, these were typically low in size. However, there were few studies that used wait-list control comparison conditions and effects of BPD-specific psychotherapies may be greater if compared to no treatment.

- Risk of bias: Moderate. Although a few studies had a low risk of bias, the majority of studies had a moderate or high risk of bias.

- Applicability: The studies included individuals with BPD, but some studies excluded patients who were at significant suicide risk or who had other co-occurring conditions, which would limit applicability. Most samples were white, although some studies did not describe the race or ethnicity of participants. Study populations were primarily young adult women in the U.S., Canada, United Kingdom, Australia, or Europe. Differences in health care delivery systems may result in some differences from practice in the U.S. Most studies were conducted in outpatients and there may be less applicability to inpatient settings.

- Directness: Direct. Some of the outcomes such as functioning addressed patient-oriented outcomes whereas others such as BPD severity addressed symptom related outcomes that are also of importance to patients.

- Consistency: Inconsistent. Findings for a specific treatment differed for measured outcomes, and findings for specific outcomes differed for various psychotherapies. Overall, however, there were consistent improvements in all treatment arms on at least some outcomes even when differences between the treatment groups did not show statistically significant differences.

- Precision: Imprecise. For many of the psychotherapy comparisons, the studies did not meet the optimal information size (i.e., number of participants in a meta-analysis) and were downgraded for imprecision.

- Dose-response relationship: No information on dose-response relationships was available.

- Confounding factors (including likely direction of effect): Present. Confounding factors may increase the observed effect. Subjects and treating clinicians are aware of the treatment arm to which subjects were assigned. This may cause confounding effects due to expectancy.

- Publication bias: Unable to be assessed. The relatively small number of studies for each comparison and the heterogeneity of study designs make it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Moderate. The writing group assessed the overall strength of research evidence for psychotherapy in BPD as moderate. Although the relatively small number of studies for each comparison and the heterogeneity of study designs make it difficult to assess the strength of research evidence for specific psychotherapies, in the vast majority of studies, all treatment arms showed improvement with psychotherapy even when differences between the treatment groups did not show statistically significant differences. When compared to TAU or other active comparison arms, superiority was noted on at least some outcomes for a number of specific psychotherapies (e.g., DBT, DDP, GPM, MBT, SFT, STEPPS, TFP).

Grading of the Overall Supporting Body of Research Evidence for Harms of Psychotherapy in BPD

On the basis of the lack of data on harms in studies of psychotherapies in BPD, no grading of the body of research evidence is possible.

Pharmacotherapy

Statement 6 – Clinical Review Before Medication Initiation

APA *recommends* **(1C)** that a patient with borderline personality disorder have a review of co-occurring disorders, prior psychotherapies, other non-pharmacological treatments, past medication trials, and current medications before initiating any new medication.

Evidence for this statement comes from general principles of assessment and clinical care in psychiatric practice. A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any studies related to this recommendation in the context of BPD treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Clinical Review Before Medication Initiation in Patients with BPD

On the basis of the limitations of the evidence for assessment of a patients with possible BPD, no grading of the body of research evidence is possible.

Statement 7 – Pharmacotherapy Principles

APA *suggests* **(2C)** that any psychotropic medication treatment of borderline personality disorder be time-limited, aimed at addressing a specific measurable target symptom, and adjunctive to psychotherapy.

Evidence for this statement comes primarily from the systematic review conducted by RTI on the efficacy and comparative effectiveness of SGAs, anticonvulsants, and antidepressants in patients with BPD (Gartlehner et al. 2021). Few studies were designed to specifically address benefits of pharmacotherapy as an adjunct to psychotherapy. One small study found an adjunctive benefit of

olanzapine as an add-on to DBT (Soler et al. 2005), but small studies of adjunctive fluoxetine in patients with (Bellino et al. 2006) and without (Simpson et al. 2004) MDD did not find a benefit for BPD. Older literature suggested possible effects of lithium, the monoamine oxidase inhibitor, tranylcypromine, and the anticonvulsant, carbamazepine (Cowdry and Gardner 1988; de la Fuente and Lotstra 1994; Gardner and Cowdry 1986; Links et al. 1990). However, sample sizes were small, and BPD was diagnosed using different criteria than at present.

Second-Generation Antipsychotics Versus Placebo

Nine double-blinded RCTs evaluated the efficacy of four second-generation antipsychotics (aripiprazole, olanzapine, quetiapine extended release [ER], ziprasidone) compared with placebo (Black et al. 2014; Bogenschutz and Nurnberg 2004; Linehan et al. 2008; Nickel et al. 2006; Pascual et al. 2008; Schulz et al. 2008; Soler et al. 2005; Zanarini and Frankenburg 2001; Zanarini et al. 2011b). Overall, these studies provided data on 1,124 participants. Two studies were rated as having a moderate (Black et al. 2014; Nickel et al. 2006) risk of bias and seven as having a high risk of bias (Bogenschutz and Nurnberg 2004; Linehan et al. 2008; Pascual et al. 2008; Schulz et al. 2008; Soler et al. 2005; Zanarini and Frankenburg 2001; Zanarini et al. 2011b). Reasons for ratings of high risk of bias were lack of intention to treat analysis and high attrition. Four trials employed fixed-dose designs assessing aripiprazole (15 mg/day) (Nickel et al. 2006), olanzapine (2.5 mg/day or 5 mg/day) (Linehan et al. 2008; Zanarini and Frankenburg 2001), and quetiapine ER (150 or 300 mg/day) (Black et al. 2014); five trials used flexible-dose designs for olanzapine (2.5 to 20 mg/day) (Bogenschutz and Nurnberg 2004; Schulz et al. 2008; Soler et al. 2005; Zanarini et al. 2011b) and ziprasidone (40 to 200 mg/day) (Pascual et al. 2008). Follow-up durations ranged from eight weeks to six months. All trials, except one (Nickel et al. 2006), were funded by the pharmaceutical industry.

The majority of trial participants were female and white; mean ages across studies ranged from 21 to 34 years. Participants were moderately ill at baseline, with mean ZAN-BPD scores ranging from 14.6 to 17.7 and scores on the Clinical Global Impression scale modified for BPD from 4.3 to 4.8. Studies, in general, excluded patients with psychiatric comorbidities such as schizophrenia, MDD, alcohol or substance use disorder, or bipolar disorder.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—25 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

Three studies assessed changes in the severity of BPD on the ZAN-BPD (Black et al. 2014; Schulz et al. 2008; Zanarini et al. 2011b). Two multinational, flexible-dose trials on olanzapine, rated as having a high risk of bias, reported mixed results (Gunderson et al. 2011; Schulz et al. 2008). A three-armed trial (N=451) included a fixed-dose arm with olanzapine 2.5 mg/day (n=150), which did not achieve significant improvements compared with placebo on the ZAN-BPD (Zanarini et al. 2011b). A flexibly dosed arm showed significantly greater improvements for participants treated with olanzapine 5 to 10 mg/day than those treated with placebo, although the absolute difference in points was small (1.5 points) (Zanarini et al. 2011b). By contrast, another large trial (N=314) reported no significant differences between olanzapine 5 to 20 mg/day and placebo on the ZAN-BPD (Schulz et al. 2008).

A fixed-dose trial assessing quetiapine ER (N=95), rated as having a moderate risk of bias, reported significant improvements on the ZAN-BPD scale for low-dose (150 mg/day) but not moderate-dose (300 mg/day) treatment with quetiapine ER compared with placebo (treatment effects NR) (Black et al. 2014).

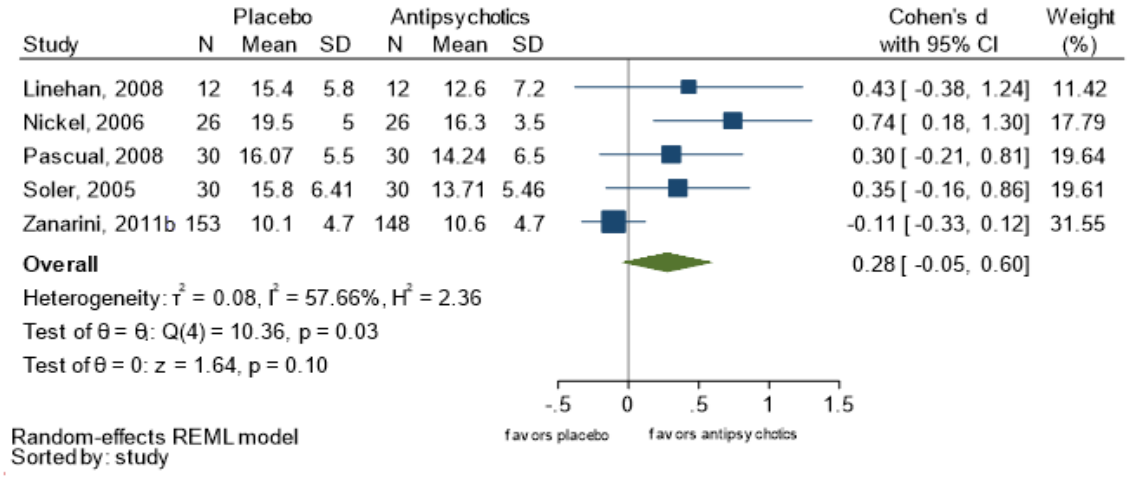
Severity of Symptoms Associated With Borderline Personality Disorder

Results assessing changes in the severity of symptoms associated with BPD reported mixed results regarding improvements in anger, impulsiveness, aggression, and depressive symptoms. A random-effects meta-analysis on the reduction of depressive symptoms favored second-generation antipsychotics over placebo but rendered no significant difference (Figure C—1).

One study (N=52), rated as having a moderate risk of bias, reported significant improvements for aripiprazole on the State-Trait Anger Expression Inventory (Nickel et al. 2006). By contrast, two RCTs (N=95 and N=60), one moderate risk of bias and the other high, detected no significant improvements for quetiapine ER (Black et al. 2014) and ziprasidone (Pascual et al. 2008) on the Barratt Impulsiveness Scale.

Regarding improvement of aggression, one moderate risk of bias RCT (N=451) (Zanarini et al. 2011b) and two RCTs (N=40 and N= 24) (Bogenschutz and Nurnberg 2004; Linehan et al. 2008) rated as having a high risk of bias reported no significant differences between olanzapine and placebo on the Modified Overt Aggression Scale. By contrast, an RCT (N=95), rated as having a moderate risk of bias, detected significant improvements for quetiapine ER compared with placebo on the Modified Overt Aggression Scale (Black et al. 2014).

Figure C—1. Standardized mean differences of changes of depressive symptoms for second-generation antipsychotics versus placebo.



Abbreviations. CI, confidence interval; N, sample size; REML, restricted maximum likelihood; SD, standard deviation.
Source. Linehan et al. 2008; Nickel et al. 2006; Pascual et al. 2008; Soler et al. 2005; Zanarini et al. 2011b

Global Impression and Functioning

Five RCTs assessed differences between second-generation antipsychotics and placebo on the Symptom Checklist-90-Revised and provided mixed results (Black et al. 2014; Bogenschutz and Nurnberg 2004; Nickel et al. 2006; Pascual et al. 2008; Zanarini et al. 2011b). Three RCTs (N=451, N=52, N=95), rated as having a moderate risk of bias, reported significantly greater improvements on the Symptom Checklist-90-Revised for participants treated with second-generation antipsychotics (aripiprazole, olanzapine, quetiapine) compared with participants in the placebo groups (Black et al. 2014; Nickel et al. 2006; Zanarini et al. 2011b). Two RCTs, one on olanzapine (N=40; Bogenschutz and Nurnberg 2004), the other on ziprasidone (N=60; Pascual et al. 2008), both rated as having a high risk of bias, favored second-generation antidepressants over placebo but rendered no significant differences between active treatments and placebo on the Symptom Checklist-90-Revised. Studies provided insufficient data for meta-analyses.

Likewise, two trials (N=40 and N=60), rated as having a high risk of bias, provided mixed results about improvements with olanzapine versus placebo on the Clinical Global Impression scale (Bogenschutz and Nurnberg 2004; Soler et al. 2005). Bogenschutz and Nurnberg (2004) reported a significant improvement with olanzapine, whereas Soler and colleagues (2005) found no significant differences in treatment effects for olanzapine and placebo on the Clinical Global Impression scale.

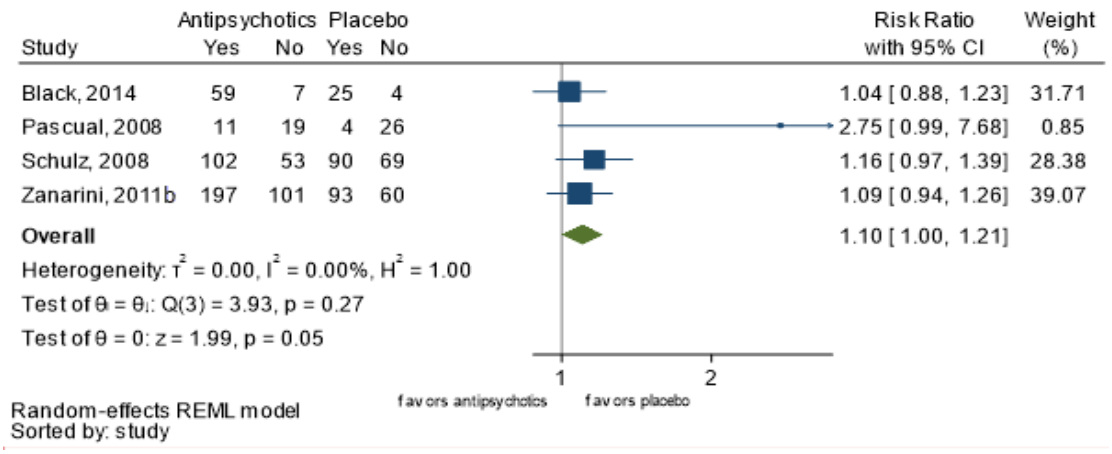
An 18-month follow-up of the trial by Nickel and colleagues (2006; N=52) reported that the significant difference on the Symptom Checklist-90-Revised between aripiprazole and placebo could be maintained (Nickel et al. 2007).

Three trials, two moderate (Black et al. 2014; Zanarini et al. 2011b) and one high risk of bias (Bogenschutz and Nurnberg 2004) with a total of 586 participants, reported no significant differences in functional capacity comparing quetiapine ER or olanzapine with placebo.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The incidence of adverse events was generally higher in the groups that received second-generation antipsychotics (Black et al. 2014; Pascual et al. 2008; Schulz et al. 2008; Zanarini et al. 2011b). A random-effects meta-analysis showed a small, but significantly higher risk of adverse events for participants treated with antipsychotics compared with placebo (Figure C—2).

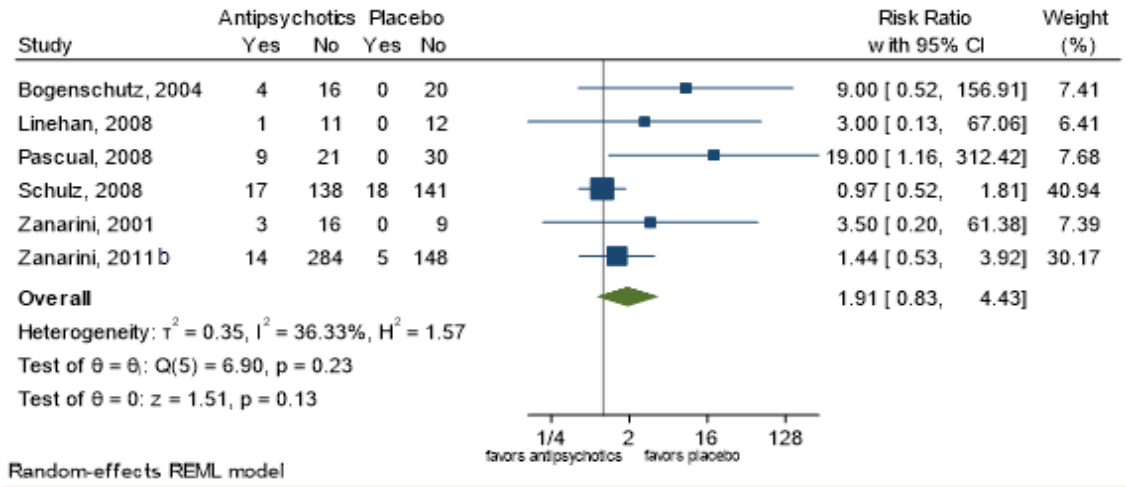
Figure C—2. Random effects meta-analysis of the incidence of adverse events comparing second-generation antipsychotics with placebo.



Abbreviations. CI, confidence interval; REML, restricted maximum likelihood.
Source. Black et al. 2014; Pascual et al. 2008; Schulz et al. 2008; Zanarini et al. 2011b

Likewise, withdrawals due to adverse events were numerically higher for participants on second-generation antipsychotics than placebo (Bogenschutz and Nurnberg 2004; Linehan et al. 2008; Moher et al. 2015; Pascual et al. 2008; Schulz et al. 2008; Zanarini and Frankenburg 2001; Zanarini et al. 2011b). A random-effects meta-analysis, however, did not reach a significant difference (Figure C—3).

Figure C—3. Random effects meta-analysis of withdrawal due to adverse events comparing second-generation antipsychotics with placebo.



Abbreviations. CI, confidence interval; REML, restricted maximum likelihood.
Source. Bogenschutz and Nurnberg 2004; Linehan et al. 2008; Pascual et al. 2008; Schulz et al. 2008, Zanarini and Frankenburg 2001; Zanarini et al. 2011b

The incidence of serious adverse events, when reported, was numerically lower for second-generation antipsychotics than placebo. Sample sizes, however, were too small to detect rare but serious adverse events reliably.

Table C—25. Certainty-of-evidence ratings for second-generation antipsychotics versus placebo.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with placebo	Difference in effect with SGA
Severity of BPD assessed with ZAN-BPD follow-up: range 8 weeks to 12 weeks	860 (3 RCTs) (Black et al. 2014; Schulz et al. 2008; Zanarini et al. 2011b)	⨁⨁◯◯ LOW[a] for no effect of SGA	-	The mean score at endpoint was 10.3 points*	mean 1.2 points lower
Anger assessed with: STAXI follow-up: mean 8 weeks	52 (1 RCT) (Nickel et al. 2006)	⨁⨁◯◯ LOW[b] for effect of SGA	-	The mean score at endpoint was 26.2 points	mean 7.7 points lower (p<0.001)
Aggression assessed with: MOAS follow-up: range 8 weeks to 12 weeks	610 (4 RCTs) (Black et al. 2014; Bogenschutz and Nurnberg 2004; Linehan et al. 2008; Zanarini et al. 2011b)	⨁⨁◯◯ LOW[a,c] for no effect of SGA	-	The mean score at endpoint was 18.6 points*	mean 14.7 points lower (ns)
Depression assessed with: Ham-D and MADRS follow-up: range 8 weeks to 21 weeks	497 (5 RCTs) (Gunderson et al. 2011; Linehan et al. 2008; Nickel et al. 2006; Pascual et al. 2008)	⨁⨁◯◯ LOW[d,e] for no effect of SGA	-	The mean score at endpoint was NR	mean 0.28 SDs (Cohen's d) greater (-0.05 to 0.60)
Impulsiveness assessed with: BIS follow-up: range 8 weeks to 12 weeks	155 (2 RCTs) (Black et al. 2014; Pascual et al. 2008)	⨁⨁◯◯ LOW[d,f] for no effect of SGA	-	The mean score at endpoint was 69.1 points*	mean 1.4 points lower (ns)
General Psychopathology assessed with: SCL-90 follow-up: range 8 weeks to 12 weeks	698 (5 RCTs) (Black et al. 2014; Bogenschutz and Nurnberg 2004; Nickel et al. 2006; Pascual et al. 2008; Zanarini et al. 2011b)	⨁⨁⨁◯ MODERATE[a] for effect of SGA	-	The mean score at endpoint was 10.3 points*	mean 1.2 points lower (ns)
Functioning assessed with: GAF and SDS follow-up: mean 8 weeks to 12 weeks	586 (3 RCTs) (Black et al. 2014; Bogenschutz and Nurnberg 2004; Zanarini et al. 2011b)	⨁⨁⨁◯ MODERATE[g] for no effect of SGA	-	The mean score at endpoint was 63.2*	mean 2.9 higher (ns)

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with placebo	Difference in effect with SGA
Incidence of Adverse Events	920 (4 RCTs) (Black et al. 2014; Pascual et al. 2008; Schulz et al. 2008; Zanarini et al. 2011b)	⊕⊕⊕◯ MODERATE[a] for higher risk with antipsychotics	**RR 1.10** (1.00 to 1.21)	571 per 1,000	57 more per 1,000 (0 fewer to 120 more)
Withdrawal due to Adverse Events	917 (5 RCTs) (Bogenschutz and Nurnberg 2004; Pascual et al. 2008; Schulz et al. 2008; Zanarini and Frankenburg 2001; Zanarini et al. 2011b)	⊕⊕◯◯ LOW[a,h] for similar risks	**RR 1.91** (0.83 to 4.43)	69 per 1,000	63 more per 1,000 (12 fewer to 237 more)
Incidence of Serious Adverse Events	957 (6 RCTs) (Black et al. 2014; Bogenschutz and Nurnberg 2004; Nickel et al. 2006; Pascual et al. 2008; Schulz et al. 2008; Zanarini et al. 2011b)	⊕◯◯◯ VERY LOW[i] for higher risk with placebo	**RR 0.46** (0.23 to 0.95)**	44 per 1,000	24 fewer per 1,000 (34 fewer to 2 fewer)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

*****Effect estimate from largest study or the study with the lowest risk of bias (Zanarini et al. 2011b or Black et al. 2014).

******Effect estimate from Zanarini et al. 2011b. The other studies reported that no serious adverse events occurred.

[a] The majority of studies were high risk of bias; downgraded 2 steps for study limitations.

[b] Small study, does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

[c] Schulz et al. 2008 assessed MOAS but did not report data; downgraded 1 step for reporting bias.

[d] At least half of studies were high risk of bias; downgraded 1 step for study limitations.

[e] Inconsistent effects, largest study shows substantially smaller treatment effect; downgraded 1 step for inconsistency.

[f] Small study, does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

[g] Does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.

[h] Few events; downgraded 1 step for imprecision.

[i] Very few events; downgraded 2 steps for imprecision.

***Abbreviations*.** BIS, Barratt Impulsiveness Scale, BPD, borderline personality disorder; CI, confidence interval; GAF, Global Assessment of Functioning; GRADE Grading of Recommendations Assessment, Development, and Evaluation; Ham-D, Hamilton Rating Scale for Depression; MADRS, Montgomery-Åsberg Depression Rating Scale; MOAS, Modified Overt Aggression Scale; NR, not reported; ns, not significant; RCT, randomized controlled trial; RR, risk ratio; SCL-90, Symptom Checklist-90, SD, standard deviation; SDS, Sheehan Disability Scale; SGA, second-generation antipsychotics; STAXI: State-Trait Anger Expression Inventory; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Grading of the Overall Supporting Body of Research Evidence for Benefits of Second-Generation Antipsychotics in BPD

- Magnitude of effect: Low. There was a small benefit of SGAs on general psychopathology but no effect on other outcomes.

- Risk of bias: High. Of the RCT studies on SGAs, two had a moderate risk of bias and seven had a high risk of bias, suggesting that the body of evidence has a high risk of bias.

- Applicability: Studies included individuals with a diagnosis of BPD, but many excluded individuals taking other medications or who had other co-occurring disorders, which are common among clinical populations. The symptom severity of patients in these trials was also less than is typically seen in clinical populations. Demographically, the study samples were primarily young adult white women. Some but not all studies included a mix of races and ethnicities. Medication doses that were studied were generally consistent with clinical practice.

- Directness: Direct. Some of the outcomes such as functioning addressed patient-oriented outcomes whereas others, such as BPD severity, addressed symptom related outcomes that are also of importance to patients.

- Consistency: Inconsistent. In many of the studies, there was at least one outcome measure that showed a statistically significant effect. However, these were not consistent for specific SGAs or for SGAs as a group.

- Precision: Imprecise. For many of the outcomes, the optimal information size (i.e., number of participants in a meta-analysis) was not met and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Insufficient information. Although two studies included treatment arms with two different doses of medication, there was inconsistent evidence for a dose response relationship.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted.

- Publication bias: Unable to be assessed. The relatively small number of studies of each SGA and the heterogeneity of study designs make it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. There is a high risk of bias of the majority of the studies, inconsistency of some of the findings, and some limits on the applicability of the studies to typical clinical practice.

Grading of the Overall Supporting Body of Research Evidence for Harms of Second-Generation Antipsychotics in BPD

- Magnitude of effect: Low. Although study withdrawals due to adverse effects were comparable for SGAs and placebo, there was a small increase in adverse effects with SGAs and a very small increase in serious adverse effects with placebo.

- Risk of bias: High. Of the RCT studies on SGAs, two had a moderate risk of bias and seven had a high risk of bias, suggesting that the body of evidence has a high risk of bias.

- Applicability: Studies included individuals with a diagnosis of BPD, but many excluded individuals taking other medications or who had other co-occurring disorders, which are common among clinical populations. Demographically, the study samples were primarily young adult white women. Some but not all studies included a mix of races and ethnicities. Medication doses that were studied were generally consistent with clinical practice.

- Directness: Direct as well as indirect. Outcomes included adverse effects and serious adverse effects but also study withdrawal due to adverse effects.

- Consistency: Inconsistent. Findings were different for adverse effects, serious adverse effects, and study withdrawal due to adverse effects.

- Precision: Imprecise. For many of the outcomes, the optimal information size (i.e., number of participants in a meta-analysis) was not met and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Insufficient information. Although two studies included treatment arms with two different doses of medication, there was inconsistent evidence for a dose response relationship.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted.

- Publication bias: Unable to be assessed. The relatively small number of studies of each SGA and the heterogeneity of study designs make it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. There is a high risk of bias of the majority of the studies, inconsistency of some of the findings, and some limits on the applicability of the studies to typical clinical practice.

Second-Generation Antipsychotic Versus Antidepressant

One industry-funded RCT (N=45; Zanarini et al. 2004c), rated as having a moderate risk of bias, assessed differences in efficacy between olanzapine (2.5 to 7.5 mg/day), fluoxetine (10 to 30 mg/day), and a combination of fluoxetine and olanzapine. The study duration was eight weeks. All trial participants were females between 18 and 40 years of age; the majority were white. The severity of disease at baseline was not reported.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—26 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

The study did not report any relevant outcomes.

Severity of Symptoms Associated With Borderline Personality Disorder

After eight weeks, participants treated with olanzapine or a combination of olanzapine and fluoxetine had significantly greater improvements in aggression (Modified Overt Aggression Scale) and depressive symptoms (Montgomery-Åsberg Depression Rating Scale) than participants treated with fluoxetine alone (Zanarini et al. 2004c).

Global Impression and Functioning

The study did not report any relevant outcomes.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report data on the incidence of adverse or serious adverse events. Only two participants (one in the fluoxetine and one in the olanzapine plus fluoxetine group) withdrew because of adverse events (Zanarini et al. 2004c).

Table C—26. Certainty-of-evidence ratings of studies comparing second-generation antipsychotic with second-generation antidepressant.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with antidepressants	Difference in effect with SGA
Olanzapine vs. Fluoxetine					
Aggression assessed with: MOAS follow-up: mean 8 weeks	30 (1 RCT) (Zanarini et al. 2004c)	⊕◯◯◯ LOW[c,b] for greater effect of olanzapine	-	The mean score at endpoint was **7.83** points	mean **4.3 points lower** (p=0.003)
Depression assessed with: MADRS follow-up: mean 8 weeks	30 (1 RCT) (Zanarini et al. 2004c)	⊕◯◯◯ LOW[b] for greater effect of olanzapine	-	The mean score at endpoint was **6.2** points	mean **1.0 points lower** (p<0.001)
Olanzapine+Fluoxetine vs. Fluoxetine					
Aggression assessed with: MOAS follow-up: mean 8 weeks	29 (1 RCT) (Zanarini et al. 2004c)	⊕◯◯◯ LOW[b] for greater effect of olanzapine+fluoxetine	-	The mean score at endpoint was **7.83** points	mean **4.8 points lower** (p<0.001)
Depression assessed with: MADRS follow-up: mean 8 weeks	29 (1 RCT) (Zanarini et al. 2004c)	⊕◯◯◯ LOW[b] for greater effect of olanzapine+fluoxetine	-	The mean score at endpoint was **6.2** points	mean **1.8 points lower** (p=0.02)
Withdrawals due to Adverse Events follow-up: mean 8 weeks	29 (1 RCT) (Zanarini et al. 2004c)	⊕◯◯◯ VERY LOW[a,b] for similar risks	**RR 0.94** (0.06 to 13.68)	71 per 1,000	**4 fewer per 1,000** (67 fewer to 906 more)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[a] Unclear how withdrawal due to adverse events was determined; downgraded 1 step for indirectness.

[b] Small study, does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. CI, confidence interval; GRADE Grading of Recommendations Assessment, Development, and Evaluation; MADRS, Montgomery-Åsberg Depression Scale; MOAS, Modified Overt Aggression Scale; RCT, randomized controlled trial; RR, risk ratio; SGA, second-generation antipsychotics.

Second-Generation Antipsychotics Versus Second-Generation Antipsychotics

One RCT (N=51; Bozzatello et al. 2017) and one retrospective cohort study (N=116; García-Carmona et al. 2021) compared second-generation antipsychotics with other second-generation antipsychotics.

The RCT, rated as having a high risk of bias, assessed differences in efficacy between asenapine (5 to 10 mg/d) and olanzapine (5 to 10 mg/d) (Bozzatello et al. 2017). The study duration was 12 weeks. All trial participants were between 18 and 50 years of age; the majority were female (63%) with race being unreported.

The high risk of bias retrospective cohort study compared the effectiveness of oral second-generation antipsychotics (not specified) and long-acting injectable second-generation antipsychotics (aripiprazole, paliperidone, risperidone) (García-Carmona et al. 2021). The study used data from 116 outpatients in Spain with follow-up data from one to three months.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—27 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

After 12 weeks, the RCT reported no significant difference on the BPD Severity Index between the asenapine and olanzapine groups (Bozzatello et al. 2017).

Severity of Symptoms Associated With Borderline Personality Disorder

After 12 weeks, the RCT reported no significant differences on the Barratt Impulsiveness Scale, the Self-Harm Inventory, and the Modified Overt Aggression Scale between the asenapine and olanzapine groups (Bozzatello et al. 2017). The retrospective cohort study reported no significant differences for suicidal behavior for individuals who received long-acting injectable antipsychotics compared with those who were on oral antipsychotics (García-Carmona et al. 2021).

Global Impression and Functioning

After 12 weeks, the RCT reported no significant difference on the Clinical Global Impression Scale–Severity between the asenapine and olanzapine groups (Bozzatello et al. 2017).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

In the RCT, the incidence of adverse events was nearly equal in both groups (5 in the olanzapine group and 4 in the asenapine group). The study did not report data on the incidence of serious adverse events. Only four participants (2 in each group) withdrew because of adverse events (Bozzatello et al. 2017).

Table C—27. Certainty-of-evidence ratings of studies comparing second-generation antipsychotics with second-generation antipsychotics.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with olanzapine	Difference in effect with asenapine
Severity of BPD assessed with: BPDSI follow-up: mean 12 weeks	51 (1 RCT) (Bozzatello et al. 2017)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was 49.12	mean 2.23 lower (ns)
Aggression assessed with: MOAS follow-up: mean 12 weeks	51 (1 RCT) (Bozzatello et al. 2017)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was 4.8	mean 1.4 higher (ns)
Impulsiveness assessed with: BIS follow-up: mean 12 weeks	51 (1 RCT) (Bozzatello et al. 2017)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was 72.9	mean 8.2 lower (ns)
Self-harm assessed with: SHI follow-up: mean 12 weeks	51 (1 RCT) (Bozzatello et al. 2017)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was 10	mean 2 lower (ns)
Global Impression assessed with: CGI-S follow-up: mean 12 weeks	51 (1 RCT) (Bozzatello et al. 2017)	⊕◯◯◯ VERY LOW[a,b] for similar effects	-	The mean score at endpoint was 3.9	mean 0.2 lower (ns)
Incidence of Adverse Events follow-up: mean 12 weeks	40 (1 RCT) (Bozzatello et al. 2017)	⊕◯◯◯ VERY LOW[a,b] for similar risks	**RR 1.38** (0.43 to 4.40)	263 per 1,000	100 more per 1,000 (150 fewer to 895 more)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[a] High attrition; downgraded 1 step for risk of bias.

[b] Small study, does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; CGI-S, Clinical Global Impression-Severity; CI, confidence interval; GRADE Grading of Recommendations Assessment, Development, and Evaluation; MOAS, Modified Overt Aggression Scale Checklist; no., number; ns, not significant; RCT, randomized controlled trial; RR, risk ratio; SHI, Self-Harm Inventory.

Anticonvulsants Versus Placebo

Nine double-blinded RCTs evaluated the efficacy of three anticonvulsant medications (divalproex sodium, lamotrigine, topiramate) compared with placebo (Crawford et al. 2018; Frankenburg and Zanarini 2002; Hollander et al. 2001; Loew et al. 2006; Moen et al. 2012; Nickel et al. 2004, 2005; Reich et al. 2009; Tritt et al. 2005). Overall, these studies provided data on 523 participants.

Two studies were rated as having a low risk of bias (Loew et al. 2006; Tritt et al. 2005), three as having a moderate risk of bias (Crawford et al. 2018; Nickel et al. 2004, 2005), and four as having a high risk of bias (Frankenburg and Zanarini 2002; Hollander et al. 2001; Moen et al. 2012; Reich et al. 2009). Reasons for ratings of high risk of bias were lack of intention-to-treat analysis and high attrition.

Four trials employed fixed-dose designs assessing lamotrigine (200 mg/day) (Tritt et al. 2005) or topiramate (200 and 250 mg/day) (Loew et al. 2006; Nickel et al. 2004, 2005); five trials used flexible-dose designs for divalproex sodium (Frankenburg and Zanarini 2002; Hollander et al. 2001; Moen et al. 2012) or lamotrigine (Crawford et al. 2018; Reich et al. 2009). Follow-up durations ranged from 8 to 52 weeks. Four trials were funded by the pharmaceutical industry (Frankenburg and Zanarini 2002; Hollander et al. 2001; Moen et al. 2012; Reich et al. 2009); the others reported no funding or were supported by public institutions.

The majority of trial participants were female and white, and mean ages ranged from 25 to 38 years. Participants were moderately ill at baseline, with mean scores on the ZAN-BPD ranging from 11.3 to 20.2. Studies, in general, excluded patients with psychiatric comorbidities such as schizophrenia, MDD, alcohol or substance use disorder, and bipolar disorder. An exception was the trial by Frankenburg and Zanarini (2002), which included participants with BPD and bipolar disorder.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—28 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

Divalproex Sodium

A small RCT (N=15; Moen et al. 2012), rated as having a high risk of bias, assessed the efficacy of divalproex sodium ER compared with placebo in participants who were already on 12-week DBT, which included individual therapy sessions, a skills training group, and telephone coaching calls. The study reported no significant differences between participants on divalproex sodium ER or placebo on the BEST scale after 12 weeks of treatment.

Lamotrigine

The publicly funded LABILE (Lamotrigine and Borderline Personality Disorder: Investigating Long-Term Effects) trial (N=276; Crawford et al. 2018), rated as having a moderate risk of bias, and a small, industry-funded RCT (N=28; Reich et al. 2009), rated as having a high risk of bias, assessed the efficacy of lamotrigine (200 to 400 mg/day) compared with placebo on the ZAN-BPD. Both trials reported no significant differences between participants in the lamotrigine and the placebo groups after 12 weeks of treatment. The primary endpoint of the LABILE trial was at 52 weeks, which also yielded no significant difference on the ZAN-BPD between treatment groups (Crawford et al. 2018).

Topiramate

None of the included trials reported relevant outcomes.

Severity of Symptoms Associated With Borderline Personality Disorder

Divalproex Sodium

Two small RCTs, rated as having a high risk of bias, reported results regarding the efficacy of divalproex sodium (flexible dose to achieve serum levels of 80 and of 50 to 100 mg/l, respectively) to reduce aggression (Frankenburg and Zanarini 2002; Hollander et al. 2001). One trial (N=30; Frankenburg and Zanarini 2002) reported significant improvements for divalproex sodium compared with placebo on the Modified Overt Aggression Scale and the Symptom Checklist-90-Revised subscale for anger and hostility after 24 weeks of treatment. This study enrolled participants with BPD and bipolar II disorder. The other trial (N=16; Hollander et al. 2001) also favored divalproex sodium over placebo but found no significant differences on the Aggression Questionnaire and the Modified Overt Aggression Scale after 10 weeks.

An RCT (N=15; Moen et al. 2012), rated as having a high risk of bias, reported no significant differences between participants on divalproex sodium ER or placebo on the Barratt Impulsiveness Scale after 12 weeks of treatment.

Lamotrigine

Three trials assessed improvements of BPD-specific symptoms under lamotrigine treatment (Crawford et al. 2018; Reich et al. 2009; Tritt et al. 2005). The LABILE trial (N=276; Crawford et al. 2018), rated as having a moderate risk of bias, reported no significant differences in alcohol or other substance use between participants treated with lamotrigine or placebo. In a Cochrane review of pharmacological treatments for BPD, the evidence for lamotrigine was assessed as being very uncertain in terms of effects on self-harm (Stoffers-Winterling et al. 2022).

An RCT in 27 female participants with BPD, rated as having a low risk of bias, showed significant improvements in anger as measured on four out of five subscales on the State-Trait Anger Expression Inventory after eight weeks of treatment (Tritt et al. 2005). The subscale assessing the tendency to repress anger did not improve significantly.

Likewise, a small RCT with 28 participants, rated as having a high risk of bias, reported significantly greater reductions on the Affective Lability Scale for the lamotrigine group compared with the placebo treatment group (Reich et al. 2009).

Topiramate

Two RCTs (N=31 and 44) with similar protocols conducted by the same author team, rated as having a moderate risk of bias, investigated the efficacy of topiramate (titrated from 50 mg/day to 250 mg/day) to reduce anger and aggression in women (Nickel et al. 2004) and men (Nickel et al. 2005) with BPD. After eight weeks, both women and men experienced significant improvements in four out of five subscales of the State-Trait Anger Expression Inventory. In both trials, the subscale assessing the tendency to repress anger did not improve significantly (Nickel et al. 2004, 2005).

Global Impression and Functioning

Divalproex Sodium

Two very small RCTs (N=16 and N=15), rated as having a high risk of bias, reported no significant differences between divalproex sodium and placebo on the Clinical Global Impression-Improvement scale and the Symptom Checklist-90-Revised after 10 and 12 weeks of treatment (Hollander et al. 2001; Moen et al. 2012).

Lamotrigine

The LABILE trial (N=276; Crawford et al. 2018) reported no significant differences on the Social Functioning Questionnaire between participants treated with lamotrigine or placebo after 52 weeks of treatment.

Topiramate

One RCT (N=56; Loew et al. 2006), rated a having a low risk of bias, assessed the efficacy of topiramate (titrated from 50 mg/day to 200 mg/day) in women with BPD ages 18 to 35 years. After 10 weeks, participants in the topiramate group had significantly greater improvements on the Global Severity Index of the Symptom Checklist-90-Revised, the Short Form-36, and the Inventory of Interpersonal Problems.

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

Divalproex Sodium

None of the three included studies reported on the incidence of adverse events. Two trials reported similar proportions of withdrawals because of adverse events between divalproex sodium and placebo treatment groups (Frankenburg and Zanarini 2002; Hollander et al. 2001).

Lamotrigine

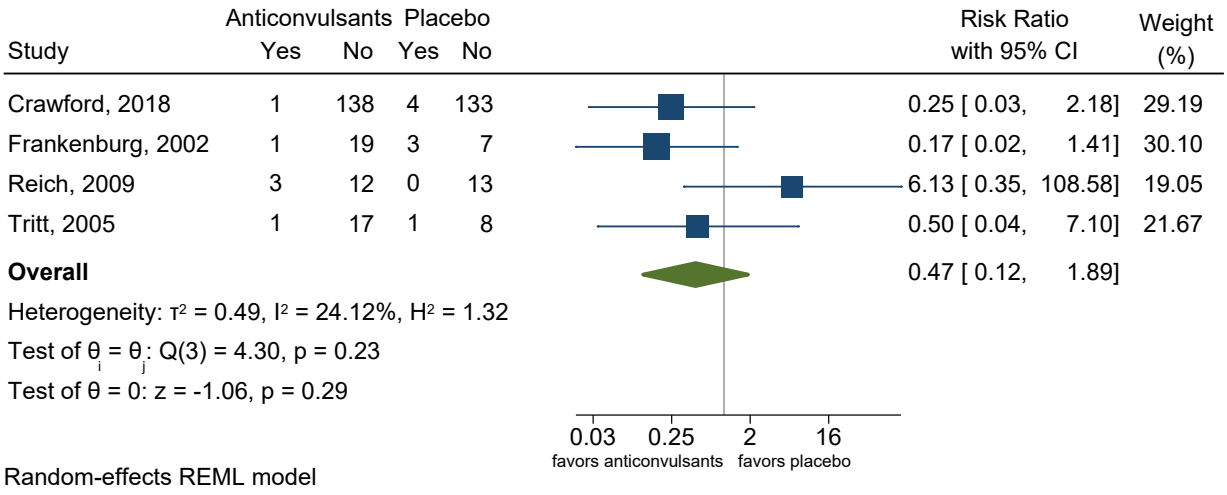
The incidence of adverse events, serious adverse events, and withdrawals due to adverse events was similar between lamotrigine and placebo treatment groups (Crawford et al. 2018; Reich et al. 2009; Tritt et al. 2005).

Topiramate

None of the trials reported the incidence of adverse events or serious adverse events. Two publications stated that no participants withdrew because of adverse events during eight weeks of treatment (Nickel et al. 2004, 2005).

A meta-analysis of anticonvulsant medications as a class rendered no significant differences in withdrawals because of adverse events after 8 to 52 weeks of treatment (Figure C—4).

Figure C—4. Random effects meta-analysis of withdrawal due to adverse events comparing anticonvulsant medications with placebo.



***Abbreviations*.** CI, confidence interval; REML, restricted maximum likelihood.
Source. Crawford et al. 2018; Frankenburg and Zanarini 2002; Reich et al. 2009; Tritt et al. 2005

Table C—28. Certainty-of-evidence ratings of studies comparing anticonvulsants with placebo.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with placebo	Difference in effect with anticonvulsants
Divalproex sodium					
Severity of BPD assessed with: BEST follow-up: mean 12 weeks	15 (1 RCT) (Moen et al. 2012)	⊕◯◯◯ VERY LOW[a,b] for no effect of divalproex sodium	-	The mean score at endpoint was 30.0 points	mean 1.3 points lower (ns)
Aggression assessed with: MOAS; SCL-90-R subscale for anger and hostility follow-up: range 10 weeks to 24	46 (2 RCTs) (Frankenburg and Zanarini 2002; Hollander et al. 2001)	⊕◯◯◯ VERY LOW[a,c,d] for effect of divalproex sodium	-	The mean score on MOAS was 3.2 points*	mean 0.6 points lower (p=0.03)
Impulsiveness assessed with: BIS-Motor follow-up: mean 12 weeks	15 (1 RCT) (Moen et al. 2012)	⊕◯◯◯ VERY LOW[a,b] for no effect of divalproex sodium	-	The mean score at endpoint was 18.2 points	mean 5.7 points higher (ns)
General Psychopathology assessed with: SCL-90-R, CGI-I follow-up: range 10 weeks to 12 weeks	31 (2 RCTs) (Hollander et al. 2001; Moen et al. 2012)	⊕◯◯◯ VERY LOW[a,d] for no effect of divalproex sodium	-	The mean score at endpoint on SCL-90 was 114.2 points*	mean 22.8 points higher (ns)
Withdrawals due to adverse events follow-up: range 10 to 24 weeks	46 (2 RCTs) (Frankenburg and Zanarini 2002; Hollander et al. 2001)	⊕◯◯◯ VERY LOW[a,d] for similar risks	**RR 0.26** (0.03 to 2.35)	136 per 1,000*	**101 fewer per 1,000** (132 fewer to 184 more; ns)
Lamotrigine					
Severity of BPD assessed with: ZAN-BPD follow-up: range 12 weeks to 52 weeks	304 (2 RCTs) (Crawford et al. 2018; Reich et al. 2009)	⊕⊕⊕◯ MODERATE[e] for no effect of lamotrigine	-	The mean score at endpoint was 11.5 points*	mean 0.5 points lower (ns)

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with placebo	Difference in effect with anticonvulsants
Affective lability assessed with: ALS follow-up: mean 12 weeks	28 (1 RCT) (Reich et al. 2009)	⨁◯◯◯ VERY LOW[b,f] for effect of lamotrigine	-	The mean at endpoint score was 1.52 points	mean 0.27 points lower (p=0.012)
Alcohol and substance use assessed with ASSIST follow-up: mean 52 weeks	160 (1 RCT) (Crawford et al. 2018)	⨁⨁◯◯ LOW[b] for no effect of lamotrigine	-	The mean score at endpoint was 23 points	mean 4 points higher (ns)
Anger assessed with: STAXI follow-up: mean 8 weeks	27 (1 RCT) (Tritt et al. 2005)	⨁⨁◯◯ LOW[b] for effect of lamotrigine	-	The mean score at endpoint was **NR**	**NR** (4 of 5 subscales significantly improved)
Functioning assessed with: SFQ follow-up: mean 52 weeks	276 (1 RCT) (Crawford et al. 2018)	⨁⨁⨁◯ MODERATE[e,g] for no effect of lamotrigine	-	The mean score at endpoint was 12.3 points	mean 0.1 points higher (ns)
Incidence of Adverse Events follow-up: range 10 weeks to 52 weeks	304 (2 RCTs) (Crawford et al. 2018; Reich et al. 2009)	⨁⨁◯◯ LOW[g] for similar risks	**RR 0.86** (0.71 to 1.03)	630 per 1,000*	**88 fewer per 1,000** (183 fewer to 19 more; ns)
Incidence of Serious Adverse Events follow-up: mean 52 weeks	276 (1 RCT) (Crawford et al. 2018)	⨁⨁◯◯ LOW[h] for similar risks	**RR 0.82** (0.52 to 1.31)	230 per 1,000	**41 fewer per 1,000** (111 fewer to 71 more; ns)
Withdrawal due to Adverse Events follow-up: range 10 weeks to 52 weeks	328 (3 RCTs) (Crawford et al. 2018; Reich et al. 2009; Tritt et al. 2005)	⨁◯◯◯ VERY LOW[i,h] for similar risks	**RR 3.79** (0.82 to 17.57)	12 per 1,000	**35 more per 1,000** (2 fewer to 206 more; ns)
Topiramate					

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with placebo	Difference in effect with anticonvulsants
Anger assessed with: STAXI follow-up: mean 8 weeks	75 (2 RCTs) (Nickel et al. 2004; Nickel et al. 2005)	⨁⨁◯◯ LOW[d] for effect of topiramate	-	The mean score at endpoint was NR	NR (4 of 5 subscales significantly improved)
General Psychopathology assessed with: SCL-90 follow-up: range 8 weeks to 12 weeks	56 (1 RCT) (Loew et al. 2006)	⨁⨁◯◯ LOW[b] for effect of topiramate	-	The mean score at endpoint was 70.1 points	mean 5.9 points lower (p<0.001)
Withdrawal due to Adverse Events	75 (2 RCTs) (Nickel et al. 2004; Nickel et al. 2005)	⨁◯◯◯ VERY LOW[d,j] for similar risks	RR 1.95 (0.77 to 4.94)	0 per 1,000	0 fewer per 1,000 (0 fewer to 0 fewer)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
*Effect estimate from largest study or the study with the lowest risk of bias.
[a] High attrition; downgraded 1 step for risk of bias.
[b] Small study, does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[c] Conflicting results of two studies; downgraded 1 step for inconsistency.
[d] Small studies, do not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.
[e] Sample size probably does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 1 step for imprecision.
[f] Trial with high risk of bias¸ downgraded 1 step for risk of bias.
[g] Few events; downgraded 2 steps for imprecision.
[h] Very few events; downgraded 2 steps for imprecision.
[i] Proportions vary substantially; downgraded 1 step for inconsistency.
[j] One study does not report data on withdrawal due to adverse events; downgraded 1 step for outcomes reporting bias.
Abbreviations. ALS, Affective Lability Scale; ASSIST, Alcohol, Smoking, and Substance Involvement Screening Test; BEST, Borderline Evaluation of Severity Over Time; BIS-Motor, Barratt Impulsiveness Scale-Motor; BPD, borderline personality disorder; CGI-I, Clinical Global Impressions–Improvement; CI, confidence interval; GRADE Grading of Recommendations Assessment, Development, and Evaluation; MOAS, Modified Overt Aggression Scale Checklist; No., number; NR, not reported; ns, not significant; RCT, randomized controlled trial; RR, risk ratio; SCL-90, Symptom Checklist-90; SCL-90-R, Symptom Checklist-90-Revised; SFQ, Social Functioning Questionnaire; STAXI, State-Trait Anger Expression Inventory; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Grading of the Overall Supporting Body of Research Evidence for Benefits of Divalproex in BPD

- Magnitude of effect: Minimal. There was a very small benefit of divalproex on aggression but no effect on other outcomes.

- Risk of bias: High. Of the RCT studies on divalproex, both had a high risk of bias.

- Applicability: Studies were conducted in the U.S. and included individuals with a diagnosis of BPD, but excluded individuals with co-occurring disorders or who were suicidal. Demographically, the study samples were primarily young adult white women, but a mix of races and ethnicities were included. Medication doses that were studied were smaller than in usual clinical practice, limiting the generalizability of the findings.

- Directness: Indirect. Outcomes in one study were not well delineated; in the other study, outcomes were either global or addressed aggressive behavior.

- Consistency: Consistent. Studies were generally consistent and, with the exception of aggressive behavior in one study, showed significant effects of divalproex.

- Precision: Imprecise. The optimal information size (i.e., number of participants in a meta-analysis) was not met due to small samples, and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Unable to be assessed. Studies did not include information on dose-response relationships.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted but some may have been present due to the high risk of bias in the study design.

- Publication bias: Unable to be assessed. The small number of studies makes it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. There is a high risk of bias in both studies, inconsistency of some of the findings, and limits on the applicability of the studies to typical clinical practice.

Grading of the Overall Supporting Body of Research Evidence for Harms of Divalproex in BPD

- Magnitude of effect: None noted. Study withdrawal rates due to adverse effects were comparable for placebo and divalproex in one study. No data on adverse effects was reported in the other study.

- Risk of bias: High. Of the RCT studies on divalproex, both had a high risk of bias.

- Applicability: Studies were conducted in the U.S. and included individuals with a diagnosis of BPD, but excluded individuals with co-occurring disorders or who were suicidal. Demographically, the study samples were primarily young adult white women, but a mix of races and ethnicities were

included. Medication doses that were studied were smaller than in usual clinical practice, limiting the generalizability of the findings.

- Directness: Indirect. Outcomes in one study were not well delineated; in the other study, outcomes were either global or addressed aggressive behavior.

- Consistency: Consistent. Studies were generally consistent and with the exception of aggressive behavior in one study, showed significant effects of divalproex.

- Precision: Imprecise. The optimal information size (i.e., number of participants in a meta-analysis) was not met due to small samples, and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Unable to be assessed. Studies did not include information on dose-response relationships.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted but some may have been present due to the high risk of bias in the study design.

- Publication bias: Unable to be assessed. The small number of studies makes it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. There is a high risk of bias in both studies, inconsistency of some of the findings, and limits on the applicability of the studies to typical clinical practice.

Grading of the Overall Supporting Body of Research Evidence for Benefits of Lamotrigine in BPD

- Magnitude of effect: Minimal. There was a very small benefit of lamotrigine on affective lability and anger, in one small study each, but no effect on other outcomes. In one large study that assesses BPD severity and functioning, lamotrigine had no significant effect.

- Risk of bias: Moderate. Of the RCT studies on lamotrigine, the largest study had a moderate risk of bias whereas the two smaller studies had a low and a high risk of bias.

- Applicability: Studies were conducted in the U.S., the U.K, Germany, and Austria. They included individuals with a diagnosis of BPD, but the smaller studies excluded individuals with co-occurring disorders or who were suicidal. Demographically, the study samples were primarily young adult white women but in the largest study 25% of participants were male and 11% non-white race. Medication doses that were studied were comparable to those used in usual clinical practice.

- Directness: Direct. The primary outcome in the largest study was BPD severity, although the smaller studies had indirect measures of anger and affective lability as primary outcomes.

- Consistency: Inconsistent. The smaller studies showed some benefits on affective lability and anger whereas the larger study showed no effect of lamotrigine on BPD severity, self-harm, or functioning.

- Precision: Imprecise. The optimal information size (i.e., number of participants in a meta-analysis) was not met due to small samples in two studies, and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Unable to be assessed. Studies did not include information on dose-response relationships.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted.

- Publication bias: Unable to be assessed. The small number of studies makes it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. For most outcomes, data was only available from a single study. There was also inconsistency of some of the findings, and variability in the risk of bias in the studies.

Grading of the Overall Supporting Body of Research Evidence for Harms of Lamotrigine in BPD

- Magnitude of effect: None detected. There was a similar effect of lamotrigine on withdrawal due to adverse effects as well as on the incidence of adverse effects and serious adverse effects.

- Risk of bias: Moderate. Of the RCT studies on lamotrigine, the largest study had a moderate risk of bias whereas the two smaller studies had a low and a high risk of bias.

- Applicability: Studies were conducted in the U.S., the U.K, Germany, and Austria. They included individuals with a diagnosis of BPD, but the smaller studies excluded individuals with co-occurring disorders or who were suicidal. Demographically, the study samples were primarily young adult white women but in the largest study 25% of participants were male and 11% non-white race. Medication doses that were studied were comparable to those used in usual clinical practice.

- Directness: Direct. The studies measured the incidence of adverse effects and serious adverse effects.

- Consistency: Consistent. The studies were consistent in showing a comparable incidence of adverse effects and serious adverse effects as well as similar rates of study withdrawal due to adverse effects.

- Precision: Imprecise. The optimal information size (i.e., number of participants in a meta-analysis) was not met due to small samples in two studies, and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Unable to be assessed. Studies did not include information on dose-response relationships.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted.

- Publication bias: Unable to be assessed. The small number of studies makes it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. Based on the variability in the risk of bias in the studies and imprecision, the overall strength of research evidence was rated as low.

Grading of the Overall Supporting Body of Research Evidence for Benefits of Topiramate in BPD

- Magnitude of effect: Minimal. There was a very small benefit of topiramate on general psychopathology in one small study and anger in two small studies.

- Risk of bias: Moderate. Of the RCT studies on topiramate, two had a moderate risk of bias and one had a low risk of bias.

- Applicability: Studies were conducted in Germany and Austria. They included individuals with a diagnosis of BPD, but the smaller studies excluded individuals with co-occurring disorders or who were suicidal. Demographically, the study samples were primarily young adults, with only women in two studies and only men in the third study. No data was obtained on race or ethnicity. Medication doses that were studied were comparable to those used in usual clinical practice.

- Directness: Indirect. The primary outcomes were symptom measures but not specific to BPD severity or functioning.

- Consistency: Consistent. The studies were consistent in showing some minimal benefits of topiramate.

- Precision: Imprecise. The optimal information size (i.e., number of participants in a meta-analysis) was not met due to small samples, and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Unable to be assessed. Studies did not include information on dose-response relationships.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted.

- Publication bias: Unable to be assessed. The small number of studies makes it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. Two of the studies had a moderate risk of bias, results were downgraded for imprecision, and there were significant issues with applicability of the study samples.

Grading of the Overall Supporting Body of Research Evidence for Harms of Topiramate in BPD

- Magnitude of effect: None noted. No study withdrawals due to adverse effects were noted in the two studies that examined this outcome.

- Risk of bias: Moderate. Of the RCT studies on topiramate, two had a moderate risk of bias and one had a low risk of bias.

- Applicability: Studies were conducted in Germany and Austria. They included individuals with a diagnosis of BPD, but the smaller studies excluded individuals with co-occurring disorders or who were suicidal. Demographically, the study samples were primarily young adults, with only women in two studies and only men in the third study. No data was obtained on race or ethnicity. Medication doses that were studied were comparable to those used in usual clinical practice.

- Directness: Indirect. The primary outcome related to adverse effects was study withdrawals.

- Consistency: Consistent. The two studies that measured withdrawals due to adverse effects were consistent in showing no study withdrawals for this reason.

- Precision: Imprecise. The optimal information size (i.e., number of participants in a meta-analysis) was not met due to small samples, and the certainty of evidence was downgraded for imprecision.

- Dose-response relationship: Unable to be assessed. Studies did not include information on dose-response relationships.

- Confounding factors (including likely direction of effect): Not identified. No specific confounding effects were noted.

- Publication bias: Unable to be assessed. The small number of studies makes it difficult to assess publication bias. However, publication bias seems possible because of the tendency for negative clinical trial results to go unpublished.

- Overall strength of research evidence: Low. Two of the studies had a moderate risk of bias, results were downgraded for imprecision, and there were significant issues with applicability of the study samples.

Antidepressants Versus Placebo

One industry-funded RCT (N=25; Simpson et al. 2004), rated as having a high risk of bias, assessed differences in efficacy between fluoxetine (20 to 40 mg/day) and placebo. The study duration was 12 weeks. All trial participants were female; the majority were white. Participants in both treatment groups received individual DBT and were part of 2-hour weekly skills groups.

Detailed information on main study characteristics and treatment effects is presented in Appendix D. Table C—29 presents certainty-of-evidence ratings.

Severity of Borderline Personality Disorder

The study did not report any relevant outcomes.

Severity of Symptoms Associated With Borderline Personality Disorder

After a mean of 10 weeks, authors reported no significant difference between fluoxetine and placebo on the State-Trait Anger Expression Inventory and the Modified Overt Aggression Scale (Simpson et al. 2004).

Global Impression and Functioning

After 10 weeks, there were no significant differences between both groups in the Global Assessment of Functioning scale (Simpson et al. 2004).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report any relevant adverse events.

Table C—29. Certainty-of-evidence ratings of studies comparing antidepressants with placebo.

Outcomes	No. of participants (studies)	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Effect with placebo	Difference in effect second-generation antidepressants
Anger assessed with: STAXI follow-up: mean 10 weeks	25 (1 RCT) (Simpson et al. 2004)	⨁◯◯◯ VERY LOW[a,b] for no effect of fluoxetine	-	The mean score at endpoint was 27.6 points	mean 7.1 lower (ns)
Aggression assessed with: MOAS follow-up: mean 10 weeks	25 (1 RCT) (Simpson et al. 2004)	⨁◯◯◯ VERY LOW[a,b] for no effect of fluoxetine	-	The mean score at endpoint was NR	NR (ns)
Functioning assessed with: GAF follow-up: mean 10 weeks	25 (1 RCT) (Simpson et al. 2004)	⨁◯◯◯ VERY LOW[a,b] for no effect of fluoxetine	-	The mean score at endpoint was 59.3 points	mean 0.6 higher (ns)

Note. The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[a] No intention-to-treat analysis; downgraded 1 step for risk of bias.

[b] Small study, does not meet optimal information size (i.e., number of participants in a meta-analysis); downgraded 2 steps for imprecision.

Abbreviations. CI, confidence interval; GAF, Global Assessment of Functioning; GRADE Grading of Recommendations Assessment, Development, and Evaluation; MOAS, Modified Overt Aggression Scale; NR, not reported; ns, not significant; RCT, randomized controlled trial; STAXI, State-Trait Anger Expression Inventory.

Grading of the Overall Supporting Body of Research Evidence for Antidepressants in BPD

Only a single study met inclusion criteria related to antidepressants in BPD and, thus, no grading of the body of research evidence is possible.

Repetitive Transcranial Magnetic Stimulation Versus Sham Treatment

One RCT (N=9; Cailhol et al. 2014), rated as having a moderate risk of bias, assessed differences in efficacy between 10 sessions of repetitive Transcranial Magnetic Stimulation (rTMS) and sham rTMS. The study duration was three months. The majority of trial participants were females between 20 and 45 years of age with race being unreported. The severity of disease at baseline was reported by the BPD Severity Index. The study was publicly funded. Detailed information on main study characteristics and treatment effects is presented in Appendix D.

Severity of Borderline Personality Disorder

After three months, there were no significant differences on the BPD Severity Index between the rTMS and the sham rTMS groups (Cailhol et al. 2014).

Severity of Symptoms Associated With Borderline Personality Disorder

The study did not report any relevant outcomes (Cailhol et al. 2014).

Global Impression and Functioning

After three months, differences on the Symptom Checklist-90 and the Global Assessment Scale favored rTMS over sham treatment, but the difference did not reach statistical significance because of the small sample size (N=9) (Cailhol et al. 2014).

Incidence of Adverse Events, Serious Adverse Events, and Withdrawal Due to Adverse Events

The study did not report data on the incidence of adverse or serious adverse events. No participants withdrew due to adverse events (Cailhol et al. 2014).

Grading of the Overall Supporting Body of Research Evidence for TMS in BPD

Only a single study met inclusion criteria related to TMS in BPD and, thus, no grading of the body of research evidence is possible.

Statement 8 – Pharmacotherapy Review

APA *recommends* **(1C)** that a patient with borderline personality disorder have a review and reconciliation of their medications at least every 6 months to assess the effectiveness of treatment and identify medications that warrant tapering or discontinuation.

Evidence for this statement comes from general principles of clinical care in psychiatric practice. In addition, medication reconciliation and de-prescribing, where indicated, are recommended best practices in hospital as well as outpatient settings (Institute for Safe Medication Practice 2023; The Joint Commission 2022). A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any studies related to this recommendation in the context of BPD treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Pharmacotherapy Review in Patients with BPD

On the basis of the limitations of the evidence for pharmacotherapy review in patients with possible BPD, no grading of the body of research evidence is possible.

Appendix D. Evidence Tables for Individual Studies Supporting Guideline Statements

Psychoeducation

Psychoeducation vs. Wait-list

Table D—1. Study characteristics and main results of psychoeducation compared with wait-list control.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Zanarini and Frankenburg (2008)	Design: RCT Setting: NR Country: United States Funding: Eli Lilly	N=50 G1 (20): Delayed psychoeducation G2 (30): Psychoeducation 12 weeks	Inclusion: Females; age 18-30 years; met DIB-R and DSM-IV criteria for BPD Exclusion: Currently in any type of psychiatric treatment; schizophrenia, schizoaffective disorder, bipolar I disorder, or SUD	Mean (SD) age: 19 (1.4) % Female: 100 % Race/ethnicity: White: 86	Primary outcome: NR No significant difference between G1 and G2 on ZAN-BPD Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 0% Differential attrition: <10 percentage points	Moderate
Zanarini et al. (2018)	Design: RCT Setting: NR Country: United States Funding: NIMH, government funding	N=80 G1 (40): No Psychoeducation G2 (40): Internet-based psychoeducation 12 weeks Follow-up: 12 months	Inclusion: Met DIB-R and DSM-IV criteria for BPD Exclusion: Schizophrenia, schizoaffective disorder, or intellectual disability; acutely suicidal or fully manic at time of assessment; current physical condition that can cause serious psychiatric symptoms (lupus, MS, etc.); serious substance abuse	Mean (SD) age: G1: 21 (3.1) G2: 22 (3.7) % Female: 100 % Race/ethnicity: White: 69 Black: 11 Hispanic: 10 Asian: 8 Other: 3	Primary outcome: NR G2 significantly more effective than G1 on SAS (0.5 vs. 0.09, p=0.049) after 12 weeks G2 significantly more effective than G1 on ZAN-BPD scale after 12 months (4.46 vs. 0.0, p=0.035); no significant differences on any other outcome measures after 12 months Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 4% Differential attrition: <10 percentage points	Moderate

Abbreviations. AE, adverse event; BPD, Borderline Personality Disorder; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; MS, multiple sclerosis; N, sample size; NIMH, National Institute of Mental Health; NR, not reported; RCT, randomized controlled trial; SAS, Social Adjustment Scale; SD, standard deviation; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder; SUD, substance use disorder.

Psychosocial Interventions

Interpersonal Psychotherapy vs. Wait-List Plus Clinical Management

Table D—2. Study characteristics and main results of interpersonal psychotherapy compared with wait-list plus clinical management.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bozzatello and Bellino (2020)	Design: RCT Setting: Outpatient, single center Country: Italy Funding: Government	N=43 G1 (21): WL plus clinical management G2 (22): IPT adapted for treating BPD: 50-minute sessions over 40 weeks; 22 sessions in the first 20 weeks and 20 sessions in the last 20 weeks 10 months	Inclusion: Age 18-60 years attending the Center for Personality Disorders who met the DSM-5 criteria for BPD Exclusion: Dementia or other cognitive disorders, schizophrenia or other psychotic disorders, or bipolar disorders; co-occurring major depressive episode and/or substance abuse; taken psychotropic medications and/or psychotherapy 3 months previously; females of childbearing age if they were not using birth control	Median age: 35 % Female: 67 % Race/ethnicity: NR	Primary outcome: NR G2 significantly lower severity of BPD on BPDSI (36.1 vs. 44.6, p=0.01), symptom scores on BIS-11 (64.8 vs. 77.4, p=0.03), and functioning scores on CGI-S (3.1 vs. 4.1, p=0.009) and SOFAS (68.2 vs. 57.1, p=0.02) than G1 after 10 months, but not on the SHI Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 14% Differential attrition: ≤10 percentage points	Moderate

Abbreviations. AE, adverse event; BIS-11, Barratt Impulsiveness Scale, version 11; BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; CGI-S, Clinical Global Impression-Severity; DSM-5, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; G1, Group 1; G2, Group 2; IPT, interpersonal psychotherapy; N, sample size; NR, not reported; RCT, randomized controlled trial; SOFAS, Social Occupational Functioning Assessment Scale; SHI, Self-Harm Inventory; WL, wait-list.

Interpersonal Psychotherapy Plus Fluoxetine vs. Clinical Management Plus Fluoxetine

Table D—3. Study characteristics and main results of interpersonal psychotherapy plus fluoxetine compared with clinical management plus fluoxetine.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bellino et al. (2006)	Design: RCT Setting: Outpatient, single center Country: Italy Funding: None	N=39 G1 (19): Clinical management plus fluoxetine 20-40 mg/day; initial fixed 20 mg/day with opportunity to increase to 40 mg/day beginning week 2 G2 (20): IPT in weekly 1-hour sessions plus fluoxetine 20-40	Inclusion: DSM-IV BPD diagnosis; met criteria for major depressive episode Exclusion: Lifetime diagnosis of delirium, dementia, amnestic or other cognitive disorders, or schizophrenia or other psychotic disorders; major depressive episode as an expression of bipolar	Mean (SD) age: 26 (3.7) % Female: 60 (Reported as: The ratio of men to women was 3 to 5) % Race/ethnicity: NR	Primary outcome: NR G2 significantly more effective than G1 for improving symptoms of depression (measured by the Ham-D [9.1 vs. 12, p=0.005]) No significant differences between G2 and G1 for anxiety	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		mg/day plus; initial fixed 20 mg/day with opportunity to increase to 40 mg/day beginning week 2 24 weeks	disorder; current substance abuse disorder; treatment with psychotropic drugs or psychotherapy during 2 months prior to study; female patients not using adequate birth control		for clinical global impressions (measured by CGI-S) or anxiety (measured by Ham-A) Attrition: 17.9% (7/39) G1: 20.0% (4/20) G2: 15.8% (3/19)	
Bellino et al. (2010)	Design: RCT Setting: Outpatient, single center Country: Italy Funding: None	N=55 G1 (28): Clinical management plus fluoxetine 20-40 mg/day; initial fixed 20 mg/day with the ability to increase to a maximum dose of 40 mg/day beginning in week 2 plus 15-20 minutes of clinical management every 2 weeks, dealing with clinical issues G2 (27): IPT plus fluoxetine 20-40 mg/day; initial fixed 20 mg/day with the ability to increase to a maximum dose of 40 mg/day beginning in week 2 plus IPT adapted to BPD according to Markowitz 's model (IPT-BPD) 32 weeks	Inclusion: DSM-IV-TR BPD diagnosis Exclusion: Lifetime diagnosis of delirium, amnestic disorder, or other cognitive disorders, schizophrenia or other psychotic disorders, bipolar disorder, or Axis I or II disorders; those receiving psychotropic drugs in the last 2 months and/or psychotherapy in the last 6 months; those of childbearing age who were not using an adequate method of birth control	Mean (SD) age: G1: 26 (7.2) G2: 26 (6.4) % Female: 67 % Race/ethnicity: NR	Primary outcome: NR No significant differences between G2 and G1 on BPDSI, Ham-A, Ham-D, CGI-S, and SOFAS Attrition: 20% (11/55) G1: 21.4% (6/28) G2: 18.5% (5/27)	Moderate

Abbreviations. BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; CGI, Clinical Global Impression Scale; CGI-S, Clinical Global Impression-Severity; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*; G1, Group 1; G2, Group 2; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; IPT, interpersonal psychotherapy; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SOFAS, Social Occupational Functioning Assessment Scale.

Acceptance and Commitment Therapy vs. Treatment as Usual

Table D—4. Study characteristics and main results of acceptance and commitment therapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Morton et al. (2012)	Design: RCT Setting: Outpatient, multicenter Country: Australia Funding: NR	N=41 G1 (20): TAU: case management provided mostly by public mental health services G2 (21): ACT: 12 group sessions in psychoeducational format (2 hours/week) 13 weeks	Inclusion: ≥4 criteria of BPD (DSM-IV Axis I and Axis II diagnoses using the SCID-I and the SCID-II, respectively); a registered client of a public sector adult mental health service Exclusion: Current positive or negative psychotic symptoms other than reactive psychotic symptoms associated with BPD; a significant risk of violent and/or threatening behavior to other participants; intellectual disability, cognitive impairment, or difficulty speaking English, severe enough to interfere with participation	Mean (SD) age: G1: 34 (9.0) G2: 36 (9.3) % Female: 93 % Race/ethnicity: NR	Primary outcome: BEST at 13 weeks G2 significantly more effective than G1 on BEST (-11.8 vs. -2.4, p=0.028), BHS (-4.7 vs. +0.7[a], p=0.006), and DERS (-18.7 vs. +5.6[a], p=0.008) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 22% Differential attrition: G1: 30% (6/20) G2: 14% (3/21)	Moderate

Note. [a] The control group worsened over time, hence the positive change on the scale.

Abbreviations. ACT, acceptance and commitment therapy; AE, adverse event; BEST, Borderline Evaluation of Severity Over Time; BHS, Beck Hopelessness Scale; BPD, borderline personality disorder; DERS, Difficulty in Emotion Regulation Scale; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition*; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; RCT, randomized controlled trial; SCID-I, Structured Clinical Interview for DSM-IV Axis I Disorders; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SD, standard deviation; TAU, treatment as usual.

Manual-Assisted Cognitive Therapy vs. Treatment as Usual

Table D—5. Study characteristics and main results of manual-assisted cognitive therapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Weinberg et al. (2006)	Design: RCT Setting: Outpatient, single center Country: United States Funding: Other, foundation	N=30 G1 (15): TAU G2 (15): MACT: 6 sessions adjunctive to ongoing TAU, modified to focus on deliberate self-harm in patients with BPD 6 sessions (duration NR)	Inclusion: Females; age 18-40 years; met DSM-IV and DIB-R criteria for BPD; history of repetitive deliberate self-harm with at least 1 episode during the month before enrollment Exclusions: Comorbid psychotic disorders, bipolar I disorder, or SUD; elevated suicide risk	Mean (SD) age: G1: 26 (7.7) G2: 30 (8.6) % Female: 100 % Race/ethnicity: White: 93 Nonwhite: 7	Primary outcome: NR G2 significantly more effective than G1 in reducing the frequency (1.98 vs. 6.69, p<0.001) and severity (0.51 vs. 1.01, p<0.001) of deliberate self-harm 6 months post-treatment Attrition: 0% (0/30) G1: 0% (0/15) G2: 0% (0/15)	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
			(scoring ≥9 on BHS); describing a concrete immediate suicide plan			

Abbreviations. BHS, Beck Hopelessness Scale; BPD, borderline personality disorder; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; MACT, manual-assisted cognitive therapy; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SUD, substance use disorder; TAU, treatment as usual.

Cognitive-Behavioral Therapy vs. Treatment as Usual

Table D—6. Study characteristics and main results of cognitive-behavioral therapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Davidson et al. (2006) BOSCOT	Design: RCT Setting: Outpatient, multicenter Country: United Kingdom Funding: Other, foundation	N=106 G1 (52): TAU: inpatient and outpatient hospital services, community-based services, and primary and community care services G2 (54): CBT: up to 30 sessions over 1 year (1 hour/session) with weekly supervision from CBT experts at each site 24 months	Inclusion: Age 18-65 years; met criteria for at least 5 items of the BPD using the DSM-IV Axis II Personality Disorders; received either in-patient psychiatric services or an assessment at accident and emergency services Exclusion: Currently receiving in-patient treatment for a mental state disorder or systematic psychological therapy or specialist service; evidence of an organic illness, mental impairment, alcohol or drug dependence, schizophrenia, or bipolar disorder	Mean (SD) age: 32 (9.1) G1: 31 (9.4) G2: 32 (9.0) % Female: 84 % Race/ethnicity: White: 100	Primary outcome: Suicidal acts, psychiatric hospitalization, accident, and emergency attendance at 24 months G2 significantly lower number of suicidal acts per person (0.87 vs. 1.73, p=0.02) and greater improvements on STAI (5.4 vs. 0.5, p=0.01) than G1 after 24 months No significant differences in suicidal acts, STAI, BDI-II, EuroQuol-5D, SFQ, or for the number of hospitalizations after 12 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 15% Differential attrition: ≤10 percentage points	Moderate

Abbreviations. AE, adverse event; BDI-II, Beck Depression Inventory; BOSCOT, Borderline Personality Disorder Study of Cognitive Therapy; BPD, borderline personality disorder; CBT, cognitive-behavioral therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; EuroQuol-5D, European Quality of Life-5 Dimension; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SFQ, Social Functioning Questionnaire; STAI, State-Trait Anxiety Inventory; TAU, treatment as usual.

Dialectical Behavior Therapy vs. Wait-list/Treatment as Usual

Table D—7. Study characteristics and main results of dialectical behavior therapy compared with wait-list or treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bohus et al. (2004)	Design: Nonrandomized clinical trial Setting: Inpatient, single center Country: Germany Funding: Government, DFG other, BPDRF Foundation	N=60 G1 (20): WL (TAU) G2 (40): DBT: individual therapy (2 hour/week), group skills training (2 hour/week), group psychoeducation (1 hour/week), peer group meetings (2 hour/week), mindfulness group (1 hour/week), individual body-oriented therapy (1.5 hour/week), and therapist team consultations meetings (2 hour/week) 3 months	Inclusion: Met DSM-IV criteria for BPD using SCID-II and DIB-R; 1 suicide attempt or minimum of 2 NSSI acts within the last 2 years Exclusion: Comorbid schizophrenia, bipolar I disorder, substance abuse, or intellectual disability; living >250 miles from inpatient center; ongoing outpatient DBT or DBT post-discharge	Mean (SD) age: G1: 30 (5.4) G2: 29 (7.2) % Female: 100 % Race/ethnicity: NR	Primary outcome: NR G2 significantly more effective than G1 on GSI (0.56 vs. 0.07, p=0.005), GAF (11.4 vs. 1.3, p=0.003), BDI (NR vs. 10.4, p=0.002), STAI (-8.2 vs. +1.2, p<0.001), Ham-A (0.6 vs. NR, p=0.01), and self-mutilation (62% vs. 31%, p=0.039) No significant differences on DES and STAXI Incidence of AEs: NR Withdrawals due to Aes: NR Attrition: 17% Differential attrition: G1: 5% (1/20) G2: 22% (9/40)	High
Carter et al. (2010)	Design: RCT Setting: Outpatient, single center Country: Australia Funding: NR	N=76 G1 (35): 6 months WL while receiving TAU G2 (38): DBT: team-based approach including individual therapy, weekly group-based skills training, and telephone access to an individual therapist and therapist supervision groups 12 months	Inclusion: Females; age 18-65 years; met DSM-IV criteria for BPD; history of multiple episodes of deliberate self-harm with at least 3 self-reported episodes in preceding 12 months Exclusion: Presence of a disabling organic condition, schizophrenia, bipolar disorder, psychotic depression, florid antisocial behavior, or developmental disability	Mean (SD) age: 25 (6.1) % Female: 100 % Race/ethnicity: NR Above results reported among N analyzed	Primary outcome: Deliberate self-harm and hospitalizations because of self-harm at 6 months No significant differences on number of self-harm episodes, proportion of participants with self-harm, and hospitalizations Incidence of Aes: NR Withdrawals due to Aes: NR Attrition: 30% Differential attrition: G1: 11% (4/35) G2: 47% (18/38)	Low

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Feigenbaum et al. (2012)	Design: RCT Setting: Outpatient, single center Country: United Kingdom Funding: Government, C&IHA, NTRHA	N=42 G1 (16): TAU: standard care of a range of individualized service provisions according to the patients' needs through the local crisis services G2 (26): DBT: goal setting and commitment building (3-6 weeks), individual therapy (1 hour/week), group skills training (2.5 hour/week), and out-of-hours telephone consultation 12 months	Inclusion: Men and women; age 18-65 years; DSM-IV criteria for cluster B personality disorder Exclusion: Forensic history with evidence of current high and immediate risk to others; in long-term psychotherapeutic treatment for schizophrenia or bipolar disorder; substance abuse; severe cognitive impairment	Mean (SD) age: G1: 35 (7.4) G2: 35 (7.8) % Female: 73 % Race/ethnicity: NR	Primary outcome: CORE-OM at 12 months No significant differences on CORE-OM, DSH, DES, BDI, STAXI, and OAS-M Incidence of Aes: NR Withdrawals due to Aes: NR Attrition: 29% Differential attrition: G1: 13% (2/16) G2: 39% (10/26)	High
Gregory and Sachdeva (2016)[a]	Design: Retrospective cohort Setting: Outpatient, single center Country: United States Funding: ApsaA	N=41 G1 (16): TAU: Unstructured psychotherapy G2 (25): DBT: skill group sessions including learning mindfulness, emotion regulation, and distress tolerance followed by individual sessions 12 months	Inclusion: Age >18 years; SCID-II; Individual Assessment Profile Exclusion: Schizophrenia, intellectual disabilities, or dementia	Mean (SD) age: G1: 29 (11.5) G2: 37 (10.2) % Female: 81 % Race/ethnicity: Caucasian: 88 Other: 12	Primary outcome: BEST at 12 months G2 significantly more effective than G1 on BDI (-5.5 vs. -0.6, p<0.001) No significant differences on BEST and SDS, and the number of suicide attempts and number of self-injuries Incidence of Aes: NR Withdrawals due to Aes: G1: 0% (0/16) G2: 0% (0/25) Attrition: 53% Differential attrition: <10 percentage points	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
McMain et al. (2017)	Design: RCT Setting: Outpatient, single center Country: Canada Funding: Ontario Mental Health Foundation	N = 84 G1 (42): WL G2 (42): Brief DBT: skills training only 20 weeks Follow up: 32 weeks	Inclusion: Age 18-60 years; met DSM-IV criteria for BPD; 2 suicidal and/or NSSI episodes in the past 5 years, with 1 occurring within 10 weeks prior to enrollment Exclusion: Met DSM-IV criteria for a psychotic disorder, bipolar I disorder, or dementia; evidence of intellectual disability; participation in a DBT program within the past year	Mean (SD) age: 30 (8.6) % Female: 79 % Race/ethnicity: NR	Primary outcome: Frequency of suicidal or NSSI episodes at 32 weeks G2 significantly more effective than G1 to reduce suicidal and self-injurious episodes on the LSASI (7.65 vs. 5.77, p=0.04) and to improve symptoms on the STAXI (8.44 vs. 4.79, p<0.001) and the DERS (20.80 vs. 4.74, p<0.01) G2 significantly more clinically relevant improvements on SCL-90-R than G1 (43.8% vs. 18.4%, p=0.024) No significant differences on DSHI, BSL-23, BDI, and BIS-11 Incidence of Aes: NR Withdrawal due to Aes: NR Attrition: 16% Differential attrition: <10 percentage points	Moderate
Verheul et al. (2003) ; van den Bosch et al. (2005)	Design: RCT Setting: Outpatient, with various settings Country: The Netherlands Funding: Dutch health insurance company	N=64 G1 (33): TAU: clinical management from the original referral source; generally no more than 2 sessions per month with a psychologist, a psychiatrist, or a social worker G2 (31): Weekly DBT: individual cognitive behavioral psychotherapy sessions with the primary therapist, skills-training groups (2-2.5 hours/session), and supervision and	Inclusions: Female; age 18-70 years; BPD; residing near Amsterdam; referred by psychologist or psychiatrist willing to sign an agreement to commit to delivering 12 months of TAU Exclusions: DSM-IV diagnosis of bipolar disorder or (chronic) psychotic disorder; insufficient command of the Dutch language; severe cognitive impairments	Mean (SD) age: 35 (7.7) % Female: 100 % Race/ethnicity: NR	Primary outcome: NR G2 more effecive than G1 to reduce self-mutilating behavior (35% vs. 57%, p=0.003); numerically lower frequency of suicidal attempts (7% vs. 26%, p=0.06) for G2 than G1 Incidence of adverse events: NR Withdrawal due adverse events: NR Attrition: 19%	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		consultation meetings for the therapists 52 weeks				

Note. [a] Gregory and Sachdeva (2016) is a three-arm trial. The two relevant arms to DBT vs. TAU are reported in this table; other eligible comparisons are reported in Tables 11 and 20.
Abbreviations. AE, adverse event; APsaA, American Psychoanalytic Association; BDI: Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BIS-11, Barrett Impulsiveness Scale -11; BPD, borderline personality disorder; BPDFR, Borderline Personality Disorder Research Foundation; BSL-23, Borderline Symptom List-23; C&IHA, Camden and Islington Health Authority; CORE-OM, Clinical Outcomes in Routine Evaluation–Outcome Measure; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; DES, Dissociations Experiences Scale; DFG, German Research Foundation; DIB-R, Diagnostic Interview for Borderlines-Revised; DSH, deliberate self-harm; DSHI, Deliberate Self-Harm Inventory; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; GSI, Global Severity Index; Ham-A, Hamilton Rating Scale for Anxiety; LSASI, Lifetime Suicide Attempt Self-Injury Interview; N, sample size; NR, not reported; NSSI, nonsuicidal self-injury; NTRHA, North Thames Regional Health Authority; OAS-M, Overt Aggression Scale-Modified; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90-R, Symptom Checklist-90-Revised; SD, standard deviation; SDS, Sheehan Disability Scale; STAI: State-Trait-Anxiety Inventory; STAXI, State-Trait Anger Expression Inventory; TAU, treatment as usual; WL, wait-list.

Dialectical Behavior Therapy vs. Mentalization-Based Treatment

Table D—8. Study characteristics and main results of dialectical behavior therapy compared with mentalization-based treatment for borderline personality disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Barnicot and Crawford (2019)	Design: Non-randomized clinical trial Setting: Outpatient, multicenter Country: United Kingdom Funding: Government, NIH	N=90 G1 (58): DBT G2 (32): MBT 12 months	Inclusion: Met DSM-IV criteria for BPD; were about to begin either outpatient DBT or MBT Exclusion: Intellectual disability; difficulty communicating in English; insufficient capacity to provide informed consent	Mean (SD) age: 31 (13.0) % Female: 72 % Race/ethnicity: White: 64 Black and minority: 36	Primary outcome: NR No significant differences between G1 and G2 on BEST, DERS, and DES, and for the number of self-harm incidents at 12 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 13% Differential attrition: <10 percentage points	High

Abbreviations. AE, adverse event; BEST, Borderline Evaluation of Severity Over Time; BPD, borderline personality disorder; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; DES, Dissociative Experiences Scale; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; MBT, mentalization-based treatment; N, sample size; NIH, National Institutes of Health; NR, not reported; SD, standard deviation.

Dialectical Behavior Therapy vs. General Psychiatric Management

Table D—9. Study characteristics and main results of dialectical behavior therapy compared with general psychiatric management for borderline personality disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
McMain et al. (2009)	Design: RCT Setting: Both in-patient and outpatient Country: Canada Funding: Government	N = 180 G1(90): Weekly individual therapy and medication management G2 (90): DBT: weekly individual therapy, skills group sessions, and phone coaching with explicit focus on self-harm and suicidal behavior 1 year Follow up: 36 months	Inclusion: Met DSM-IV criteria for BPD; age 18-60 years; ≥2 episodes of suicidal or NSSI episodes in the past 5 years and ≥1 of which was in the 3 months preceding enrollment Exclusion: DSM-IV psychotic disorder, bipolar I disorder, delirium, dementia, or SUD in the preceding 30 days; medical condition that precluded psychiatric medications; any serious medical condition likely to require hospitalization within the next year (e.g., cancer)	Mean (SD) age: 30 (9.9) % Female: 86 % Race/ethnicity: NR	Primary outcome: Suicidal episodes, NSSI at 12 months No significant differences between G1 and G2 for number of suicidal events and NSSI, and on SCL-90-R, ZAN-BPD, BDI, and IIP after 12 months G1 significantly greater improvements on BDI (17.4 vs. 12.7, p=0.004) than G2 at 36-month follow-up; no significant differences at 36 months for any of the other outcomes Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 38% Differential attrition: <10 percentage points	High

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BPD, borderline personality disorder; DBT, dialectical behavior therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; IIP, Inventory of Interpersonal Problems; N, sample size; NR, not reported; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; SCL-90-R, Symptom Checklist-90-Revised; SD, standard deviation; SUD, substance use disorder; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Dialectical Behavior Therapy vs. Systems Training for Emotional Predictability and Problem-Solving Behavior Therapy

Table D—10. Study characteristics and main results of dialectical behavior therapy compared with systems training for emotional predictability and problem-solving behavior therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Guillén Botella et al. (2021)	Design: Non-randomized clinical trial Setting: Outpatient, multicenter Country: Spain	N=72 G1 (45): Weekly individual and group DBT G2 (27): Weekly STEPPS group therapy plus weekly individual therapy	Inclusion: Met DSM-5 criteria for BPD Exclusion: Moderate or severe intellectual disability, schizophrenia, or bipolar disorder	Mean (SD) age: 32 (8.8) % Female: 94 % Race/ethnicity: Caucasian: 100	Primary outcome: BSL-23 at 6 months G1 significantly more effective than G2 on the sum of the BSL-23 (23.56 vs. 29.29, p=0.03) after 6 months No significant differences for any other measure Incidence of AEs: NR	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: NR	6 months			Withdrawals due to AEs: NR Attrition: 32% Differential attrition: <10 percentage points	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; BSL-23, Borderline Symptom List-23; DBT, dialectical behavior therapy; DSM-5, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; SD, standard deviation; STEPPS, systems training for emotional predictability and problem solving.

Dialectical Behavior Therapy vs. Dynamic Deconstructive Psychotherapy

Table D—11. Study characteristics and main results of dialectical behavior therapy compared with dynamic deconstructive psychotherapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Gregory and Sachdeva (2016)[a]; Sachdeva et al. (2013)	Design: Retrospective cohort Setting: Outpatient, single center Country: United States Funding: APsaA	N=52 G1 (25): DBT: weekly 1-hour individual and 2-hour group sessions and telephone skills coaching G2 (27): DDP: combined elements of translational neuroscience, object relations theory, and deconstructionist philosophy; weekly 1-hour individual sessions 12 months	Inclusion: Age >18 years; BPD by SCID-II and Individual Assessment Profile Exclusion: Schizophrenia, intellectual disabilities, or dementia	Mean (SD) age: G1: 29 (11.5) G2: 37 (10.2) % Female: 81 % Race/ethnicity: Caucasian: 88 Other: 12	Primary outcome: BEST scores at 12 months G2 significantly greater improvement than G1 on severity (BEST: 33.0 vs. 41.8, p=0.04), self-injuries (SBQ: 1.3 vs. 2.4, p=0.02), depression (BDI: 17.1 vs. 27.6, p=0.009), and disability (SDS: 3.8 vs. 6.1, p=0.049) No differences between G1 and G2 in suicide attempts Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 53% Differential attrition: G1: 69% (11/16) G2: 64% (16/25) G3: 33% (9/27)	High

Note. [a] Gregory and Sachdeva (2016) is a three-arm trial. The two relevant arms to DBT vs. DDT are reported in in this table; other eligible comparisons are reported in Tables 7 and 20.

Abbreviations. AE, adverse event; APsaA, American Psychoanalytic Association; BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BPD, borderline personality disorder; DBT, dialectical behavior therapy; DDP, dynamic deconstructive psychotherapy; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; RCT, randomized controlled trial; SBQ, Suicidal Behaviors Questionnaire; SCID-II, Structured Clinical Interview for DSM-IV Axis II Personality Disorders; SD, standard deviation; SDS, Sheehan Disability Scale.

Dialectical Behavior Therapy vs. Transference-Focused Psychotherapy vs. Supportive Therapy

Table D—12. Study characteristics and main results of dialectical behavior therapy compared with transference-focused psychotherapy and supportive therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Clarkin et al. (2007)	Design: RCT Setting: Outpatient, multicenter Country: United States Funding: Other, foundation	N=90 G1 (22): Weekly supportive treatment sessions G2 (23): TFP: twice weekly individual sessions G3 (17): DBT: weekly individual and group sessions and available telephone consultation 12 months	Inclusion: Age 18-50 years; met DSM-IV criteria for BPD Exclusion: Comorbid psychotic disorders, bipolar I disorder, delusional disorder, delirium, dementia, and/or amnestic, other cognitive disorders, or SUD	Mean (SD) age: 31 (7.9) % Female: 92 % Race/ethnicity: White: 68 Black: 10 Hispanic: 9 Asian: 6 Other: 8	Primary outcome: Suicidal behavior at 12 months No significant differences among G1, G2, and G3 in suicidal behavior, BDI, BSI, or GAF at 12 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 31% Differential attrition: G1: 27% (8/30) G2: 23% (7/30) G3: 43% (13/30)	High

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BPD, borderline personality disorder; BSI, Brief Symptom Inventory; DBT, dialectical behavior therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; G3, Group 3; GAF, Global Assessment of Functioning; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SUD, substance use disorder; TFP, transference-focused psychotherapy.

Components of Dialectical Behavior Therapy vs. Other Components of Dialectical Behavior Therapy

Table D—13. Study characteristics and main results of components of dialectical behavior therapy compared with other components of dialectical behavior therapy for borderline personality disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Andión et al. (2012)	Design: Non-randomized clinical trial Setting: Outpatient, single center Country: Spain	N=51 G1 (37): Weekly individual DBT therapy G2 (14): Combined weekly individual and group DBT therapy sessions 12 months intervention, followed through 18 months	Inclusion: Age 18-50 years; ≥1 suicide attempt and/or 1 self-harm behavior during the previous month; met criteria for DSM-IV Axis II and Axis I Disorders Exclusion: Intellectual disability, schizophrenia, or	Mean (SD) age: 26 (6.5) % Female: 100 % Race/ethnicity: NR	Primary outcome: Suicide attempts and self-harm at 12 and 18 months No significant differences between groups on any outcome Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 10% Differential attrition: <10 percentage points	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: Government, health department, La Caixa		bipolar I disorder; previous DBT treatment			
Linehan et al. (2015)	Design: RCT Setting: Outpatient, multicenter Country: United States Funding: Government, NIMH	N=99 G1 (33): Weekly standard DBT: skills training, individual therapy, telephone coaching, and a therapist consultation team G2 (33): DBT individual therapy with no group skills training G3 (33): Weekly DBT group skills training with no individual therapy 1-year treatment followed through 2 years	Inclusion: Women; age 18-60 years; met DSM-IV criteria for BPD, Axis II; ≥2 suicide attempts and/or NSSI episodes in the past 5 year, ≥1 suicide attempt or NSSI act in the 8-week period before entering the study, and ≥1 suicide attempt in the past year Exclusion: <70 IQ score; DSM-IV criteria for current psychotic or bipolar disorders; seizure disorder requiring medication; required primary treatment for another life-threatening condition	Mean (SD) age: 30 (8.9) % Female: 100 % Race/ethnicity: White: 71 Asian American: 5 Biracial: 22 Other: 2	Primary outcome: Frequency and severity of suicide attempts and NSSI episodes at 12 and 24 months No significant difference between groups in suicide attempts, NSSI acts, or suicide ideation During the treatment year, G1 and G3 significantly greater improvement in depression than G2 (12.3 and 10.4 vs. 18.2, p=0.02 on Ham-D) with no differences between groups in anxiety (Ham-A) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 26% Differential attrition: G1: 18% (6/33) G2: 33% (11/33) G3: 27% (9/33)	High
Lyng et al. (2020)	Design: Prospective cohort Setting: Outpatient, multicenter Country: Ireland Funding: NR	N=88 G1 (54): Weekly standard DBT: individual therapy, group skills training, phone consultation (as needed), and therapist consultation team meeting G2 (34): Weekly DBT group skills training 6 months	Inclusion: DSM-IV-TR BPD or equivalent diagnosis of emotionally unstable personality disorder by community psychiatrist Exclusion: Enduring psychotic disorder or primary (i.e., main reason for seeking treatment) alcohol or substance abuse disorder; suicide attempt in the previous 6 months and/or ongoing medically	Mean (SD) age: 33 (range 18-59) % Female: 83 % Race/ethnicity: NR	Primary outcome: Borderline symptoms, general psychopathology, suicidal ideation at 6 months No significant differences between groups in BPD symptomatology, suicide ideation, and symptom severity index G2 significantly greater improvement on the BHS (8.0 vs. 11.91, p=0.02) and on the DERS (96.24 vs. 115.12, p=0.02) Incidence of AEs: NR Withdrawals due to AEs: NR	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
			serious self-harm; other weekly counseling		Attrition: 25.0% Differential attrition: G1: 17% (9/54) G2: 38% (13/34)	

Abbreviations. AE, adverse event; BHS, Beck Hopelessness Scale; BPD, borderline personality disorder; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*; G1, Group 1; G2, Group 2; G3, Group 3; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; IQ, intelligence quotient; N, sample size; NIMH, National Institute of Mental Health; NR, not reported; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; SD, standard deviation.

Component of Dialectical Behavior Therapy Skills Training vs. Another Component of Dialectical Behavior Therapy Skills Training

Table D—14. Study characteristics and main results of component of dialectical behavior therapy skills training compared with another component of dialectical behavior therapy skills training for borderline personality disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Carmona i Farrés et al. (2019a)	Design: RCT Setting: Outpatient, single center Country: Spain Funding: Other, mixed	N=70 G1 (35): Weekly group DBT interpersonal effectiveness skills training for 10 sessions G2 (35): Weekly group DBT mindfulness skills training for 10 sessions 10 weeks	Inclusion: Age 18-50 years; DSM-IV BPD diagnosis; no comorbidities with schizophrenia, drug-induced psychosis, organic brain syndrome, SUD, bipolar disorder, intellectual disability, or major depressive episode in course; no concurrent psychotherapy at study enrollment; no previous training in mindfulness, other meditation-contemplative practices, or any other mind-body practices Exclusion: NR	Mean (SD) age: G1: 33.29 (8.54) G2: 30.51 (6.9) % Female: 90 % Race/ethnicity: NR	Primary outcome: Emotional dysregulation (DERS) and impulsivity (BIS-11) at 10 weeks No significant differences between G1 and G2 on DERS at 10 weeks G2 significantly greater improvement on the BIS-11 (75.3 vs. 79.3, p=0.03) at 10 weeks Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 28% G1: 20% G2: 37%	High
Carmona i Farrés et al. (2019b)	Design: RCT Setting: Outpatient, single center	N=65 G1 (32): Weekly group DBT interpersonal	Inclusion: Age 18-50 years; DSM-IV BPD diagnosis; no comorbidities with schizophrenia, drug-induced	Mean (SD) age: G1: 33.75 (8.78) G2: 31.03 (6.76)	Primary outcome: Default mode network activation and deactivation during an executive task	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Country: Spain Funding: Other, mixed	effectiveness skills training for 10 sessions G2 (33): Weekly group DBT mindfulness skills training for 10 sessions 10 weeks	psychosis, organic brain syndrome, SUD, bipolar disorder, intellectual disability, or major depressive episode in course; no concurrent psychotherapy at study enrollment; right-handed; IQ within the normal range Exclusion: NR	% Female: 89.2 % Race/ethnicity: NR	No significant differences between G1 and G2 on BSL-23, BDI, STAI-T, or STAI-S at 10 weeks (decreases on the outcome measures in both groups) No between-group differences in default mode network activation or deactivation Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 23% Differential attrition: ≤10 percentage points	
Elices et al. (2016)	Design: RCT Setting: Outpatient, single center Country: Spain Funding: Government	N=64 G1 (32): Weekly group DBT interpersonal effectiveness skills training G2 (32): Weekly group DBT mindfulness training 10 weeks	Inclusion: Age 18-45 years; BPD criteria according to SCID-II and DIB-R Exclusion: Lifetime diagnosis of schizophrenia, drug-induced psychosis, organic brain syndrome, or bipolar disorder; participation in any psychotherapy during the study or having received DBT in the past; having meditation/yoga experience	Mean (SD) age: 32 (6.9) G1: 32 (6.82) G2: 32 (7.25) % Female: 86 % Race/ethnicity: NR	Primary outcome: Borderline severity at 10 weeks G2 significantly reduced BPD symptoms on the BSL-23 than G1 at 10 weeks (33.5 vs. 52.5, p=0.001) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 31% Differential attrition: G1: 22% (7/32) G2: 41% (13/32)	High
Schmidt et al. (2021)	Design: RCT Setting: Outpatient, single center Country: Spain Funding: Other, mixed	N=102 G1 (52): Weekly group DBT interpersonal effectiveness skills training for 10 sessions G2 (50): Weekly group DBT mindfulness skills training for 10 sessions	Inclusion: Age 18-50 years; DSM-IV BPD diagnosis; no comorbidities with schizophrenia, drug-induced psychosis, organic brain syndrome, SUD, bipolar disorder, intellectual disability, or major depressive episode in course; no concurrent psychotherapy at study enrollment; no previous experience in mindfulness	Mean (SD) age: G1: 33 (8.0) G2: 32 (8.0) % Female: 93 % Race/ethnicity: NR	Primary outcome: Borderline severity (BSL-23) at 10 weeks G2 significantly greater improvements on the BSL-23 at 10 weeks (37.38 vs. 48.90, p=0.000) and on the EQ at 10 weeks (31.28 vs. 27.48, p=0.001) No differences between groups on the DERS Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: NR Differential attrition: NR	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		10 weeks	meditation and a DBT skills training Exclusion: NR			

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BIS-11, Barrett Impulsiveness Scale-11; BPD, borderline personality disorder; BSL-23, Borderline Symptom List-23; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; EQ, European Quality of Life; G1, Group 1; G2, Group 2; IQ, intelligence quotient; N, sample size; NR, not reported; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Personality Disorders; SD, standard deviation; STAI-S, State-Trait Anxiety Inventory-State; STAI-T, State-Trait Anxiety Inventory-Trait; SUD, substance use disorder.

Dialectical Behavior Therapy vs. Cognitive Therapy

Table D—15. Study characteristics and main results of dialectical behavior therapy compared with cognitive therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Lin et al. (2019)	Design: RCT Setting: Other: university counseling centers Country: Taiwan Funding: Government	N=82 G1 (40): CT group: weekly CT group, phone consultation as when needed, and closed social media community for group members G2 (42): Weekly DBT group skills training, phone consultation as when needed, and closed social media community for group members 8 weeks	Inclusion: College students who met criteria for BPD per the BPDFS; ≥21 on Ko's Depression Inventory; ≥1 suicide attempt in the past 6 months Exclusions: Lifetime diagnosis of schizophrenia, schizoaffective disorder, bipolar disorder, or psychotic disorder; current severe depression and suicide risk indicating need for inpatient care and crisis intervention; current neurological signs and substance abuse during the last 6 months	Mean (SD) age: G1: 20.47 (0.71) G2: 20.40 (0.76) % Female: 87.8 % Race/ethnicity: NR	Primary outcome: Suicide attempt at 32 weeks No significant difference between G1 and G2 on suicide reattempt (CMSADS-L Short form) and Ko's Depression Inventory at 32 weeks Compared with G1, G2 significant improvements in the BPDFS (5.87 vs. 4.91, p<0.01) and suicide ideation (ASIQ-S) (42.96 vs. 40.27, p<0.01) at 32 weeks Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 1% Differential attrition: <10 percentage points	Moderate

Abbreviations. AE, adverse event; ASIQ-S, Adult Suicidal Ideation Questionnaire-Shortened Version; BPD, borderline personality disorder; BPDFS, Borderline Personality Disorder Features Scale; CMSADS-L, Chinese Version of the Modified Schedule of Affective Disorders and Schizophrenia-Lifetime; CT, cognitive therapy; DBT, dialectical behavior therapy; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation.

Dialectical Behavior Therapy vs. Community Therapy by Experts

Table D—16. Study characteristics and main results of dialectical behavior therapy compared with community therapy by experts.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Linehan et al. (2006)	Design: RCT Setting: Outpatient, single center Country: United States Funding: Government, NIMH	N=111 G1 (49): Community treatment by experts; developed specifically for this study to control for factors previously uncontrolled for in DBT studies G2 (52): DBT 1 year	Inclusions: Women; age 18-45 years; BPD; 2 suicide attempts or self-injuries in the past 5 years with ≥1 in the past 8 weeks Exclusions: Comorbid schizophrenia, schizoaffective disorder, bipolar, psychotic disorder, or intellectual disability; seizure disorder requiring medication; mandate to treatment; need for primary treatment for another debilitating condition	Mean (SD) age: 29 (7.5) % Female: 100 % Race/ethnicity: White: 87.0 Black: 4.0 Asian American: 2.0 Native American or Alaskan Native: 1.0 Other: 5.0	Primary outcome: NR G2 more effective than G1 in preventing suicide attempts (23% vs. 46%, p=0.01), emergency room visits for suicide ideation (10.6% vs. 18.4%, p=0.02), and hospital admissions for suicide ideation (14.9% vs. 18.4%, p=0.004) No significant differences between groups in NSSI, Ham-D, and RLI Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: Overall: 18% G1: 27.5% (14/51) G2: 10.0% (6/60)	High

Abbreviations. AE, adverse event; BPD, borderline personality disorder; DBT, dialectical behavior therapy; G1, Group 1; G2, Group 2; Ham-D, Hamilton Rating Scale for Depression; N, sample size; NIMH, National Institute of Mental Health; NR, not reported; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; RLI, Reasons for Living Inventory; SD, standard deviation.

Dialectical Behavior Therapy Plus REMS Treatments vs. REMS Treatments

Table D—17. Study characteristics and main results of dialectical behavior therapy plus REMS treatments compared with REMS treatments.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bianchini et al. (2019)	Design: RCT Setting: Inpatient, single center Country: Italy Funding: NR	N=21 G1 (11): REMS (Residenze per l'Esecuzione delle Misure di Sicurezza, a small-scale intensive therapeutic unit) G2 (10): DBT plus REMS treatments 12 months	Inclusions: Met criteria for BPD as measured by PAI; history of violence to others Exclusions: Cognitive deficit (IQ <70) and/or comorbid neurological diseases	Mean (SD) age: 42 (8.14) % Female: 0 % Race/ethnicity: NR	Primary outcome: NR No between group comparisons at the end of treatment Significant change on only 2 outcomes, DERS and BIS-11, within the intervention group Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: NR	Moderate

Abbreviations. AE, adverse event; BPD, borderline personality disorder; BIS-11, Barrett Impulsiveness Scale-11; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; G1, Group 1; G2, Group 2; IQ, intelligence quotient; N, sample size; NR, not reported; PAI, Personality Assessment Inventory; RCT, randomized controlled trial; REMS, Residenze per l'Esecuzione delle Misure di Sicurezza; SD, standard deviation.

Dialectical Behavior Therapy vs. Conversational Model

Table D—18. Study characteristics and main results of dialectical behavior therapy compared with conversational model.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Walton et al. (2020)	Design: RCT Setting: Outpatient, single center Country: Australia Funding: None	N=166 G1 (83): Conversational Model involving twice-weekly individual therapy G2 (83): Weekly DBT: individual therapy, group training, and access to to after-hours coaching 14 months	Inclusions: Age 18-65 years; DSM-IV BPD; ≥3 suicidal and/or NSSI episodes in the previous 12 months Exclusions: Disabling organic conditions, current acute psychotic illness, antisocial behavior that posed a significant threat to staff and fellow patients, or developmental disability; substance dependent; living more than 1-hour's drive from treatment facility; inability to	Mean (SD) age: 27 (7.8) % Female: 77 % Race/ethnicity: White: 139 (86%) Aboriginal: 10 (6%) Other: 13 (8%)	Primary outcome: Suicide attempts and NSSI at 14 months and depression severity (BDI-II) at 14 months No differences between groups in suicide attempts, NSSI, BPD severity (BPDSI-IV), interpersonal problems (IIP), dissociation (DES), and mindfulness at 14 months G2 significantly greater reductions in BDI-II scores at 14 months (15.94 vs. 22.13, p=0.005) and greater improvements in emotion regulation (DERS) (87.08 vs. 105.16, p=0.008) Incidence of AEs: NR Withdrawal due to AEs: NR	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
			speak or read English; prior treatment with DBT or Conversational Model		Attrition: 29% G1: 24% (20/83) G2: 34% (28/83)	

Abbreviations. AE, adverse event; BDI-II, Beck Depression Inventory-II; BPD, borderline personality disorder; BPDSI-IV, Borderline Personality Disorder Severity Index-IV; DBT, dialectical behavior therapy; DERS, Difficulties in Emotion Regulation Scale; DES, Dissociative Experiences Scale; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition;* G1, Group 1; G2, Group 2; IIP, Inventory of Interpersonal Problems; N, sample size; NR, not reported; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; SD, standard deviation.

Dialectical Behavior Therapy Skills Training vs. Standard Group Therapy

Table D—19. Study characteristics and main results of dialectical behavior therapy compared with standard group therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Soler et al. (2009)	Design: RCT Setting: Outpatient, single center Country: Spain Funding: Government	N=60 G1 (30): Weekly standard group therapy G2 (29): Weekly group DBT skills training 12 weeks	Inclusion: Age 18-45; met DSM-IV criteria for BPD as assessed by SCID-II and DIB-R; CGI-S score of ≥4 Exclusion: Comorbid schizophrenia, drug-induced psychosis, organic brain syndrome, alcohol or other psychoactive SUD, bipolar disorder, intellectual disability, or major depressive episode in course; current psychotherapy	Mean (SD) age: G1: 29.97 (5.63) G2: 28.45 (6.55) % Female: 83.0 % Race/ethnicity: NR	Primary outcome: NR No significant differences between G1 and G2 on CGI-BPD and SCL-90-R at 12 weeks G2 significantly greater improvement on the Ham-D (11.1 vs. 16.0, p=0.001) and Ham-A (16.6 vs. 13.0, p=0.03) at 12-weeks Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 49% Differential attrition: G1: 35% (10/29) G2: 63% (19/30)	High

Abbreviations. AE, adverse event; BPD, borderline personality disorder; CGI-BPD, Clinical Global Impression Scale for Borderline Personality Disorder; CGI-S, Clinical Global Impression-Severity; DBT, dialectical behavior therapy; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition;* G1, Group 1; G2, Group 2; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; N, sample size; NR, not reported; NSSI, nonsuicidal self-injury; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90-R, Symptom Checklist-90-Revised; SD, standard deviation; SUD, substance use disorder.

Dynamic Deconstructive Psychotherapy vs. Treatment as Usual

Table D—20. Study characteristics and main results of dynamic deconstructive psychotherapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Gregory and Sachdeva (2016)[a]	Design: Retrospective cohort Setting: Outpatient, single center Country: United States Funding: APsaA	N=44 G1 (16): TAU: Unstructured psychotherapy G2 (28): DDP: combined elements of translational neuroscience, object relations theory, and deconstructionist philosophy in weekly 1-hour individual sessions 12 months	Inclusion: Age >18 years; BPD by SCID-II and Individual Assessment Profile Exclusion: Schizophrenia, intellectual disabilities, or dementia	Mean (SD) age: G1: 29 (11.5) G2: 28 (11.7) % Female: 81 % Race/ethnicity: Caucasian: 88 Other: 12	Primary outcome: BEST at 12 months G2 significantly more effective than G1 on change from baseline in BEST (14.1 vs. -2.6, p=0.006), BDI (-12.6 vs. -0.6, p<0.001), and SDS (-2.5 vs. 0.6, p<0.001) No significant differences in the number of suicide attempts and self-injuries Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 53% Differential attrition: G1: 69% (11/16) G2: 33% (9/27)	High

Note. [a] Gregory and Sachdeva (2016) is a three-arm trial. The two relevant arms to DDP vs. TAU are reported in this table; other eligible comparisons are reported in Tables 7 and 11.
Abbreviations. AE, adverse event; APsaA, American Psychoanalytic Association; BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; DDP, dynamic deconstructive psychotherapy; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SD, standard deviation; SDS, Sheehan Disability Scale; TAU, treatment as usual.

Mentalization-Based Treatment vs. Other Active Comparators

Table D—21. Study characteristics and main results of mentalization-based treatment compared with other active comparators.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Beck et al. (2020) M-GAB	Design: double-blinded RCT Setting: Outpatient, multicenter Country: Denmark	N=112 G1 (56): Standardized to at least 12 individual monthly sessions and additional contact varied across clinics and therapists and according to the needs of the patient; therapists were nurses,	Inclusions: Age 14-17 years; met ≥4 DSM-5 BPD criteria and a total score above clinical cutoff (>67) on the BPFS-C Exclusions: Comorbid pervasive developmental disorder, learning disability (IQ <75), anorexia, psychosis, schizophrenia or	Mean (SD) age: G1: 16 (1.0) G2: 16 (1.1) % Female: 99% (111/112) % Race/ethnicity: NR	Primary outcome: BPFS-C No significant differences between G2 and G1 on BPFS-C, BPFS for Parent, ZAN-BPD, Risk-Taking and Self-Harm Inventory for Adolescents, BDI for Youth, internalizing or externalizing symptoms on the Youth Self-Report,	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: Government, region Zealand other, TrygFonden	psychologists, social workers, or psychiatrists who were not trained in or practicing MBT; treatment was not manualized G2 (56): MBT delivered as a 1-year program with 3 components: MBT-introduction, -group, and -parents (90-minute sessions) 12 months	schizotypal personality disorder, antisocial personality disorder, or any mental disorder other than BPD considered the primary diagnosis; current (past 2 months) SUD (not substance abuse); current psychiatric inpatient treatment		the Child Behavior Checklist, or the Children's GAS Report of any AEs: 0% Attrition: 25.0% (28/112) G1: 19.6% (11/56) G2: 30.3% (17/56)	

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BPD, borderline personality disorder; BPFS, Borderline Personality Features Scale; BPFS-C, Borderline Personality Features Scale for Children; DSM-5, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition;* G1, Group 1; G2; Group 2; GAS, Global Assessment Scale; IQ, intelligence quotient; MBT, mentalization-based treatment; M-GAB, Mentalization-Based Treatment in Groups for Adolescents with Borderline Personality Disorder; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SUD, substance use disorder; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Mentalization-Based Treatment vs. Supportive Therapy

Table D—22. Study characteristics and main results of mentalization-based treatment compared with supportive therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bateman and Fonagy (2009); Bateman et al. (2021)	Design: RCT Setting: Outpatient, single center Country: United Kingdom Funding: Other, foundation	N=134 G1 (63): SCM individual and group sessions plus medication; therapy based on supportive approach with case management, advocacy support, and problem-oriented psychotherapeutic interventions G2 (71): MBT plus medication: 18-month manualized weekly combined individual and group psychotherapy 18 months	Inclusion: Age 18-65 years; DSM-IV BPD diagnosis; suicide attempt or episode of life-threatening self-harm within last 6 months Exclusion: Current long-term psychotherapeutic treatment; met DSM-IV criteria for psychotic disorder or bipolar I disorder; opiate dependence requiring specialist treatment; mental impairment or evidence of organic brain disorder	Mean (SD) age: G1: 31 (7.9) G2: 31 (7.6) % Female: 80 % Race/ethnicity: White: 72 Black: 18 Other: 10	Primary outcome: Suicide, self-injury, and hospitalizations at 18 months At 18 months, G2 significantly more effective than G1 in reducing life-threatening suicide attempts in previous 6-month period on the SCL-90-R (0.03 vs. 0.32, p<0.001), reducing severe self-harm incidents on the SCL-90-R (0.38 vs. 1.66, p<0.001), and reducing hospitalizations (0.03 vs. 0.19, p<0.001) (composite of all three measures, 0.5 vs. 2.2, p<0.001) At 18 months, G2 greater improvement in 6-month periods free of suicidal behavior,	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		Follow up: 6 years			severe self-harm, and hospitalizations (OR 0.28, 95% CI=0.13-0.61, p<0.002) G2 significantly greater than G1 in the number of participants who achieved the primary recovery criteria (free of self-harm, suicide attempts, or inpatient hospital stays) and who remained well over a 6-year follow-up period (75% vs. 51%, p=0.02) At 18 months, G2 significantly greater improvements in BDI (14.80 vs. 18.68, p<0.01), IIP (1.28 vs. 1.65, p<0.001), SCL-90-GSI (1.12 vs. 1.55, p<0.001), and GAF (60.9 vs. 53.2, p<0.001) Incidence of AEs: NR at 18 months Withdrawals due to AEs: NR Attrition: 26% at 18 months; 39% at 6 years post-treatment Differential attrition: <10 percentage points	
Carlyle et al. (2020)	Design: RCT Setting: Outpatient, single center Country: New Zealand Funding: NR	N=72 G1 (34): Enhanced therapeutic case management with case managers using the published manual of SCM G2 (38): MBT: manualized weekly 1-hour individual sessions and weekly 1.5-hour group sessions 18 months	Inclusion: BPD diagnosis using SCID-II Exclusion: Patients diagnosed with psychoses or primary substance dependence; insufficient proficiency in English; concurrent engagement in a structured psychological treatment for personality disorder	Mean (SD) age G1: 32 (11.7) G2: 32 (9.8) % Female: 99 % Race/ethnicity: NZ European: 79 Maori: 6 European other: 12.5 Other: 3	Primary outcome: Nonsuicidal self-harm and suicide attempts at 18 months At 18 months, no significant differences between groups on incidents of nonsuicidal self-harm, suicide attempts, or hospitalizations Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 14% Differential attrition: <10 percentage points	Moderate
Jørgensen et al. (2013)	Design: RCT Setting: Outpatient, single center	N=111 randomized; n=85 treated G1 (27): Biweekly group therapy and monthly group	Inclusion: Age ≥21 years; met DSM-IV BPD criteria as assessed by SCID-II; GAF score >34	Mean (SD) age: G1: 30 (6.8) G2: 30 (6.5) % Female: 96	Primary outcome: GAF at 24 months At 24 months, G2 significantly greater improvement than G1 in therapist-rated GAF-F (56.7 vs. 51.3, p=0.007) and GAF-S (58.5 vs. 54.0, p<0.001)	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Country: Denmark Funding: NR	psycho-educational program for 6 months G2 (58): Weekly individual and group MBT therapy and monthly group MBT psycho-educational program for 6 months 24 months	Exclusion: Met diagnostic criteria for antisocial or paranoid personality disorder at the time of assessment; severe substance abuse (daily) requiring specialist treatment	% Race/ethnicity: NR	Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 32% Differential attrition: <10 percentage points	

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BPD, borderline personality disorder; CI, confidence interval; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; GAF-F, Global Assessment of Functioning-Functioning; GAF-S, Global Assessment of Functioning-Symptoms; IIP, Inventory of Interpersonal Problems; MBT, mentalization-based treatment; N, sample size; NR, not reported; OR, odds ratio; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Personality Disorders; SCL-90-R, Symptom Checklist-90-Revised; SCL-90-GSI, Symptom Checklist-90-Global Severity Index; SCM, structured clinical management; SD, standard deviation.

Mentalization-Based Treatment vs. Psychodynamic Treatment Program

Table D—23. Study characteristics and main results of mentalization-based treatment compared with psychodynamic treatment program.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Kvarstein et al. (2015)	Design: Prospective cohort Setting: Outpatient, single center Country: Norway Funding: NR	N=345 G1 (281): 18 weeks of weekly inpatient group therapies followed by weekly outpatient group therapy G2 (64): Weekly individual and group MBT therapy and monthly group MBT psycho-educational program. 36 months	Inclusion: NR; assessed baseline diagnostic status with the M.I.N.I. version 4.4 for DSM Axis-I diagnosis and the SCID-II at baseline. Exclusion: Treated in the transition period between G1 to G2; included in a RCT during 2004-2006	Mean (SD) age: G1: 30 (7.0) G2: 26 (6.0) % Female: 83.2 % Race/ethnicity: NR	Primary outcome: GAF, CIP, BSI-18 at 36 months G2 significantly greater improvements in CIP (0.9 vs. 1.4, p<0.001), BSI-18 (0.8 vs. 0.9, p<0.001), and GAF (63.0 vs. 56.0, p<0.001) No difference between G1 and G2 in self-harm and suicide attempts at 36 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 16% Differential attrition: G1: 8% (22/281) G2: 50% (32/64)	High

Abbreviations. AE, adverse event; BPD, borderline personality disorder; BSI-18, Brief Symptom Inventory-18; CIP, Circumplex of Interpersonal Problems; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; MBT, mentalization-based treatment; M.I.N.I., Mini International Neuropsychiatric Interview; N, sample size; NR, not reported; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SD, standard deviation.

Day-Hospital Mentalization-Based Treatment vs. Specialized Psychotherapy

Table D—24. Study characteristics and main results of day-hospital mentalization-based treatment compared with specialized psychotherapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bales et al. (2015)	Design: Nonconcurrent cohort Setting: Other day-hospital Country: Netherlands Funding: NR	N=204 G1 (175): A variety of psychotherapeutic treatments in inpatient, outpatient, and day-hospital settings G2 (29): MBT in a day-hospital setting: daily group therapy, weekly individual therapy, individual crisis planning and art therapy twice a week, mentalizing cognitive group therapy, and writing therapy; medication consultation when indicated 18-month treatment phase; actual treatment was a mean of 15.5 months (3.8)	Inclusion: Age ≥18 years or older; met DSM-IV criteria for BPD, Exclusion: Schizophrenia, ADHD, bipolar disorder, psychotic disorders, or SUDs; intellectual impairment; organic brain disorder	Mean (SD) age: G1: 30 (7.9) G2: 30 (6.2) % Female: G1: 86 G2: 69 % Race/ethnicity: NR	Primary outcome: Psychiatric symptoms, personality functioning at 18 months G2 significantly greater improvements than G1 in GSI at 18 months (1.04 vs. 1.21, p=0.01) and at 36 months (0.73 vs. 1.04, p=0.02) G2 favored on SIPP-118 changes in (mal)adaptive personality functioning (results NR) Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: NR (could assume none)	High
Laurenssen et al. (2018)	Design: RCT Setting: Other inpatient and outpatient therapies, multicenter Country: Netherlands Funding: Other, organization	N=95 G1 (41): Manualized psychiatric treatment and system-oriented tailored care G2 (54): Day-hospital MBT consisting of daily group psychotherapy, weekly individual psychotherapy, individual crisis planning, art therapy twice a week, mentalizing cognitive	Inclusion: Met DSM-IV criteria for BPD; score of ≥20 on the BPDSI Exclusion: Schizophrenia or bipolar disorder; substance abuse requiring specialist treatment; organic brain disorder	Mean (SD) age: G1: 34 (10.6) G2: 34 (9.4) % Female: 79 % Race/ethnicity: NR	Primary outcome: BPDSI total score at 18 months At 18 months, no significant differences between groups on any outcome Incidence of AEs: Completers: G1: 0% (0/15) G2: 0% (0/33) Withdrawals due to AEs: NR Attrition: 50% Differential attrition: <10 percentage points	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		group therapy, and writing therapy 18 months				

Abbreviations. ADHD, attention-deficit/hyperactivity disorder; AE, adverse event; BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GSI, Global Severity Index; MBT, mentalization-based treatment; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SIPP-118, Severity Indices of Personality Problems; SUD, substance use disorder.

Systems Training for Emotional Predictability and Problem Solving vs. Treatment as Usual

Table D—25. Study characteristics and main results of systems training for emotional predictability and problem solving compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Blum et al. (2008) STEPPS	Design: RCT Setting: Other outpatient, inpatient, and community Country: United States Funding: Government, NIMH	N=165 G1 (72 [data analysis based on 65]): TAU: 20 weeks of continued usual care including individual psychotherapy, medication, and case management G2 (93 [data analysis based on 59]): STEPPS plus individual therapy: 20 sessions (2 hour/ week) of manual-based STEPPS group treatment that combines cognitive behavioral elements with skills training; components included psychoeducation about BPD, emotion management skills training, and behavior	Inclusion: Subjects with DSM-IV BPD who could designate a mental health professional and a friend or relative to serve as system members Exclusion: Non-English speaker; had a psychotic or primary neurological disorder; participated in STEPPS previously	Among the 124 who received allocated intervention: Mean (SD) age: 32 (9.5) % Female: 83 % Race/ethnicity: White: 95 Black: 2 Other: 3	Primary outcome: BPD-specific psychiatric symptoms (ZAN-BPD) measured at 20 weeks Significantly improved symptoms of BPD with G2 than G1 at 20 weeks on the ZAN-BPD (9.8 vs. 13.4, p=0.001) Significantly improved impulsivity of BPD with G2 than G1 on the BIS (72.7 vs. 76.8, p=0.004) Significantly improved depression on the BDI (22.0 vs. 25.8, p=0.03) Significantly improved global impressions and functioning with G2 than G1 on the SCL-90 (12.5 vs. 14.1, p=0.03), CGI-S (4.4 vs. 4.7, p<0.001), CGI-I (2.7 vs. 3.8, p<0.001), and GAS (50.5 vs. 43.5, p<0.001) No significant differences between groups on suicide attempts and self-harm acts (BEST); on SCL-90, BDI, CGI, and GAS between 20 weeks and 1 year; or SAS at 20 weeks or 20 weeks to 1 year Incidence of AEs: NR Withdrawals due to AEs: NR	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		management skills training 20 weeks			Attrition: 42% Differential attrition: G1: 29% (21/72) G2: 52% (48/93)	
Bos et al. (2010)	Design: RCT Setting: Outpatient, multicenter Country: Netherlands Funding: Other	N=79 G1 (37): TAU G2 (42): STEPPS plus individual therapy: 18 weekly sessionsof STEPPS and a follow-up session 3-6 months after the intervention; components included psychoeducation about BPD, emotion management skills training, and behavior management skills training 24 weeks	Inclusion: Met DSM-IV criteria for BPD by administering BPD modules and the SCID-II; BDSI-IV with scores exceeding the established cutoff on 1 or both subscales Exclusion: Did not speak Dutch; cognitively impaired (IQ<70); age <18 years; treated involuntary; presented an imminent danger to themselves or others	Mean (SD) age: G1: 32 (9.2) G2: 32 (5.6) % Female: 86 % Race/ethnicity: NR	Primary outcome: BPD-specific (BPD-40) and general psychiatric symptoms (SCL-90) at 1 year Significantly improved BPD-specific symptoms (BPD-40: 78.2 vs. 88.6, p=0.001), general psychiatric symptoms (SCL-90: 199.2 vs. 222.7, p=0.001) and quality of life (WHOQOL-Bref: 12.6 vs. 11.3, p=0.006) for G2 than G1 Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 17% Differential attrition: G1: 21% (9/42) G2: 11% (4/37)	Moderate
González-González et al. (2021)	Design: Prospective cohort Setting: Outpatient, single center Country: Spain Funding: None	N=118 G1 (98 [data analysis based on 28]): TAU G2 (20 [data analysis based on 9]): STEPPS: 20 weekly sessions of group STEPPS psychotherapy, 5 sessions of group psychotherapy for companions, monthly sessions of individual and family psychotherapy, and the possibility of therapy in case of an emergency; combined with usual	Inclusion: DSM-5 BPD diagnosis including self-harm or aggressive impulsive behaviors for the past 2 years Exclusion: Acute patients or those with a comorbid pathology; cognitive, intellectual, or psychopathological impairment for daily life activities requiring care in a rehabilitation center; receiving another psychotherapy treatment	Mean (range): 34 (18-58) % Female: 85 % Race/ethnicity: NR	Primary outcome: NR Significantly improved BPD-specific symptoms (BEST: 47.3 [14.1] vs. 28.8 [10.9], p<0.01) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 59% Differential attrition: G1: 55% (11/20) G2: 71% (70/98)	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		medication and/or psychiatric consultations 18 months				

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BDSI-IV, Borderline Syndrome Index IV; BEST, Borderline Evaluation of Severity Over Time; BIS, Barratt Impulsivity Scale; BPD, borderline personality disorder; BPD-40, Borderline Personality Disorder checklist-40; CGI, Clinical Global Impressions; CGI-I, Clinical Global Impression-Improvement; CGI-S, Clinical Global Impression-Severity; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; DSM-5, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; G1, Group 1; G2, Group 2; GAS, Global Assessment Scale; IQ, intelligence quotient; N, sample size; NIMH, National Institute of Mental Health; N, sample size; NR, not reported; RCT, randomized controlled trial; SAS, Social Adjustment Scale; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90, Symptom Checklist-90; SD, standard deviation; STEPPS, systems training for emotional predictability and problem solving; TAU, treatment as usual; WHOQOL-Bref, World Health Organization Quality of Life Bref; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Transference-Focused Psychotherapy vs. Treatment as Usual

Table D—26. Study characteristics and main results of transference-focused psychotherapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Doering et al. (2010)	Design: RCT Setting: Outpatient, multicenter Countries: Austria and Germany Funding: Other, Austrian bank	N=104 G1 (52): TAU: treatment by community psychotherapists and medication treatments as needed G2 (52): TFP: 2 50-minute sessions every week from experienced clinical psychologists or medical doctors; medications as needed 12 months	Inclusion: Female; age 18-45 years; DSM-IV BPD diagnosis; sufficient knowledge of the German language Exclusions: Antisocial personality disorder, schizophrenia, or bipolar I and II disorder with a major depressive, manic, or hypomanic episode during the previous 6 months; SUD during the previous 6 months; organic pathology or intellectual disability	Mean (SD) age: G1: 27 (7.5) G2: 28 (6.8) % Female: 100 % Race/ethnicity: NR	Primary outcome: Suicide attempts, dropout from therapy at 12 months Significantly fewer suicide attempts with G2 than G1 (13.7% vs. 21.2%, p=0.009) for LOCF analysis but not for completers analysis (p≥0.025) G2 significantly more effective than G1 for achieving fewer than 5 DSM-IV criteria for BPD (42.3% vs. 15.4%, p=0.002) and on GAF (58.62 vs. 56.06, p=0.002) No significant differences for self-harm acts, severity of symptoms, depression (BDI), and anxiety (STAI) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 34% Differential attrition: G1: 42.3% (22/52) G2: 25% (13/52)	High

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; LOCF, last observation carried forward; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; STAI, State-Trait-Anxiety Inventory; SUD, substance use disorder; TAU, treatment as usual; TFP, transference-focused psychotherapy.

Transference-Focused Psychotherapy vs. Schema-Focused Therapy

Table D—27. Study characteristics and main results of transference-focused psychotherapy compared with schema-focused therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Giesen-Bloo et al. (2006); Spinhoven et al. (2007)	Design: RCT Setting: Outpatient, multicenter	N=88 G1 (43): TFP: 50-minute sessions twice a week from therapists	Inclusion: Age 18-60 years; DSM-IV BPD diagnosis; BPDSI-IV score >20	Mean (SD) age: G1: 29.5 (6.5) G2: 31.7 (8.9)	Primary outcome: BPDSI-IV at 36 months G2 significantly more effective than G1 to improve BPDSI-IV at 36 months	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Country: Netherlands Funding: Government, Dutch Health Care Insurance Board	G2 (45): SFT: 50-minute sessions twice a week from trained therapists 3 years	Exclusion: Psychotic disorders, bipolar disorder, dissociative identity disorder, antisocial personality disorder, or ADHD; addiction requiring clinical detoxification; psychiatric disorders secondary to medical conditions	% Female: 93 % Race/ethnicity: NR	(16.24 vs. 21.87, p=0.005, RR=2.33, 95% CI 1.24-4.37) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 39% Differential attrition: G1: 51% (22/43) G2: 27% (12/45)	

Abbreviations. ADHD, attention-deficit/hyperactivity disorder; AE, adverse event; BPD, borderline personality disorder; BPDSI-IV, Borderline Personality Disorder Severity Index-IV; CI, confidence interval; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; RCT, randomized controlled trial; RR, risk ratio; SD, standard deviation; SFT, schema-focused therapy; TFP, transference-focused psychotherapy.

Special Populations

BPD and Substance Use Disorder: Comprehensive Validation Therapy Plus 12-Step vs. Dialectical Behavior Therapy

Table D—28. Study characteristics and main results of comprehensive validation therapy plus 12-step compared with dialectical behavior therapy in patients with BPD and substance use disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Linehan et al. (2002)	Design: RCT Setting: Outpatient, single center Country: United States Funding: Government, NIDA, NIH	N=24 G1 (12): Weekly individual CVT plus 12S: weekly "12-and-12" NA group and case management and phone consultation as needed, and opiate agonist therapy G2 (12): Weekly DBT individual and group skills training, case management	Inclusion: Female; age 18-45 years; BPD diagnosis according to PDE and SCID-II; current opiate dependence according to SCID-I; no indication of treatment coercion (e.g., court-ordered/agency-ordered to retain housing) Exclusion: Did not meet criteria for BPD; met criteria for bipolar mood disorder; pregnant; did not complete pre-treatment and/or medical evaluation.	Mean (SD) age: 36 (7.3) % Female: 100% % Race/ethnicity: Caucasian: 66% African American: 26 Mixed (Asian and Hispanic American): 4	Primary outcome: Percentage of opiate-positive urine specimens At the end of 12 month-treatment, G2 significantly lower percentage of opiate-positive urine specimens than G1 (*t*=2.32, p<0.02); no significant differences at 12 months for any other outcomes No significant differences between G1 and G2 for percentage of opiate-positive urine specimens or parasuicidal behavior and on BSI or GAS at 16 months Incidence of AEs: NR	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		and phone consultation as needed, and opiate agonist therapy 12 months			Withdrawals due to AEs: NR Attrition: 21% Differential attrition: G1: 0% (0/12) G2: 42% (5/12)	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; BSI, Brief Symptom Inventory; CVT plus 12S, comprehensive validation therapy plus 12-Step; DBT, dialectical behavior therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAS, Global Assessment Scale; N, sample size; NA, Narcotics Anonymous; NIDA, National Institute on Drug Abuse; NIH, National Institute of Health; NR, not reported; PDE, Personality Disorders Exam; RCT, randomized controlled trial; SCID-I, Structured Clinical Interview for DSM-IV Axis I Disorders; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SD, standard deviation.

BPD and Substance Use Disorder: Substance Use Disorder Treatment vs. Mentalization-Based Treatment Plus Substance Use Disorder Treatment

Table D—29. Study characteristics and main results of substance use disorder treatment compared with mentalization-based treatment plus substance use disorder treatment in patients with BPD and substance use disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Philips et al. (2018)	Design: RCT Setting: Outpatient, multicenter Country: Sweden Funding: Multiple	N=46 G1 (22): Standard SUD treatment G2 (24): Standard SUD treatment plus combined individual and group MBT 18 months	Inclusions: Males and females; age 18-65 years; DSM-IV BPD and SUD diagnoses; currently undergoing treatment at a SUD treatment clinic Exclusions: Schizophrenia, schizoaffective disorder, bipolar disorder type I, cognitive impairment, autism spectrum disorders, or psychopathy; participation in psychotherapy outside of the study; not being able to communicate in the Swedish language	Mean (SD) age: 36.7 (9.6) % Female: 80.4 % Race/ ethnicity: NR	Primary outcome: BPDSI-IV, deliberate self-harm, suicide attempts, IIP, reflective functioning scale, GSI, at 18 months No significant difference between groups on any outcome measure at 18 months Attrition: 48% Differential attrition: <10%	High

Abbreviations. BPD, borderline personality disorder; BPDSI-IV, Borderline Personality Disorder Severity Index-IV; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GSI, Global Severity Index; IIP, Inventory of Interpersonal Problems; MBT, mentalization-based treatment; N, sample size; NR, not reported; RCT, randomized clinical trial; SD, standard deviation; SUD, substance use disorder.

BPD and Alcohol Use Disorder: Dynamic Deconstructive Psychotherapy vs. Treatment as Usual in the Community

Table D—30. Study characteristics and main results of dynamic deconstructive psychotherapy compared with treatment as usual in the community in patients with BPD and alcohol use disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Gregory et al. (2008; 2009; 2010)	Design: RCT Setting: Outpatient, single center Country: United States Funding: Other, university	N=30 G1 (15): TAU: combination of individual psychotherapy, medication management, alcohol and drug counseling, professional and self-help groups, and/or case management G2 (15): DDP: weekly, 1-hour sessions administered by the PI or by 1 of 5 psychiatry residents 12 months	Inclusions: Age 18-45 years; DSM-IV BPD diagnosis; active alcohol abuse or dependence Exclusions: Schizophrenia or schizoaffective disorder, intellectual disability, or a neurological condition that may produce secondary psychiatric symptoms (e.g., stroke, multiple sclerosis, partial complex seizures, traumatic brain injury)	Mean (SD) age: 29 (7.7) % Female: 80 % Race/ethnicity: White: 90 Black: 3.3 Hispanic or Latino: 3.3 American Indian or Alaska Native: 3.3	Primary outcome: Parasuide behavior, alcohol misuse, and institutional care at 12 months No significant difference between G2 and G1 for parasuicide behavior, alcohol misuse, and dissocation at 12 months G2 significant improvements in depression (21,0 vs. 25.9, p<0.05) and in core symptoms of BPD (on the BEST) (33.6 vs. 38.4, p<0.05) at 12 months Attrition at 12 months: 37% Differential attrition: <10%	High

Abbreviations. BEST, Borderline Evaluation of Severity Over Time; BPD, borderline personality disorder; DDP, dynamic deconstructive psychotherapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; N, sample size; PI, principal investigator; RCT, randomized clinical trial; SD, standard deviation; TAU, treatment as usual.

BPD and Eating Disorder: Cognitive-Behavioral Therapy vs. Dialectical Behavior Therapy

Table D—31. Study characteristics and main results of cognitive-behavioral therapy compared with dialectical behavior therapy in patients with BPD and eating disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Navarro-Haro et al. (2018)	Design: Non-RCT Setting: Outpatient, multicenter Country: Spain	N=118 G1 (47): Weekly individual CBT, weekly group session, and pharmacological treatment	Inclusion: Age ≥18 years; DSM-IV BPD and eating disorder diagnoses Exclusion: Psychotic disorder and/or bipolar I disorder; alcohol or other SUD; organic disease that could interfere	Mean (SD) age: 27 (8.8) % Female: 100 % Race/ethnicity: NR	Primary outcome: Suicide attempt frequency, NSSI at 6 months G2 significantly more improved on BDI than G1 (23.9 vs. 29.8, p=0.02) at 6 months No significant differences between groups for suicide attempts, NSSI, or on GAF after 6 months; no significant differences between	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: Government, national agency	G2 (71): Weekly individual DBT, weekly DBT group skills training, and pharmacological treatment 6 months	with the psychological treatment		groups for depression, emotional regulation, or resilience at 6 years Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 8% at 6 months and 41.5% at 6 years Differential attrition: ≤10 percentage points	

Abbreviations. AE, adverse event; BDI, Beck Depression Inventory; BPD, borderline personality disorder; CBT, cognitive-behavioral therapy; CGI, Clinical Global Impression; DBT, dialectical behavior therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; N, sample size; NR, not reported; NSSI, nonsuicidal self-injury; RCT, randomized clinical trial; SD, standard deviation; SUD, substance use disorder.

BPD and Eating Disorder: Specialist Supportive Clinical Management vs. Modified Mentalization-Based Treatment

Table D—32. Study characteristics and main results of specialist supportive clinical management compared with modified mentalization-based treatment in patients with BPD and eating disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Robinson et al. (2016) NOURISHED	Design: RCT Setting: Outpatient, multicenter Country: United Kingdom Funding: Government, NIHOUR	N=68 G1 (34): SSCM: 1 session every 1-4 weeks for 20-26 sessions over 1 year G2 (34): MBT: 1 individual and 1 group session per week for 1 year 12 months	Inclusion: Age ≥18 years; DSM-IV eating disorder and BPD diagnoses or "BPD symptoms" from DSM-IV (impulsivity in ≥2 potentially self-damaging areas, recurrent suicidal or self-mutilating behavior) Exclusion: Current psychosis; current inpatient or day-patient (3+ days/week); currently in individual or group psychological therapy; received MBT less than 6 months prior to randomization; organic brain disease leading to significant cognitive impairment; BMI <15	Mean (SD) age: 31 (9.9) % Female: 93 % Race/ethnicity: White: 84	Primary outcome: EDE global score at 18 months No significant differences between G1 and G2 on ZAN-BPD at 18 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 78% Differential attrition: G1: 85% (29/34) G2: 71% (24/34)	High

Abbreviations. AE, adverse event; BMI, body mass index; BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition;* EDE, Eating Disorder Examination; G1, Group 1; G2, Group 2; MBT, mentalization-based therapy; N, sample size; NIHOUR, National Institute for Health Research; NOURISHED, Nice OUtcomes for Referrals with Impulsivity, NR, not reported; RCT, randomized clinical trial; SSCM, specialist supportive clinical management; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder; SD, standard deviation.

BPD and Major Depressive Disorder: Cognitive Therapy Plus Fluoxetine vs. Interpersonal Therapy Plus Fluoxetine

Table D—33. Study characteristics and main results of cognitive therapy plus fluoxetine compared with interpersonal therapy plus fluoxetine in patients with BPD and major depressive disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bellino et al. (2007)	Design: RCT Setting: Outpatient, single center Country: Italy Funding: None	N=32 G1 (16): Weekly CT plus fluoxetine G2 (16): Weekly IPT plus fluoxetine 24 weeks	Inclusion: Met DSM-IV-TR criteria for BPD and a major depressive episode Exclusion: Lifetime diagnosis of delirium, dementia, amnestic or other cognitive disorders, schizophrenia or other psychotic disorders, or bipolar disorder; current substance abuse disorder; treated with psychotropic drugs or psychotherapy during the 2 months prior to the study	Based on report among completers: Mean (SD) age: 31 (5.8) % Female: 73 % Race/ethnicity: NR	Primary outcome: Ham-D at 24 weeks No significant differences between G1 and G2 on Ham-D, Ham-A, BDI-II, CGI-S, SOFAS, SAT-P at 24 weeks Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 19% Differential attrition: G1: 25% (4/16) G2: 13% (2/16)	Moderate

Abbreviations. AE, adverse event; BDI-II , Beck Depression Inventory-II; BPD, borderline personality disorder; CGI-S, Clinical Global Impression-Severity; CT, cognitive therapy; DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*; G1, Group 1; G2, Group 2; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; IPT, interpersonal therapy; N, sample size; NR, not reported; RCT, randomized clinical trial; SAT-P, Satisfaction Profile; SD, standard deviation; SOFAS, Social Occupational Functioning Assessment Scale.

BPD and Major Depressive Disorder: Interpersonal Psychotherapy Plus Fluoxetine vs. Clinical Management Plus Fluoxetine

Table D—34. Study characteristics and main results of interpersonal psychotherapy plus fluoxetine compared with clinical management plus fluoxetine in patients with BPD and major depressive disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bellino et al. (2006)	Design: RCT Setting: Outpatient, single center Country: Italy Funding: None	N=39 G1 (19): Clinical management plus fluoxetine 20-40 mg/day; initial fixed 20 mg/day with opportunity to increase to 40 mg/day beginning week 2	Inclusion: DSM-IV BPD diagnosis; met criteria for major depressive episode Exclusion: Lifetime diagnosis of delirium, dementia, amnestic or other cognitive disorders, or schizophrenia or other psychotic disorders; major depressive episode as an expression of bipolar disorder; current substance abuse	Mean (SD) age: 26 (3.7) % Female: 60 (reported as: the ratio of men to women was 3 to 5) % Race/ethnicity:NR	Primary outcome: NR G2 significantly more effective than G1 for improving symptoms of depression (measured by the Ham-D [9.1 vs. 12, p=0.005]) No significant differences between G2 and G1 for anxiety for clinical global impressions (measured by CGI-S) or anxiety (measured by Ham-A)	Moderate

		G2 (20): IPT in weekly 1-hour sessions plus fluoxetine 20-40 mg/day plus; initial fixed 20 mg/day with opportunity to increase to 40 mg/day beginning week 2 24 weeks	disorder; treatment with psychotropic drugs or psychotherapy during 2 months prior to study; female patients not using adequate birth control		Attrition: 17.9% (7/39) G1: 20.0% (4/20) G2: 15.8% (3/19)	

Abbreviations. BPD, borderline personality disorder; CGI-S, Clinical Global Impression-Severity; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; IPT, interpersonal therapy; N, sample size; RCT, randomized clinical trial; SD, standard deviation.

BPD and Posttraumatic Stress Disorder: Dialectical Behavior Therapy Alone vs. Dialectical Behavior Therapy Plus Dialectical Behavior Therapy-Prolonged Exposure

Table D—35. Study characteristics and main results of dialectical behavior therapy alone compared with dialectical behavior therapy plus dialectical behavior therapy-prolonged exposure in patients with BPD and posttraumatic stress disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Harned et al. (2014, 2018)	Design: RCT Setting: Outpatient, single center Country: United States Funding: Government, NIMH	N=26 G1 (9): DBT: weekly individual, group training, and therapist consultation team meeting and as needed phone consultation G2 (17): DBT-PE: weekly PE protocol and DBT as well as group DBT skills training and as needed phone consultation 1 year	Inclusion: Female; age 18-60 years; DSM-IV BPD and PTSD diagnoses; can remember at least some part of index trauma; recent and recurrent intentional self-injury; lives within commuting distance of the clinic Exclusion: Met criteria for a psychotic disorder, bipolar disorder, or intellectual disability; legally mandated to treatment; required primary treatment for another debilitating condition (i.e., life-threatening anorexia nervosa)	Mean (SD) age: 33 (12) % Female: 100 % Race/ethnicity: White: 81 Biracial: 15 Asian-American: 4	Primary outcome: PTSD (PSS-I), intentional self-injury (SASII) at 15 months Numerically greater improvements in suicide attempts, NSSI, PSS-I, SASII, Ham-A, Ham-D, and GSI across both groups; no statistical tests performed Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 42% Differential attrition: ≤10 percentage points	High

Abbreviations. AE, adverse event; BPD, borderline personality disorder; DBT, dialectical behavior therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GSI, Global Severity Index; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; N, sample size; NIMH, National Institute of Mental Health; NR, not reported; NSSI, nonsuicidal self-injury; PE, prolonged exposure; PSS-I, PTSD Symptom Scale-Interview; PTSD, posttraumatic stress disorder; RCT, randomized clinical trial; SASII, Suicide Attempt Self-Injury Interview; SD, standard deviation.

Adolescents With BPD: Manualized Good Clinical Care vs. Cognitive Analytic Therapy

Table D—36. Study characteristics and main results of manualized good clinical care compared with cognitive analytic therapy in adolescents with BPD.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Chanen et al. (2008)	Design: RCT Setting: Outpatient, single center Country: Australia Funding: Government, NHMRC other, VicHealth, Colonial	N=86 G1 (42): Weekly group standardized good clinical care G2 (44): Weekly CAT 24 months	Inclusion: Age 15-18 years; met 2-9 DSM-IV criteria for BPD; any personality disorder or disruptive behavior disorder symptom; low socioeconomic status; depressive symptoms; history of abuse or neglect Exclusions: Learning disability, psychiatric disorder, pervasive developmental disorder, or severe primary Axis I disorder; >9 sessions of specialist mental health treatment in the previous 12 months; sustained psychosis and met criteria for Early Psychosis Prevention and Intervention Centre	Mean (SD) age: NR % Female: NR % Race/ethnicity: NR	Primary outcome: Psychopathology, parasuicidal behavior, global functioning at 24 months No significant differences between G1 and G2 for parasuicidal behavior or on BPD Total Score, SOFAS at 24 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 9% Differential attrition: ≤10 percentage points	Moderate

Abbreviations. AE, adverse event; BPD, borderline personality disorder; CAT, cognitive analytic therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; N, sample size; NHMRC, National Health and Medical Research Council; NR, not reported; RCT, randomized clinical trial; SD, standard deviation; SOFAS, Social Occupational Functioning Assessment Scale.

Adolescents with BPD and Substance Use Disorder: Individual Drug Counselling vs. Integrative Borderline Personality Disorder-Oriented Adolescent Family Therapy

Table D—37. Study characteristics and main results of individual drug counselling compared with integrative borderline personality disorder-oriented adolescent family therapy in adolescents with BPD and substance use disorder.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Santisteban et al. (2015)	Design: RCT Setting: Outpatient, single center Country: United States	N=40 G1 (20): Twice weekly individual drug counseling and monthly family meeting with caregivers	Inclusion: Age 14-17 years; DSM-IV BPD and substance use diagnoses Exclusion: NR	Mean (SD) age: G1: 16 (0.8) G2: 16 (0.8) % Female: 38 % Race/ethnicity: Hispanic: 85	Primary outcome: Substance use, BPD behaviors at 12 months No significant differences between G1 and G2 for substance use or BPD behavior at 12 months Incidence of AEs: NR Withdrawals due to AEs: NR	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: Government	G2 (20): Twice weekly I-BAFT, including family therapy, individual therapy, and skills-building interventions 7 months			Attrition: 33% Differential attrition: G1: 40% (8/20) G2: 25% (5/20)	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; I-BAFT, integrative borderline personality disorder-oriented adolescent family therapy; N, sample size; NR, not reported; RCT, randomized clinical trial; SD, standard deviation.

Pharmacotherapy

Second-Generation Antipsychotics vs. Placebo

Table D—38. Study characteristics and main results of second-generation antipsychotics compared with placebo.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Black et al. (2014)	Design: Double-blinded RCT Setting: Outpatient, multicenter Country: United States Funding: AstraZeneca	N=95 G1 (29): Placebo G2 (33): Quetiapine ER (150 mg/day) G3 (33): Quetiapine ER (300 mg/day) 8 weeks	Inclusion: Males and females; age 18-45 years; DSM-IV criteria for personality disorders; ≥ 9 on the ZAN-BPD Exclusion: History of psychotic disorder, neurological condition, or cognitive impairment; current SUD or abuse; medically unstable; history of lack of response to a second-generation antipsychotic; pregnant or lactating; acutely suicidal	Mean (SD) age: G1: 30 (8.8) G2: 28 (8.0) G3: 30 (8.1) % Female: 30 % Race/ethnicity: European-Caucasian: 78 Other: 21	Primary outcome: ZAN-BPD at 8 weeks G2 (but not G3) significantly more effective than G1 on ZAN-BPD (p=0.03) G3 (but not G2) significantly more effective on SCL-90 than G1 (p=0.03) G2 and G3 significantly more effective on MOAS (p=0.01) No significant differences on BIS, MADRS, and SDS Incidence of AEs: G1: 86% (25/29) G2: 88% (29/33) G3: 91% (30/33) Withdrawals due to AEs: NR Attrition: 33% Differential attrition:	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
					G1: 21% (6/29) G2: 33% (11/33) G3: 42% (14/33)	
Bogenschutz and Nurnberg (2004)	Design: Double-blinded RCT Setting: Outpatient, single center Country: United States Funding: Eli Lilly	N=40 G1 (20): Placebo G2 (20): Olanzapine (2.5-20 mg/day) 12 weeks	Inclusion: Age 18-60 years; DSM-IV BPD diagnosis; medically stable Exclusion: Other psychiatric disorders, SUD, or actively suicidal	Mean (SD) age: 32 (10.3) % Female: 63 % Race/ethnicity: White: 58 Hispanic: 25 Asian/Pacific Islander: 8 Other: 10	Primary outcome: CGI-BPD at 12 weeks Significantly greater improvement of G2 than G1 on the CGI-BPD (p=0.03) No significant differences on SCL-90, Ham-A, Ham-D, MOAS, and GAF Incidence of AEs: NR Withdrawal due to AEs: G1: 0% (0/20) G2: 20% (4/20) Attrition: 43% Differential attrition: G1: 35% (7/20) G2: 50% (10/20)	High
Linehan et al. (2008)	Design: Double-blinded RCT Setting: University hospital Country: United States Funding: Eli Lilly	N=24 G1 (12): Placebo G2 (12): Olanzapine (5 mg/day) 6 months	Inclusion: Females; age 18-60 years; met SCID-II and Borderline Personality Disorder Examination criteria for BPD; MOAS irritability subscale ≥6 Exclusion: Schizophrenia, bipolar I disorder, schizoaffective disorder, MDD with psychotic features or other psychotic disorder, intellectual disability, seizure disorder, or SUD	Mean (SD) age: 37 (9.0) % Female: 100 % Race/ethnicity: White: 79 Black: 4 Native American: 4 Latino: 4 Other: 8	Primary outcome: NR No significant differences between G1 and G2 on MOAS and Ham-D and for self-inflicted injury Incidence of AEs: NR Withdrawal due to AEs: G1: 0% (0/12) G2: 8% (1/12) Attrition: 33% Differential attrition: ≤10 percentage points	High
Nickel et al. (2006; 2007)	Design: Double-blinded RCT	N=52 G1 (26): Placebo	Inclusion: Males and females; age ≥16 years; DSM-IV BPD diagnosis	Mean (SD) age: G1: 21 (4.6) G2: 22 (3.4)	Primary outcome: SCL-90-R, Ham-D, Ham-A, STAXI at 8 weeks G2 significantly greater improvements than G1 on SCL-90-R (15.0 vs. 4.9, p<0.001), Ham-	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Setting: University hospitals Country: Austria, Germany Funding: None	G2 (26): Aripiprazole (15 mg/day) 8 weeks Follow up: 18 months	Exclusion: Schizophrenia; current use of other psychotropic medication; past termination of aripiprazole; current psychotherapy; pregnancy; suicidal ideation; severe somatic illness; alcohol or drug abuse	% Female: 83 % Race/ethnicity: NR	D (6.4 vs. 2.1, p=0.002), Ham-A (7.0 vs. 3.3, p=0.007), and STAXI (13.6 vs. 5.7, p<0.001) 18-month follow-up for SCL-90-R: 17.9 vs. 1.4, p<0.01 Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 25% Differential attrition: ≤10 percentage points	
Pascual et al. (2008)	Design: Double-blinded RCT Setting: Outpatient, single center Country: Spain Funding: Pfizer, government funding	N=60 G1 (30): Placebo G2 (30): Ziprasidone (40-200 mg/day) 12 weeks	Inclusion: Males and females; age 18-45 years; DSM-IV BPD diagnosis; current use of medically accepted contraception for females Exclusion: Schizophrenia, drug induced psychosis, organic brain syndrome, alcohol or other SUD, bipolar disorder, intellectual disability, or major depressive episode in course; CGI-S ≥4	Mean (SD) age: G1: 29 (6.3) G2: 29 (6.0) % Female: 82 % Race/ethnicity: NR	Primary outcome: CGI-BPD at 12 weeks No significant differences on CGI-BPD, SCL-90, Ham-A, Ham-D, and clinical psychotic symptoms Incidence of AEs: G1: 13% (4/30) G2: 37% (11/30) Withdrawal due to AEs: G1: 0% (0/30) G2: 30% (9/30) Attrition: 52% Differential attrition: ≤10 percentage points	High
Schulz et al. (2008)	Design: Double-blinded RCT Setting: Outpatient, multicenter Country: Multicountry Funding: Eli Lilly	N=314 G1 (159): Placebo G2 (155): Olanzapine (2.5-20 mg/day) 12 weeks	Inclusion: Males and females; age 18-65 years; DSM-IV BPD diagnosis; ZAN-BPD total score of 9 Exclusion: Schizophrenia, bipolar I disorder, bipolar II disorder, delusional disorder, MDD, SUD, PTSD, panic disorder, or OCD; BMI <17; use of antidepressants, mood stabilizer, or antipsychotic medication within 1 week of	Mean (SD) age: G1: 32 (9.6) G2: 32 (9.5) % Female: 71 % Race/ethnicity: White: 87	Primary outcome: ZAN-BPD at 12 weeks No significant differences on ZAN-BPD, SCL-90-R, and MADRS SDS, GAF, MOAS: data NR Incidence of AEs: G1: 57% (90/159) G2: 66% (102/155) Withdrawal due to AEs: G1: 11% (18/159)	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
			randomization; new psychotherapy treatment		G2: 11% (17/155 Attrition: 43% Differential attrition: ≤10 percentage points	
Soler et al. (2005)	Design: Double-blinded RCT Setting: Outpatient, single center Country: Spain Funding: Eli Lilly	N=60 G1 (30): DBT plus placebo G2 (30): DBT plus olanzapine (5-20 mg/day) 12 weeks	Inclusion: Females; age 18-45 years; DSM-IV BPD diagnosis without comorbid, unstable axis I disorder; CGI-S score ≥4; not receiving psychotherapy Exclusion: NR	Mean (SD) age: G1: 26 (5.4) G2: 28 (6.3) % Female: 87 % Race/ethnicity: NR	Primary outcome: NR Significantly greater improvements for G2 than G1 on Ham-D (8.79 vs. 4.87, p=0.004) and the frequency of aggressive behavior (data NR, p=0.03) No significant differences on Ham-A, CGI-S, and episodes of suicide attempts and self-injury Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 30% Differential attrition: ≤10 percentage points	High
Zanarini and Frankenburg (2001)	Design: Double-blinded RCT Setting: Outpatients, single center Country: United States Funding: Eli Lilly	N=28 G1 (9): Placebo G2 (19): Olanzapine (2.5 mg/day) 6 months	Inclusion: Females; age 18-40 years; DSM-IV BPD diagnosis Exclusion: Major depressive disorder; previous treatment with olanzapine; currently on psychotropic medications; actively abusing alcohol or drugs	Mean (SD) age: G1: 26 (4.5) G2: 28 (7.7) % Female: 100 % Race/ethnicity: White: 71 Nonwhite: 29	Primary outcome: SCL-90 at 6 months G2 significantly greater improvements than G1 on 4 domains of the SCL-90 (interpersonal sensitivity, anxiety, anger/hostility, paranoia); overall score of SCL-90 NR Incidence of AEs: NR Withdrawal due to AEs: G1: 0% (0/9) G2: 16% (3/19) Attrition: 68% Differential attrition: G1: 89% (8/9) G2: 58% (11/19)	High
Zanarini et al. (2011b)	Design: Double-blinded RCT	N=451	Inclusion: Males and females; age 18-65 years; DSM-IV BPD	Mean (SD) age:	Primary outcome: ZAN-BPD at 12 weeks	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Setting: Outpatient, multicenter Country: Multicountry Funding: Eli Lilly	G1 (153): Placebo G2 (150): Olanzapine (2.5 mg/day) G3 (148): Olanzapine (5-10 mg/day) 12 weeks	diagnosis; ZAN-BPD total score ≥9 Exclusion: Schizophrenia, schizoaffective disorder, schizophreniform disorder, bipolar I disorder, bipolar II disorder, delusional disorder, MDD, SUD within the previous 3 months, PTSD, panic disorder, or OCD; actively suicidal; BMI <17; cluster A personality disorder; new psychotherapy within the 3 months prior to visit 1; use of anticholinergic medication as prophylaxis for extrapyramidal symptoms	G1: 34 (11.3) G2: 33 (11.2) G3: 33 (10.0) % Female: 74 % Race/ethnicity: White: 65 African descent: 7 East/Southeast Asian: 2 Western Asian: 0.2 Hispanic: 24.6 Other origin: 11.1	G3 significantly more effective than G1 on ZAN-BPD (-8.5 vs. -6.8, p=0.01; response: 74% vs. 60%, p=0.018) and SCL-90-R (-0.7 vs. -0.6, p<0.05) No significant differences between G1 and G3 on MADRS, GAF, and MOAS No significant differences between G1 and G2 on most outcome measures Incidence of AEs: G1: 61% (93/153) G2: 65% (98/150) G3: 67% (99/148) Withdrawal due to AE: G1: 3% (5/153) G2: 3% (5/150) G3: 6% (9/148) Attrition: 35% Differential attrition: ≤10 percentage points	

Abbreviations. AE, adverse event; BIS, Barratt Impulsiveness Scale; BMI, body mass index; BPD, borderline personality disorder; CGI, Clinical Global Impression Scale, CGI-BPD, Clinical Global Impression Scale for Borderline Personality Disorder; CGI-S, Clinical Global Impression-Severity; DBT, dialectical behavior therapy; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; ER, extended release; G1, Group 1; G2, Group 2; G3, Group 3; GAF, Global Assessment of Functioning; Ham-A, Hamilton Rating Scale for Anxiety; Ham-D, Hamilton Rating Scale for Depression; MADRS, Montgomery-Åsberg Depression Rating Scale; MDD, major depressive disorder; MOAS, Modified Overt Aggression Scale; N, sample size; NR, not reported; OCD, obsessive compulsive disorder; PTSD, posttraumatic stress disorder; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90, Symptom Checklist-90; SCL-90-R, Symptom Checklist-90-Revised; SD, standard deviation; SDS, Sheehan Disability Scale; STAXI, State-Trait Anger Expression Inventory; SUD, substance use disorder; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Second-Generation Antipsychotic vs. Antidepressant

Table D—39. Study characteristics and main results of second-generation antipsychotics compared with antidepressants.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Zanarini et al. (2004c)	Design: double-blinded RCT	N=45	Inclusions: Female; age 18-40 years; DSM-IV BPD	Mean (SD) age: 23 (5.7)	Primary outcome: NR	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Setting: Outpatient, single center Country: United States Funding: Eli Lilly	G1 (14): Fluoxetine (10-30 mg/day) G2 (16): Olanzapine (2.5-7.5 mg/day) G3 (15): Fluoxetine (10-30 mg/day) and olanzapine (2.5-7.5 mg/day) 8 weeks	diagnosis; does not meet criteria for current MDD Exclusion: Current MDD, current or lifetime schizophrenia, schizoaffective disorder, or bipolar disorder; current use of psychotropic medications; medical illness; seizure disorder; substance abuse; acutely suicidal	% Female: 100 % Race/ethnicity: White: 80	G2 and G3 signficantly more effective than G1 on MOAS (19.7 vs. 20.2 vs. 15.4, p=0.003 for G2 vs. G1, p<0.001 for G3 vs. G1) at 8 weeks G2 and G3 signficantly more effective than G1 on MADRS (13.6 vs. 11.9 vs. 8.2, p<0.001 for G2 vs. G1; p=0.02 for G3 vs. G1) at 8 weeks Incidence of AEs: NR Withdrawal due to AEs: G1: 7% (1/14) G2: 0% (0/16) G3: 7% (1/15) Attrition: 7% Differential attrition: ≤10 percentage points	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; G3, Group 3; MADRS, Montgomery-Åsberg Depression Scale; MDD, major depressive disorder; MOAS, Modified Overt Aggression Scale; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation.

Second-Generation Antipsychotics vs. Second-Generation Antipsychotics

Table D—40. Study characteristics and main results of second-generation antipsychotics compared with second-generation antipsychotics.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bozzatello et al. (2017)	Design: Double-blinded RCT Setting: Outpatient, single center Country: Italy Funding: None	N=51 G1 (26): Olanzapine (5-10 mg/day) G2 (25): Asenapine (5-10 mg/day) 12 weeks	Inclusion: Age 18-50 years; DSM-5 BPD diagnosis Exclusion: Dementia, schizophrenia or other psychotic disorders, bipolar disorders, co-occurring major depressive episode, or substance abuse; past use of psychotropic medications and/or psychotherapy	Mean (SD) age: 25 (5.3) % Female: 63 % Race/ethnicity: NR	Primary outcome: NR No significant differences between G1 and G2 on BPDSI, CGI-S, BIS, MOAS, Ham-D, and SHI at 12 weeks Incidence of AEs (completers): G1: 26% (5/19) G2: 19% (4/21) Withdrawal due to AEs: G1: 11% (2/19) G2: 10% (2/21) Attrition: 22% Differential attrition: ≤10 percentage points	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
García-Carmona et al. (2021)	Design: Retrospective cohort study Setting: Outpatient; multicenter Country: Spain Funding: None	N=116 G1 (66): Oral antipsychotics G2 (50): LAI antipsychotics 1-3 months	Inclusion: Age ≥18 years; DSM-5 BPD diagnosis; treated with an oral or a LAI second-generation antipsychotic continuously for >12 months Exclusion: Patients who were institutionalized, with an intellectual disability or other psychiatric disorder, had concomitant use of 2 LAI antipsychotics, or with missing clinical records	Mean (SE) age: G1: 42.4 (1.4) G2: 39.4 (1.7) % Female: 46 % Race/ethnicity: NR	Primary outcome: NR G1 significantly more emergency department visits than G2 (7.9 vs. 6.2, p=0.041) No significant differences in suicidal behavior and hospital admissions Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: NR Differential attrition: NR	High

Abbreviations. AE, adverse event; BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; CGI-S, Clinical Global Impression-Severity; DSM-5, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; G1, Group 1; G2, Group 2; Ham-D, Hamilton Rating Scale for Depression; LAI, long-acting injectable; MOAS, Modified Overt Aggression Scale; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SE, standard error; SHI, Self-Harm Inventory.

Anticonvulsants vs. Placebo

Table D—41. Study characteristics and main results of anticonvulsants compared with placebo.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Crawford et al. (2018) LABILE	Design: Double-blinded RCT Setting: Outpatient, multicenter Country: United Kingdom Funding: NIHOUR	N=276 G1 (139): Placebo G2 (137): Lamotrigine (200 mg/day) 52 weeks	Inclusion: DSM-IV BPD diagnosis Exclusion: Met diagnostic criteria for bipolar disorder (type I or II) or psychotic disorder; history of liver or kidney impairment	Mean (SD) age: G1: 36 (11.0) G2: 36 (11.0) % Female: 75 % Race/ethnicity: White: 89 Black: 4 Asian: 1 Other: 6	Primary outcome: ZAN-BPD at 52 weeks No significant differences on ZAN-BPD, SHI, SFQ, and EQ-5D-3L Incidence of AEs: G1: 67% (93/139) G2: 56% (77/137 Withdrawal due to AEs: G1: 1% (1/139) G2: 4% (4/137) Attrition: 29% Differential attrition: <10 percentage points	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Frankenburg and Zanarini (2002)	Design: Double-blinded RCT Setting: Community recruitment with advertisements Country: United States Funding: Abbott Laboratories	N=30 G1 (10): Placebo G2 (20): Divalproex sodium (250 mg/day) 24 weeks	Inclusion: Females; age 18-40 years; DIB-R and DSM-IV BPD and bipolar II disorder diagnoses Exclusion: Formerly treated with divalproex sodium; medically ill; seizure disorder; current substance abuse; current criteria for a major depressive episode or a hypomanic episode; current or lifetime criteria for schizophrenia, schizoaffective disorder, psychotic disorder, or bipolar I disorder	Mean (SD) age: G1: 26 (7.3) G2: 27 (7.4) % Female: 100 % Race/ethnicity: White: 67 Black: 10 Hispanic: 13 Biracial: 7	Primary outcome: MOAS, SCL-90-R (subscales on anger, interpersonal hostility, depression) at 24 weeks G2 significantly more effective than G1 on MOAS (3.0 vs. 1.9, p=0.03) and SCL-90-R subscales on anger/hostility (0.8 vs. 0.6, p=0.01) and interpersonal sensitivity (0.8 vs. 0.4, p=0.04) Incidence of AEs: NR Withdrawal due to AEs: G1: 30% (3/10) G2: 5% (1/20) Attrition: 63% Differential attrition: <10 percentage points	High
Hollander et al. (2001)	Design: Double-blinded RCT Setting: Outpatient, single center Country: United States Funding: Abbott Laboratories, NIMH	N=16 G1 (4): Placebo G2 (12): Divalproex sodium (250 mg/day) 10 weeks	Inclusion: DSM-IV BPD diagnosis Exclusion: Medical or neurological illness; psychotic disorders, substance abuse, bipolar disorder type 1 or 2, or major depressive disorder; suicidal ideation	Mean (SD) age: NR % Female: 52 % Race/ethnicity: White: 67 Black: 14 Hispanic: 19	Primary outcome: NR No significant differences on CGI-I, GAS, MOAS, and AQ Incidence of AEs: NR Withdrawal due to AEs: G1: 0% (0/4) G2: 0% (0/12) Attrition: 63% Differential attrition: G1: 100% (4/4) G2: 50% (6/12)	High
Loew et al. (2006)	Design: Double-blinded RCT Setting: Single center or multicenter	N=56 G1 (28): Placebo G2 (28): Topiramate (200 mg/day)	Inclusion: Females; age 18-35 years; DSM-IV BPD diagnosis Exclusion: Schizophrenia; current use of psychotropic medication or	Mean (SD) age: G1: 26 (5.7) G2: 25 (5.3) % Female: 100	Primary outcome: SCL-90-R, SF-36, and IIP at 10 weeks G2 significantly more effective than G1 on SCL-90-R (7.4 vs.1.8, p<0.001), SF-36 (data NR, p<0.01), and IIP (data NR)	Low

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Country: Germany and Austria Funding: None	10 weeks	psychotherapy; suicidal; substance abuse; severe somatic illness	% Race/ethnicity: NR	Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 7% Differential attrition: <10 percentage points	
Moen et al. (2012)	Design: Double-blinded RCT Setting: Outpatient, single center Country: United States Funding: Abbott	N=15 G1 (5): Placebo G2 (10): Divalproex sodium (NR) 12 weeks	Inclusion: Age 21-55 years; DSM-IV BPD diagnosis; ≥150 on the SCL-90; ≥5 on the SCID-II Exclusion: Current or past history of bipolar disorder, schizophrenia, or major depressive disorder with psychotic features; current psychotropic medication; acutely suicidal; SUD; seizure disorder and/or anticonvulsant medications	Mean (range) G1: 37 (22-51) G2: 34 (23-45) % Female: 80 % Race/ethnicity: White: 80 Black: 7 Hispanic: 7 Mixed: 7	Primary outcome: NR No significant differences on SCL-90, BIS, and BEST Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 40% Differential attrition: <10 percentage points	High
Nickel et al. (2004)	Design: Double-blinded RCT Setting: Community recruitment Country: Germany Funding: None	N=31 G1 (10): Placebo (50 mg/day) G2 (21): Topiramate (250 mg/day) 8 weeks	Inclusion: Females; age 20-35 years; DSM-IV BPD diagnosis Exclusion: Current schizophrenia, major depressive disorder, or bipolar disorder; current use of psychotropic medication, or psychotherapy; somatically ill; actively suicidal; substance abuse	Mean (SD) age: G1: 27 (NR) G2: 26 (NR) % Female: 100 % Race/ethnicity: NR	Primary outcome: STAXI at 8 weeks G2 significantly more effective than G1 on 4 out of 5 subscales on STAXI (p values from 0.05 to 0.01); no significant improvement on subscale assessing tendency to repress anger Incidence of AEs: NR Withdrawal due to AEs: G1: 0% (0/10) G2: 0% (0/21) Attrition: 6% Differential attrition: <10 percentage points	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Nickel et al. (2005)	Design: Double-blinded RCT Setting: Outpatient recruitment and community advertisement Country: Germany Funding: None	N=44 G1 (22): Placebo G2 (22): Topiramate (250 mg/day) 8 weeks	Inclusion: Males; age >18 years; DSM-IV BPD diagnosis Exclusion: Acute psychosis, severe major depressive or bipolar disorder; current use of psychotropic medication or psychotherapy; somatically ill; actively suicidal; SUD	Mean (SD) age: G1: 29 (NR) G2: 30 (NR) % Female: 0 % Race/ethnicity: NR	Primary outcome: STAXI at 8 weeks G2 significantly more effective than G1 on 4 out of 5 subscales on STAXI (p values from 0.05 to 0.01); no significant improvement on subscale assessing tendency to repress anger Incidence of AEs: NR Withdrawal due to AEs: G1: 0% (0/22) G2: 0% (0/22) Attrition: 5% Differential attrition: <10 percentage points	Moderate
Reich et al. (2009)	Design: Double-blinded RCT Setting: Outpatient, single center Country: United States Funding: GlaxoSmithKline	N=28 G1 (13): Placebo G2 (15): Lamotrigine (50-275 mg/day) 12 weeks	Inclusion: DSM-IV BPD diagnosis; ≥8 on DIB-R; "serious" score on the affective instability item of the ZAN-BPD; ≥14 on ALS Exclusion: Dementia, psychiatric disorder, bipolar disorder, psychotic disorder, or SUD; currently hospitalized; previous treatment with lamotrigine or psychotherapy; active suicidal or homicidal ideation	Mean (SD) age: G1: 35 (9.7) G2: 28 (9.5) % Female: 89 % Race/ethnicity: White: 89	Primary outcome: ALS, affective instability item of the ZAN-BPD at 12 weeks G2 significantly greater improvements on than G1 on ALS (0.71 vs. 0.40, p=0.012) and affective lability of the ZAN-BPD (1.5 vs. 1.1, p=0.043) No significant difference on ZAN-BPD Incidence of AEs: G1: 31% (4/13) G2: 40% (6/15) Withdrawal due to AEs: G1: 0% (0/13) G2: 0% (3/15) Attrition: 39% Differential attrition: <10 percentage points	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Tritt et al. (2005)	Design: Double-blinded RCT Setting: single center or multicenter Country: Germany and Austria Funding: None	N=27 G1 (9): Placebo G2 (18): Lamotrigine (200 mg/day) 8 weeks	Inclusion: Female; age 20-40 years; DSM-IV BPD diagnosis Exclusion: Schizophrenia, major depressive or bipolar disorder; current use of psychotropic medication or psychotherapy; somatically ill; actively suicidal; substance abuse	Mean (SD) age: G1: 29 (NR) G2: 29 (NR) % Female: 100 % Race/ethnicity: NR	Primary outcome: STAXI at 8 weeks G2 significantly more effective than G1 on all 5 subscales of STAXI (p values from <0.05 to <0.01; overall STAXI score: NR) G2 improved more than G1 with respect to all STAXI scales on assessments after 8 weeks of treatment Incidence of AEs: NR Withdrawal due to AEs: G1: 11% (1/9) G2: 6% (1/18) Attrition: 11% Differential attrition: ≤10 percentage points	Low

Abbreviations. AE, adverse event; ALS, Affective Liability Scale; AQ, Aggression Questionnaire; BEST, Borderline Evaluation of Severity Over Time; BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; CGI-BPD, Clinical Global Impression Scale for Borderline Personality Disorder; CGI-I, Clinical Global Impression-Improvement; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; EQ-5D-3L, European Quality of Life–5 Dimension, 3 level version; G1, Group 1; G2, Group 2; GAS, Global Assessment Scale; IIP, Inventory of Interpersonal Problems; LABILE, Lamotrigine and Borderline Personality Disorder: Investigating Long-Term Effects; MOAS, Modified Overt Aggression Scale; NIHOUR, National Institute for Health Research; NIMH, National Institute of Mental Health; N, sample size; NR, not reported; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90, Symptom Checklist-90; SCL-90-R, Symptom Checklist-90-Revised; SD, standard deviation; SF-36, Short Form Survey; SHI, Self-Harm Inventory; SFQ, Social Functioning Questionnaire; STAXI, State-Trait Anger Expression Inventory; SUD, substance use disorder; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Antidepressants vs. Placebo

Table D—42. Study characteristics and main results of antidepressants compared with placebo.

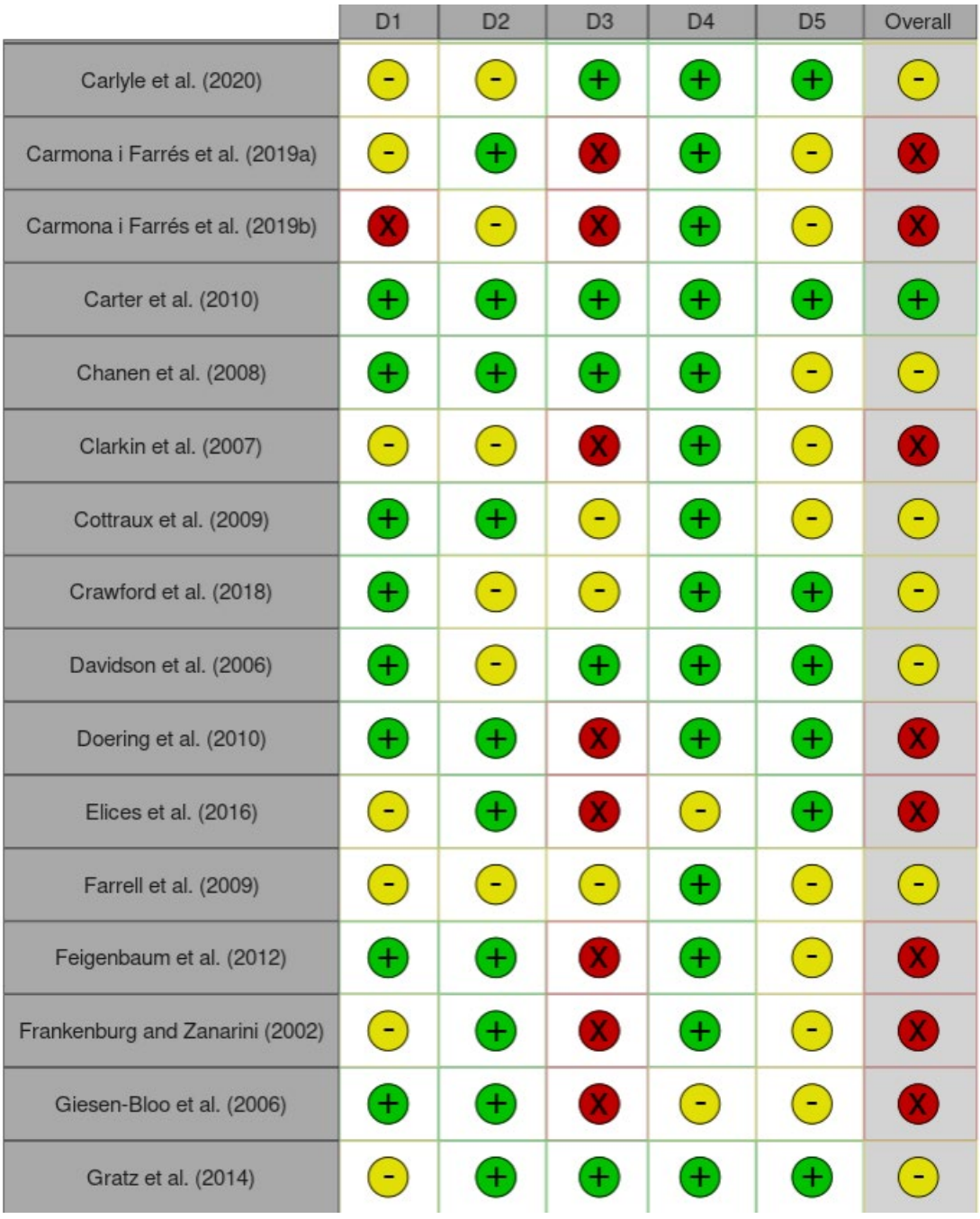
Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Simpson et al. (2004)	Design: Double-blinded RCT Setting: Outpatient, single center Country: United States Funding: Eli Lilly	N=25 G1 (13): Placebo G2 (12): Fluoxetine (40 mg/day) 12 weeks	Inclusion: Admissions to the Women's Partial Program; DSM-IV BPD diagnosis Exclusion: SUD; seizure disorder; unstable medical conditions; history of schizophrenia or bipolar	Mean (SD) age: 35 (10.1) % Female: 100 % Race/ethnicity: White: 72 Black: 20	Primary outcome: NR When corrected for multiple testing, no significant differences between G1 and G2 on STAXI, MOAS, or GAF at mean of 10 weeks Incidence of AEs: NR Withdrawal due to AEs: NR	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
			disorder; previous adequate trial of fluoxetine	Native American: 8	Attrition: 20% Differential attrition: ≤10 percentage points	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; MOAS, Modified Overt Aggression Scale; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; STAXI, State-Trait Anger Expression Inventory; SUD, substance use disorder.

Cochrane RoB. 2.0 Quality Ratings

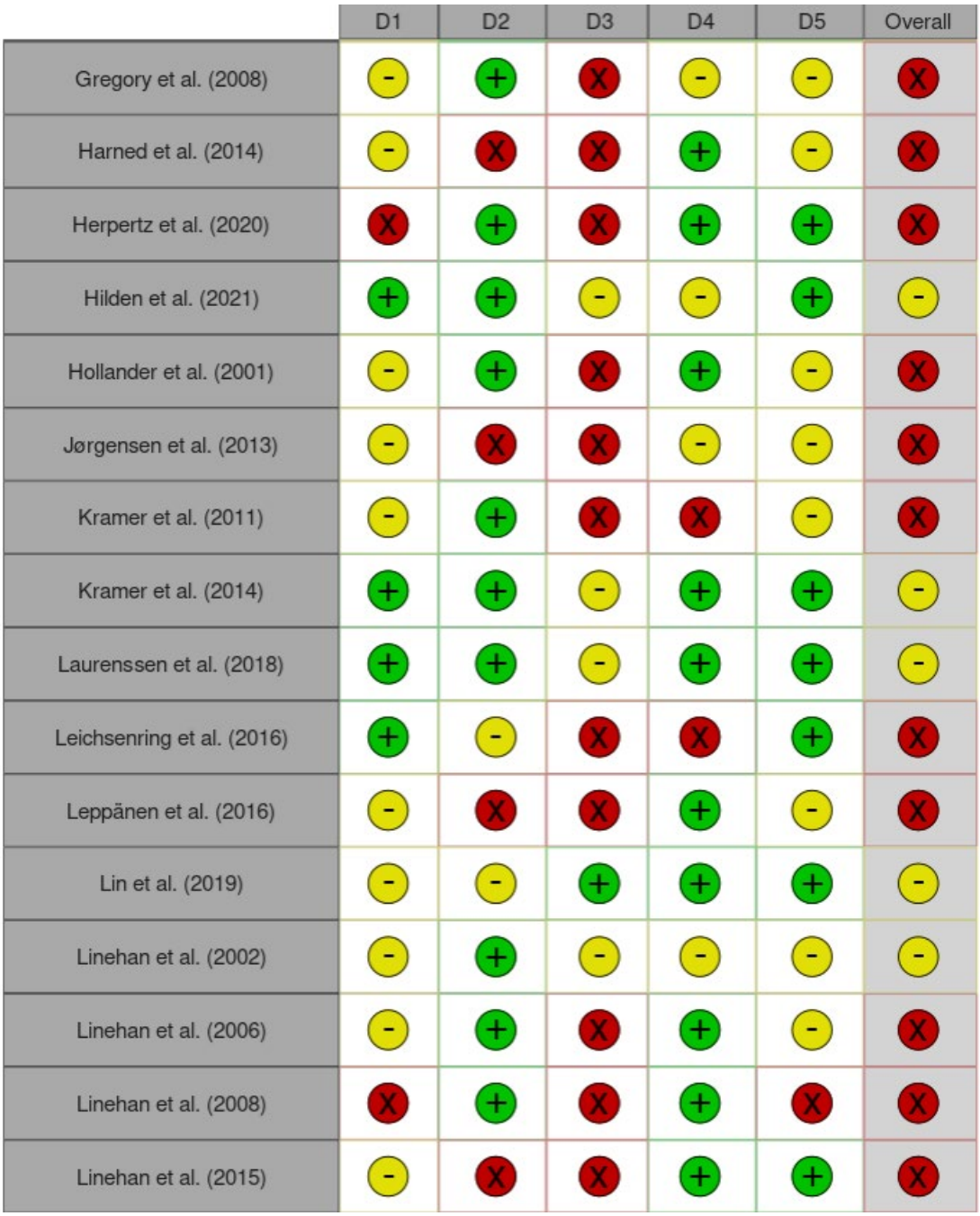
Risk of bias domains



Domains:
D1: Bias arising from the randomization process.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing outcome data.
D4: Bias in measurement of the outcome.
D5: Bias in selection of the reported result.

Judgement
High
Some concerns
Low

Risk of bias domains



Domains:
D1: Bias arising from the randomization process.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing outcome data.
D4: Bias in measurement of the outcome.
D5: Bias in selection of the reported result.

Judgement
High
Some concerns
Low

Risk of bias domains

Domains:
D1: Bias arising from the randomization process.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing outcome data.
D4: Bias in measurement of the outcome.
D5: Bias in selection of the reported result.

Judgement
High
Some concerns
Low

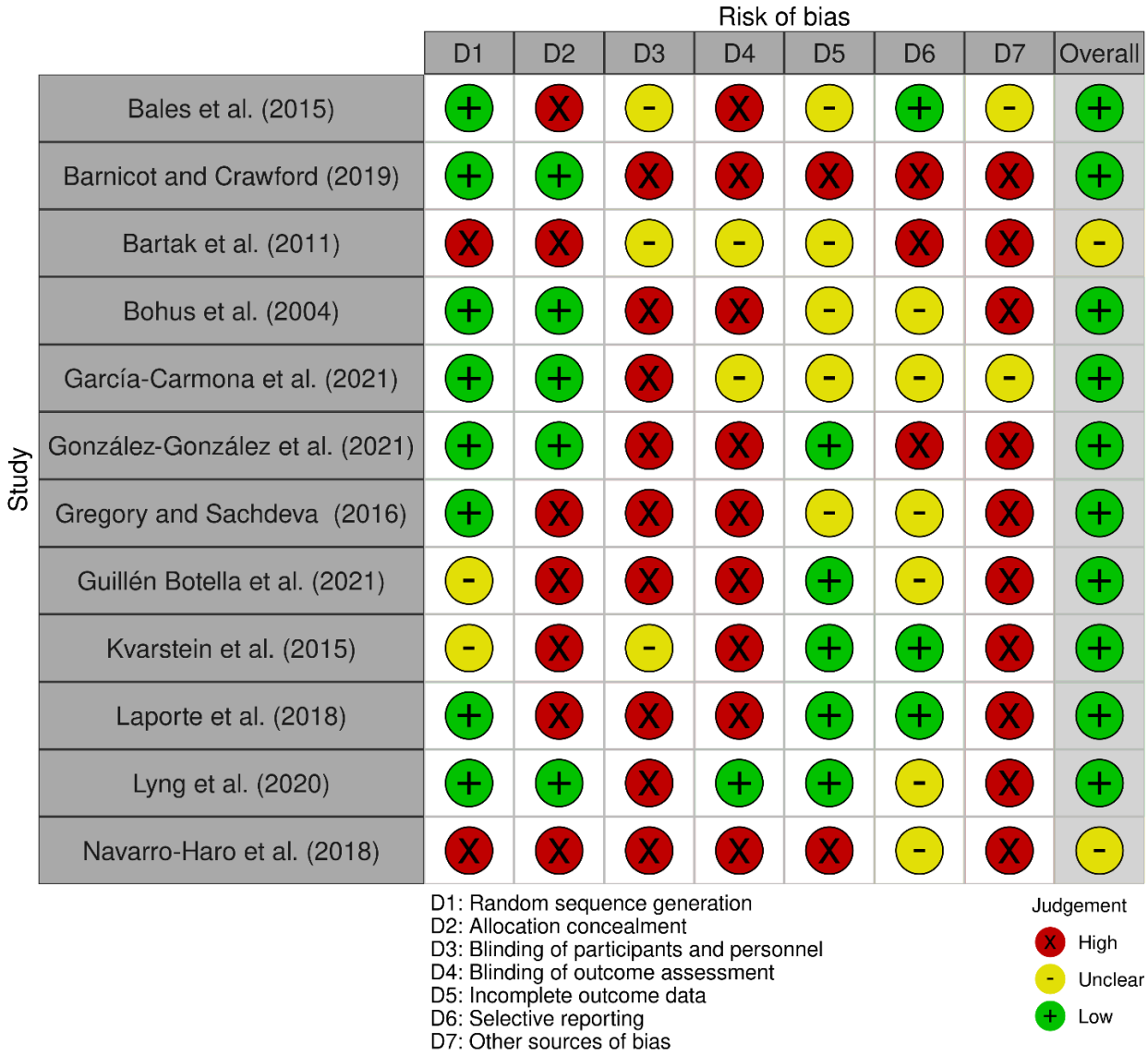
Risk of bias domains

	D1	D2	D3	D4	D5	Overall
Schmidt et al. (2021)	-	+	X	-	-	X
Schulz et al. (2008)	+	+	X	+	+	X
Simpson et al. (2004)	-	X	+	+	-	X
Sinnaeve et al. (2018)	+	+	X	X	+	X
Smits et al. (2020)	-	+	X	+	-	-
Soler et al. (2005)	-	+	X	+	-	X
Soler et al. (2009)	-	X	X	+	-	X
Tritt et al. (2005)	+	+	+	+	-	+
Verheul et al. (2003)	-	+	-	-	-	-
Walton et al. (2020)	+	+	-	+	+	-
Weinberg et al. (2006)	-	-	+	+	-	-
Zanarini and Frankenburg (2001)	-	X	X	-	-	X
Zanarini et al. (2004c)	-	-	+	+	-	-
Zanarini and Frankenburg (2008)	-	+	+	X	-	-
Zanarini et al. (2011b)	+	+	X	+	-	-
Zanarini et al. (2018)	-	+	+	-	+	-

Domains:
D1: Bias arising from the randomization process.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing outcome data.
D4: Bias in measurement of the outcome.
D5: Bias in selection of the reported result.

Judgement
X High
- Some concerns
+ Low

ROBINS-I Quality Ratings



Appendix F. Review of Benefits and Harms, Patient Preferences, Other Practice Guidelines, and Quality Measurement Considerations

Use of Guidelines to Enhance Quality of Care

Clinical practice guidelines can help enhance quality by synthesizing available research evidence and delineating recommendations for care on the basis of the available evidence. In some circumstances, practice guideline recommendations will be appropriate to use in developing quality measures. Guideline statements can also be used in other ways, such as educational activities or electronic decision support, to enhance the quality of care that patients receive. Furthermore, when availability of services is a major barrier to implementing guideline recommendations, improved tracking of service availability and program development initiatives may need to be implemented by health organizations, health insurance plans, federal or state agencies, or other regulatory programs.

Typically, guideline recommendations that are chosen for development into quality measures will advance one or more aims of the Institute of Medicine's report on "Crossing the Quality Chasm" (Institute of Medicine 2001) by facilitating care that is safe, effective, patient-centered, timely, efficient, and equitable. To achieve these aims, quality measures (Watkins et al. 2015) are needed that span the continuum of care (e.g., prevention, screening, assessment, treatment, continuing care), address the different levels of the health system hierarchy (e.g., system-wide, organization, program/department, individual clinicians), and include measures of different types (e.g., process, outcome, patient-centered experience). Emphasis is also needed on factors that influence the dissemination and adoption of evidence-based practices (Drake et al. 2008; Greenhalgh et al. 2004; Horvitz-Lennon et al. 2009a).

Often, quality measures will focus on gaps in care or on care processes and outcomes that have significant variability across specialties, health care settings, geographic areas, or patients' demographic characteristics. Administrative databases, registries, and data from electronic health record (EHR) systems can help to identify gaps in care and key domains that would benefit from performance improvements (Acevedo et al. 2015; Patel et al. 2015; Watkins et al. 2016). Nevertheless, for some guideline statements, evidence of practice gaps or variability will be based on anecdotal observations if the typical practices of psychiatrists and other health professionals are unknown. Variability in the use of guideline-recommended approaches may reflect appropriate differences that are tailored to the patient's preferences, treatment of co-occurring illnesses, or other clinical circumstances that may not have been studied in the available research. On the other hand, variability may indicate a need to strengthen clinician knowledge or address other barriers to adoption of best practices (Drake et al. 2008; Greenhalgh et al. 2004; Horvitz-Lennon et al. 2009a). When performance is compared among organizations, variability may reflect a need for quality improvement initiatives to improve overall outcomes but could also reflect case-mix differences such as socioeconomic factors or the prevalence of co-occurring illnesses.

Conceptually, quality measures can be developed for purposes of accountability, for internal or health system–based quality improvement, or both. Accountability measures require clinicians to report their rate of performance of a specified process, intermediate outcome, or outcome in a specified group of patients. Because these data are used to determine financial incentives or penalties based on

performance, accountability measures must be scientifically validated, have a strong evidence base, fill gaps in care, and be broadly relevant and meaningful to patients, clinicians, and policy makers. Development of such measures is complex and requires development of the measure specification and pilot testing (Center for Health Policy/Center for Primary Care and Outcomes Research and Battelle Memorial Institute 2011; Fernandes-Taylor and Harris 2012; Iyer et al. 2016; Pincus et al. 2016; Watkins et al. 2011). The purpose of the measure specification is to create detailed, clearly written, and precise instructions on the calculation of the measure so that, when implemented, the measure will be consistent, reliable, and effective in addressing quality in a specific target population (Centers for Medicare and Medicaid Services 2023). In contrast, internal or health system–based quality improvement measures are typically designed by and for individual providers, health systems, or payers. They typically focus on measurements that can suggest ways for clinicians or administrators to improve efficiency and delivery of services within a particular setting. Internal or health system–based quality improvement programs may or may not link performance with payment, and, in general, these measures are not subject to strict testing and validation requirements.

Regardless of the purpose of the quality measure, it must be possible to define the applicable patient group (i.e., the denominator) and the clinical action or outcome of interest that is measured (i.e., the numerator) in validated, clear, and quantifiable terms. The measure also needs to be feasible. More specifically, the health system's or clinician's performance on the measure must be readily ascertained from chart review, patient-reported outcome measures, registries, or administrative data. In addition, use of the measure should yield improvements in quality of care to justify any clinician burden (e.g., documentation burden) or related administrative costs (e.g., for manual extraction of data from charts, for modifications of EHRs to capture required data elements).

Documentation of quality measures can be challenging, and, depending on the practice setting, can pose practical barriers to meaningful interpretation of quality measures based on guideline recommendations. For example, when recommendations relate to patient assessment or treatment selection, clinical judgment may need to be used to determine whether the clinician has addressed the factors that merit emphasis for an individual patient. In other circumstances, standardized instruments can facilitate quality measurement reporting, but it is difficult to assess the appropriateness of clinical judgment in a validated, standardized manner. Furthermore, utilization of standardized assessments remains low (Fortney et al. 2017), and clinical findings are not routinely documented in a standardized format. Many clinicians appropriately use free text prose to describe symptoms, response to treatment, discussions with family, plans of treatment, and other aspects of care and clinical decision-making. Reviewing these free text records for measurement purposes would be impractical, and it would be difficult to hold clinicians accountable to such measures without advances in natural language processing technology and further increases in EHR use among mental health professionals.

Possible unintended consequences of any measures would also need to be addressed in testing the measure specifications in a variety of practice settings. For example, in many health care systems, multiple clinicians are involved in the care of a patient, and it is misleading if performance on the measure is attributed to the performance of a single clinician or group of clinicians. As another challenge,

if the measure specification requires precise wording for the measure to be met, clinicians may begin to document using standardized language that does not accurately reflect what has occurred in practice. If multiple discrete fields are used to capture information, data will be easily retrievable and reportable, but oversimplification is a possible unintended consequence of measurement and documentation burden is likely to be high (Johnson et al. 2021). Just as guideline developers must balance the benefits and harms of a particular guideline recommendation, developers of performance measures must weigh the potential benefits, burdens, and unintended consequences in optimizing quality measure design and testing.

Assessment and Determination of Treatment Plan

Statement 1 – Initial Assessment

APA *recommends* **(1C)** that the initial assessment of a patient with possible borderline personality disorder include the reason the individual is presenting for evaluation; the patient's goals and preferences for treatment; a review of psychiatric symptoms, including core features of personality disorders and common co-occurring disorders; a psychiatric treatment history; an assessment of physical health; an assessment of psychosocial and cultural factors; a mental status examination; and an assessment of risk of suicide, self-injury, and aggressive behaviors, as outlined in APA's Practice Guidelines for the Psychiatric Evaluation of Adults (3rd edition).

Benefits

Assessment of current and prior symptoms as well as previous treatment is beneficial in verifying that BPD is present and in identifying its severity and longitudinal course. Knowledge of the patient's current symptoms and functioning provides important baseline data for assessing the severity of the clinical presentation and effects of subsequent interventions. Assessment of risk factors, including risk of suicide, self-injury, and aggressive behaviors, is essential to developing a plan of treatment and determining an optimal treatment setting. Similarly, identification of co-occurring disorders and determination of the patient's goals and preferences for treatment will aid in the development of a comprehensive treatment plan.

Harms[1]

The harms of a detailed initial assessment are not well studied but are expected to be small, if any. It is possible that time used to focus on a detailed assessment could reduce time available to address other issues of importance to the patient or of relevance to diagnosis and treatment planning. Some individuals may have difficulty concentrating or may become frustrated if asked multiple questions during the evaluation. This could interfere with the therapeutic relationship between patient and clinician.

[1] Harms may include serious adverse events; less serious adverse events that affect tolerability; minor adverse events; negative effects of the intervention on quality of life; barriers and inconveniences associated with treatment; and other negative aspects of the treatment that may influence decision-making by the patient, the clinician, or both. Harms may also include opportunity costs for the clinician who may have to forgo another clinical activity that would be more beneficial for the patient.

Patient Preferences

Although there is no specific evidence on patient preferences related to assessment in individuals with BPD, clinical experience suggests that the majority of patients are cooperative with and accepting of these types of questions as part of an initial assessment.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as far outweighing the potential harms. This recommendation is also consistent with the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd edition (American Psychiatric Association 2016a). The level of research evidence is rated as low because there is minimal research on the benefits and harms of assessing these aspects of history and examination as part of an initial assessment of a patient with BPD. Nevertheless, expert opinion suggests that conducting such assessments as part of the initial psychiatric evaluation improves diagnosis and treatment planning in individuals with BPD. For additional details, see the Practice Guidelines for the Psychiatric Evaluation of Adults. For additional discussion of the research evidence, see Appendix C, Statement 1.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other organizations' practice guidelines typically assume that an evaluation has occurred and that a diagnosis of BPD has been made. The National Institute for Health and Care Excellence (NICE) guideline also notes the importance of assessing comorbid mental disorders, social problems, psychosocial and occupational functioning, coping strategies, strengths and vulnerabilities, risks to self and others, and needs for psychological treatment, social care and support, occupational rehabilitation or development, and assistance addressing needs of dependent children (National Institute for Health and Care Excellence 2009).

Quality Measurement Considerations

A detailed initial assessment of individuals with possible BPD is essential to verifying a diagnosis and establishing a comprehensive, patient-centered treatment plan. Nevertheless, it would be challenging to incorporate this recommendation into a performance-based quality measure given the breadth of content areas being assessed and the difficulty in ascertaining evaluation details from clinical charts or administrative data. However, quality-related efforts at the local level could assess whether EHR templates include prompts for documenting key elements of the assessment and whether such aspects of the evaluation are typically completed, while still allowing flexibility in the documentation of findings.

Statement 2 – Quantitative Measures

APA *suggests* **(2C)** that the initial psychiatric evaluation of a patient with possible borderline personality disorder include a quantitative measure to identify and determine the severity of symptoms and impairments of functioning that may be a focus of treatment.

Benefits

Use of a quantitative measure as part of the initial evaluation can have a number of benefits by establishing baseline information on the patient's symptom severity and associated impairment. As compared with a clinical interview, use of a quantitative measure may improve the consistency with which this information is obtained. When administered through paper-based or electronic self-report, use of quantitative measures may allow routine questions to be asked more efficiently. When used on a longitudinal basis, quantitative measures can minimize recall bias and help to determine whether treatment is having its intended effect or whether a shift in the treatment plan is needed to address symptoms, treatment-related side effects, level of distress, functioning impairments, or potential for harm to the patient or others. Ongoing use of quantitative assessments may also foster identification of residual symptoms or impairments and facilitate communication among treating clinicians.

Harms

The harms of using a quantitative measure include the time required for administration and review. Overreliance on quantitative measures may lead to overlooking other aspects of the patient's symptoms and clinical presentation. Patients may also provide inaccurate information about their symptoms, such as minimizing symptom severity or frequency, leading to an underestimation of severity of illness. Reliance on inaccurate information can have a negative impact on clinical decision-making, including recommendations for treatment. Some patients may view quantitative measures as impersonal or may feel frustrated by having to complete detailed questionnaires, resulting in possible straining of patient-clinician rapport. Changes in the workflow of clinical practices and adjustments in staffing may be needed to incorporate quantitative measures into routine care. Modification of EHRs or use of other technologies may also be required to facilitate capture of quantitative measure data.

Patient Preferences

Clinical experience suggests that the majority of patients are cooperative with and accepting of quantitative measures as part of an initial or subsequent assessment. Most patients will be able to appreciate the ways in which the use of quantitative measures will be of benefit to them. For example, in the testing of the DSM-5 Cross-Cutting Symptom Measure as part of the DSM-5 field trials, quantitative measures were found to be acceptable to patients (Clarke et al. 2014; Mościcki et al. 2013), and only a small fraction of individuals felt that measurement of symptoms would not be helpful to their treating clinician (Mościcki et al. 2013). The fact that the clinician is using a systematic approach to address the patients' symptoms and functioning may send a positive message that could improve the therapeutic relationship. Especially in developed countries, patients are used to and expect digital, computerized information exchange, including for health-related monitoring and communication. For these patients, the use of quantitative measures within the context of an electronic health record, mobile app, or other computerized technology may be more convenient.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as likely outweighing the potential harms. Quantitative measures of BPD symptoms have been used primarily in research settings and no specific scale for rating BPD symptoms can be recommended over another. Nonetheless, expert opinion suggests

that use of quantitative measures in the assessment of patients with BPD could enhance clinical decision-making and improve treatment outcomes. This statement is also consistent with Guideline VII, "Quantitative Assessment," in the APA's Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd edition (American Psychiatric Association 2016a). Although quantitative measures have been used for reporting purposes as well as research, the level of research evidence for this recommendation is rated as low because it remains unclear whether routine use of these scales in clinical practice improves overall outcomes. For additional discussion of the research evidence, see Appendix C, Statement 2.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other organizations' practice guidelines do not comment on the use of a quantitative measure, per se. However, several guidelines suggest that assessment include use of a structured or semi-structured clinical interview that focuses on diagnosis of personality disorders (Simonsen et al. 2019).

Quality Measurement Considerations

As a suggestion, this guideline statement is not appropriate for use as a performance-based quality measure or for incorporation into electronic decision support.

Statement 3 – Treatment Planning

APA *recommends* **(1C)** that a patient with borderline personality disorder have a documented, comprehensive, and person-centered treatment plan.

Benefits

Development and documentation of a comprehensive, person-centered treatment plan assures that the clinician has considered available treatment options in the context of individual patient needs, with a goal of improving overall outcome. It may also assist in forming a therapeutic relationship, eliciting patient preferences, permitting education about possible treatments, setting expectations for treatment, and establishing a framework for shared decision-making. Documentation of a treatment plan also promotes accurate communication among all those caring for the patient and can serve as a reminder of prior discussions about treatment.

Harms

The potential harms from this recommendation relate to the time spent in discussion and documentation of a comprehensive treatment plan that may reduce the opportunity to focus on other aspects of the evaluation.

Patient Preferences

Clinical experience suggests that patients are cooperative with and accepting of efforts to establish treatment plans, particularly when they are patient-centered.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as far outweighing the potential harms. The level of research evidence is rated as low because no information is available on the harms of a comprehensive, person-centered treatment plan. There is also minimal research on whether developing and documenting a specific treatment plan improves outcomes as compared with assessment and documentation as usual. However, indirect evidence, including expert opinion, supports the benefits of comprehensive treatment planning. For additional discussion of the research evidence, see Appendix C, Statement 3.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

The NICE guideline recommends development of comprehensive multidisciplinary care plans that include crisis planning, short-term treatment aims, approaches to management of comorbidities, and identification of long-term goals (National Institute for Health and Care Excellence 2009). The NICE and National Health and Medical Research Council guidelines also describe general aspects of treatment of BPD patients that are of relevance to treatment planning (National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009).

Quality Measurement Considerations

It is not known whether psychiatrists and other mental health professionals typically document a comprehensive and person-centered treatment plan, and there is likely to be variability. A quality measure could be developed to assess for the presence or absence of text in the medical record that would reflect treatment planning; however, clinical judgment would still be needed to determine whether a documented treatment plan is comprehensive and adapted to individual needs and preferences. Manual review of charts to evaluate for the presence of such a person-centered treatment plan would be burdensome and time-consuming to implement. Nevertheless, EHR note templates could include prompts to foster documentation of a patient-centered treatment plan, and local programs could engage in quality-related initiatives to improve aspects of treatment planning.

Statement 4 – Discussion of Diagnosis and Treatment

APA *recommends* **(1C)** that a patient with borderline personality disorder be engaged in a collaborative discussion about their diagnosis and treatment, which includes psychoeducation related to borderline personality disorder.

Benefits

Use of psychoeducation in patients with BPD has not been associated with a benefit in small non-representative research studies but expert opinion suggests that disclosure of diagnosis and associated psychoeducation are beneficial to patients.

Harms

The harms of psychoeducation are likely to be minimal on the basis of results from clinical trials in other psychiatric disorders that show no differences in the rate of harms experienced by individuals treated with psychoeducation as compared with usual care. It is possible that some individuals will not wish to know or would become upset by learning of their diagnosis, but this risk can be mitigated by collaborative and empathic discussion, including the benefits of treatment.

Patient Preferences

Disclosure of and discussion of a diagnosis of BPD is typically preferred by patients (Sulzer et al. 2016) and patients feel that it helps them be more informed about treatment options (Proctor et al. 2021). In addition, clinical experience suggests that most patients are interested in receiving information about their diagnosis and potential treatments as part of their care as well as being accepting of more formal and systematic approaches to psychoeducation. However, some patients may not wish to participate in psychoeducation or may experience logistical barriers (e.g., time, access to transportation, childcare, costs) in attending psychoeducation sessions.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as far outweighing the potential harms. Although patient preferences may differ, any minimal harms of psychoeducation or disclosure of diagnostic information seem to be outweighed by potential benefits of understanding of BPD and its treatment. For additional discussion of the research evidence, see Appendix C, Statement 4.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Two guidelines from other organizations also emphasize the importance of disclosing the diagnosis of BPD to the patient and providing psychoeducation, with a particular emphasis on the availability of effective treatment (National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009).

Quality Measurement Considerations

This guideline statement may not be appropriate for a performance-based quality measure because of the diversity of psychoeducational approaches and services. In addition, providing information on the diagnosis of BPD and its treatment will likely span multiple visits. Furthermore, documentation of diagnostic disclosure and psychoeducation will typically occur in free text notes, which are difficult to track for quality measurement purposes. Reminders about psychoeducation are also not well suited to incorporation into electronic health record clinical decision support. However, health organizations and health plans may wish to implement quality improvement efforts to increase diagnostic disclosure and psychoeducation among individuals with BPD.

Psychosocial Interventions

Statement 5 – Psychotherapy

APA *recommends* **(1B)** that a patient with borderline personality disorder be treated with a structured approach to psychotherapy that has support in the literature and targets the core features of the disorder.

Benefits

Use of psychotherapy in the treatment of BPD is associated with improvements in functioning and reductions in BPD severity, general psychopathology, depression, impulsivity, and suicidal and other self-harming behaviors, although different psychotherapies show different patterns of treatment benefits (moderate strength of research evidence).

Harms

The harms of psychotherapy in the treatment of BPD are not well reported in the literature. However, the harms of an effective psychotherapy delivered by a well-trained and well-supervised psychotherapist appear to be small. In contrast, the use of a psychotherapy that lacks demonstrated benefits in BPD could prevent individuals from receiving effective psychotherapy in a timely fashion, thereby influencing prognosis. Other harms of psychotherapy have been noted in individual circumstances when an evidence-based therapy is not delivered in a rigorous and systematic fashion. Such harms may result from boundary violations, alienation from support systems, apparent recollection of false memories, and undue dependency on psychotherapy, among other iatrogenic harms. In patients who have experienced prior trauma, intense or premature exploration of these experiences can increase patient distress, exacerbate symptoms, and disrupt the therapeutic relationship.

Patient Preferences

Clinical experience suggests that most patients are accepting of psychotherapy as part of a treatment plan. A meta-analysis of patient treatment preferences among individuals with a psychiatric disorder suggests a preference for psychotherapy over pharmacotherapy, with this preference being more pronounced among women and younger individuals (McHugh et al. 2013). However, patients also may have concerns about treatment cost or geographical availability that would influence their choice of psychotherapeutic approaches. In addition, some patients may prefer one type of psychotherapy over another based on personal experience or knowledge about a specific approach. Other patient and clinician factors may affect the therapeutic relationship and may also influence patient preferences.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. For additional discussion of the research evidence, see Appendix C, Statement 5. It was recognized that several psychotherapies have demonstrated efficacy in BPD. The harms of these treatments are not well studied but seem small when treatment is done by well-trained professionals using a rigorous evidence-based therapy. However, no single psychotherapy can be recommended over other effective psychotherapies in BPD. In addition, efficacies overlap among treatments, and the effects of treatment vary for different outcomes. Furthermore, patient preferences for specific therapies may differ, and

additional research evidence may influence our knowledge of effective psychotherapies for this condition. Thus, in balancing of benefits and harms, the guideline statement focuses on the use of an effective evidence-based psychotherapy for BPD rather than a specific psychotherapeutic modality.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other organizations' practice guidelines recommend use of structured psychotherapies that are intended to treat BPD (Canadian Agency for Drugs and Technologies in Health 2018; National Health and Medical Research Council 2012; Simonsen et al. 2019; The Finnish Medical Society Duodecim 2020). Outpatient treatment frequencies of up to two sessions per week and adapted to the patient's needs are recommended (National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009; Simonsen et al. 2019). Although the specific choice of a psychotherapy may depend on a number of factors including patient preference (National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009), psychotherapies that are specifically recommended are DBT (Canadian Agency for Drugs and Technologies in Health 2018; Herpertz et al. 2007; Simonsen et al. 2019; The Finnish Medical Society Duodecim 2020), MBT (Canadian Agency for Drugs and Technologies in Health 2018; Herpertz et al. 2007; Simonsen et al. 2019; The Finnish Medical Society Duodecim 2020), SFT (Canadian Agency for Drugs and Technologies in Health 2018; Herpertz et al. 2007; Simonsen et al. 2019), and TFP (Herpertz et al. 2007; Simonsen et al. 2019; The Finnish Medical Society Duodecim 2020). In women, DBT is also recommended if treatment goals for BPD include reductions in self-harm (National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009) or reductions in anger, anxiety, or depression (National Health and Medical Research Council 2012).

Quality Measurement Considerations

This guideline statement may not be appropriate for a performance-based quality measure because of the diversity of effective psychotherapeutic approaches and variations in the availability of psychotherapies. Measurement of psychotherapy utilization using structured EHR or claims data would require codes for specific types of therapy, but Current Procedural Terminology (CPT) codes refer to psychotherapy in general terms. In addition, patients may be receiving psychotherapies that include a mix of effective elements rather than rigid adherence to a specific psychotherapeutic approach, which would make it hard to specify use of a single modality. For these same reasons, reminders about psychotherapy would be difficult to incorporate into an EHR. In addition, most individuals with BPD are receiving some form of psychotherapy, and a gap in quality would need to be documented before pursuing additional quality measure development. Nevertheless, individual organizations and health plans may wish to implement programs to ensure that effective psychotherapies are being used to treat individuals with BPD.

Pharmacotherapy

Statement 6 – Clinical Review before Medication Initiation

APA *recommends* **(1C)** that a patient with borderline personality disorder have a review of co-occurring disorders, prior psychotherapies, other non-pharmacological treatments, past medication trials, and current medications before initiating any new medication.

Benefits

Review of co-occurring disorders, prior psychotherapies, other non-pharmacological treatments, past medication trials, and current medications have not been studied but are likely to be beneficial in assuring that the current treatment regimen is optimized prior to instituting a new medication. Such a review also increases awareness of possible drug-drug interactions with a new medication and may raise the possibility of discontinuing other medications or shifting the psychotherapeutic approach.

Harms

The harms of reviewing co-occurring disorders, prior psychotherapies, other non-pharmacological treatments, past medication trials, and current medications prior to starting a new medication have not been studied but are expected to be small, if any. Nevertheless, it is possible that time used to conduct such a review could delay medication initiation or reduce time available to address other issues of importance to the patient or of relevance to treatment planning.

Patient Preferences

Although there is no specific evidence on patient preferences related to conducting such a review before starting a new medication, clinical experience suggests that the majority of patients are cooperative with and accepting of careful consideration and discussion of treatment options.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as far outweighing the potential harms. The level of research evidence is rated as low because there is minimal research on the benefits and harms of assessing these aspects of history prior to initiating a new medication. Nevertheless, expert opinion suggests that conducting such an assessment would treatment planning and appropriateness of medication use in individuals with BPD. For additional details, see the Practice Guidelines for the Psychiatric Evaluation of Adults. For additional discussion of the research evidence, see Appendix C, Statement 6.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

The NICE guideline does not specify the type of review that was needed prior to instituting a medication but does note the importance of ensuring that medication is not begun in lieu of more appropriate interventions (National Institute for Health and Care Excellence 2009).

Quality Measurement Considerations

Reviewing co-occurring disorders, prior psychotherapies, other non-pharmacological treatments, past medication trials, and current medications prior to starting a new medication is likely to be beneficial to patients. Nevertheless, it would be challenging to incorporate this recommendation into a performance-based quality measure given the breadth of content areas being assessed and the difficulty in ascertaining evaluation details from clinical charts or administrative data. However, quality-related efforts at the local level could assess whether EHR templates include prompts for documenting key elements of the assessment and whether such aspects of the evaluation are typically completed, while still allowing flexibility in the documentation of findings.

Statement 7 – Pharmacotherapy Principles

APA *suggests* **(2C)** that any psychotropic medication treatment of borderline personality disorder be time-limited, aimed at addressing a specific measurable target symptom, and adjunctive to psychotherapy.

Benefits

Benefits of psychotropic medications in studies of BPD are modest (low strength of evidence) and inconsistent. Therapeutic benefits may be present for some patients that were not found in aggregated data from clinical trials, but the limitations of the evidence suggest that psychotropic medications should be used judiciously in BPD with a reliance on psychotherapy as a primary therapeutic modality. A focus on time-limited treatment that addresses a specific measurable target symptom is beneficial in assuring that treatment response will be assessed and the time of exposure to medication is minimized and dependent upon clinical response.

Harms

The harms of psychotropic medication in the treatment of BPD are, in part, dependent upon the side effect profile of the specific medication. In addition, patients may view psychotropic medications as a way to address intense feelings and emotions without engaging in the essential process of psychotherapy. The focus on time-limited treatment that addresses a specific measurable target symptom could potentially reduce long-term or non-specific use of a medication in a patient who may otherwise benefit from it.

Patient Preferences

Clinical experience suggests that most patients would prefer to minimize use of psychotropic medications due to adverse effects, costs, and other factors. Many patients, particularly women and younger individuals, prefer the use of psychotherapy to medications (McHugh et al. 2013). However, in some circumstances, patients may request medications to address specific symptoms or general experiences of distress. Some patients may also prefer one medication over another medication on the basis of prior treatment experiences or other factors.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as far outweighing the potential harms. Although the recommended approach has not been specifically studied, the harms of using psychotropic

medications in a time-limited, symptom-focused manner seem small compared to the benefits. Also, the benefits of psychotherapy clearly outweigh the benefits of psychotropic medications and the harms of psychotherapy are likely to be less than harms of pharmacotherapy, particularly when psychotherapy is evidence-based and conducted by well-trained and well-supervised psychotherapists. For additional discussion of the research evidence, see Appendix C, Statements 5 and 7.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Guidelines from other organizations also note that psychotropic medications should be used as adjuncts to psychotherapy and that use should be time-limited (National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009; Simonsen et al. 2019). The British BPD guideline recommends that the use of sedatives be time-limited (National Institute for Health and Care Excellence 2009), whereas other guidelines recommend avoiding the use of benzodiazepines in individuals with BPD (Herpertz et al. 2007; Simonsen et al. 2019; The Finnish Medical Society Duodecim 2020). Several guidelines note that use should be symptom-focused (Herpertz et al. 2007; National Health and Medical Research Council 2012; National Institute for Health and Care Excellence 2009) or intended to address co-occurring disorders (Simonsen et al. 2019).

Quality Measurement Considerations

This guideline statement would be difficult to incorporate into a meaningful performance-based quality measure. Although adjunctive use of psychotropic medications could be documented, it would be challenging to extract information from clinical documentation on whether medication was time-limited and symptom-focused in its use. By the same token, this statement would not be appropriate for use in clinical decision support in EHRs.

Statement 8 – Pharmacotherapy Review

APA *recommends* **(1C)** that a patient with borderline personality disorder have a review and reconciliation of their medications at least every 6 months to assess the effectiveness of treatment and identify medications that warrant tapering or discontinuation.

Benefits

The benefits of a review and reconciliation of medications include assuring that a complete list of medications is maintained, and potential drug-drug interactions are identified. In addition, such a review can identify medications that may warrant dose reduction or discontinuation as well as medications for which dose optimization is needed or laboratory monitoring is indicated (e.g., serum levels, metabolic studies).

Harms

The harms of medication review and reconciliation have not been studied but are likely to be small and related to time requirements.

Patient Preferences

No information is available on patient preferences related to medication review and reconciliation, but clinical experience suggests that patients are accepting and appreciative of review and discussion of treatment.

Balancing of Benefits and Harms

The potential benefits of this guideline statement were viewed as far outweighing the potential harms although evidence is limited. In addition, medication reconciliation and de-prescribing, where indicated, are recommended best practices in hospital as well as outpatient settings (Institute for Safe Medication Practice 2023; The Joint Commission 2022).

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Two other guidelines recommend periodic review of pharmacotherapies in patients with BPD with goals of tapering and discontinuing unneeded medications and avoiding polypharmacy (National Institute for Health and Care Excellence 2009; Simonsen et al. 2019).

Quality Measurement Considerations

As a recommended best practice in hospital as well as outpatient settings, medication reconciliation is already incorporated into other quality related measures in the U.S. (Institute for Safe Medication Practice 2023; The Joint Commission 2022). The addition of a measure that is specific to BPD would not be indicated.

Appendix G. Evidence Tables for Additional Studies Reviewed

Repetitive Transcranial Magnetic Stimulation vs. Sham

Table G—1. Study characteristics and main results of repetitive transcranial magnetic stimulation compared with sham.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Cailhol et al. (2014)	Design: Double-blinded RCT Setting: Outpatient, single center Country: France Funding: University Hospital of Toulouse	N=9 G1 (4): sham rTMS G2 (5): rTMS frequency: 10 Hz, 80% of motor threshold, total 2000 pulses/session; 10 sessions 2 weeks	Inclusion: Age 20-45 years; DSM-IV and DIB-R criteria for BPD Exclusion: Bipolar I disorder, alcohol dependency, current MDE or PTSD; contraindication to rTMS	Mean (SD) age: NR % Female: 89 % Race/ethnicity: NR	Primary outcome: BPDSI at 3 months No significant differences in BPDSI, MADRS, SCL-90, GAS Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 0% Differential attrition: 0%	Moderate

Abbreviations. AE, adverse event; BPD, borderline personality disorder; BPDSI, Borderline Personality Disorder Severity Index; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAS, Global Assessment Scale; Hz, hertz; MADRS: Montgomery-Åsberg Depression Scale; MDE, major depressive episode; N, sample size; NR, not reported; PTSD, posttraumatic stress disorder; RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; SCL-90, Symptom Checklist-90; SD, standard deviation.

Abandonment Psychotherapy vs. Treatment as Usual

Table G—2. Study characteristics and main results of abandonment psychotherapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Andreoli et al. (2016)	Design: RCT Setting: Outpatient, single center Country: Switzerland Funding: NR	N=170 G1 (30): TAU: intensive community treatment G2 (70): Manualized AP: 2 sessions/week delivered by nurses who had experience in the management of patients with BPD plus	Inclusion: Age 18-60 years; DSM-IV BPD and MDD diagnoses Exclusion: DSM-IV psychotic disorder, bipolar I disorder, SUD, or intellectual disability; inability to speak French	Mean (SD) age: 32 (10.1) % Female: 84 % Race/ethnicity: NR	Primary outcome: Suicidal relapse, rehospitalization, clinical remission (GAS >60) at 3 months G2 significantly more effective than G1 to reduce suicidal relapse (12.9% vs. 40.0%, p<0.005) and rehospitalization (14.3% vs. 36%, p<0.01), to achieve a 50% reduction in Ham-D (65.7% vs. 33.3%, p<0.005), and to improve GAS (62.7 vs. 36.7, p<0.01) Incidence of AEs: G1: 100% (30/30)	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		antidepressant medications 3 months			G2: 100% (70/70) Withdrawal due to AEs: NR Attrition: 12% Differential attrition: G1: 37% (11/30) G2: 6% (4/70)	

***Abbreviations*.** AE, adverse event; AP, abandonment psychotherapy; BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAS, Global Assessment Scale; Ham-D, Hamilton Rating Scale for Depression; MDD, major depressive disorder; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; SUD, substance use disorder; TAU, treatment as usual.

Schema-Focused Therapy vs. Treatment as Usual

Table G—3. Study characteristics and main results of schema-focused therapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Farrell et al. (2009)	Design: RCT Setting: Outpatient, multicenter Country: United States Funding: Government, NIMH	N=32 G1 (16): TAU: Weekly individual psychotherapy in the community G2 (16): SFT plus TAU: 30 weekly group sessions each lasting 90 minutes; combination of emotional awareness training, BPD psychoeducation, distress management training, and schema-focused change work; the sessions consisted of discussion of homework from the previous session, new information presentation, question-and-answer session, experiential or cognitive work, and homework assignment 8 months	Inclusion: Females; age 18-65 years; met criteria for BPD by the DIPD-R and the BSI; in individual psychotherapy of ≥6 months duration and stable Exclusion: Axis I diagnosis of a psychotic disorder or presence of psychosis; below average IQ (89) on the Shipley Institute of Living Scale	Mean (SD) age: G1: 36 (8.08) G2: 35 (9.30) % Female: 100 % Race/ethnicity: NR	Primary outcome: NR G2 significantly more effective than G1 for improving G1 at 14-month follow-up (6 months after end of treatment) for BPD diagnosis (measured by DIB-R [0% vs. 83%, p<0.001]), BPD symptoms (measured by BSI [15.75 vs. 33.08, p<0.001]), global severity of psychiatric symptoms (measured by SCL-90 [0.96 vs. 1.93, p<0.001]), and improved global functioning (measured by GAF [66.19 vs. 48.25, p<0.001]) Attrition: 12.5% (4/32) G1: 25% (4/16) G2: 0% (0/16)	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		Follow-up: 6 months				
Leppänen et al. (2016)	Design: RCT Setting: Outpatient, multicenter Country: Finland Funding: NR	N=71 G1 (47): TAU: Treatment in accordance with the current practices of Oulu city mental health care services; the treatments vary widely, from supportive weekly psychotherapy sessions to visits every few weeks, from occasional doctor's appointments for drug control to home rehabilitation G2 (24): SFT-based psycho-educational group integrated into individual therapy: 45-60 minute individual therapy sessions once a week, a total of 40 ninety minute psycho-educational group sessions (approximately once a week), and materials for patients to practice therapy exercises at home 12 months	Inclusion: Age ≥20 years; fulfilled the SCID-II criteria for BPD; severe symptoms of BPD which included parasuicidal behaviour (such as cutting, other forms of selfharm, impulsive overdosing of medicines); attempted suicide; considerable emotional instability affecting social and professional life; previous unsuccessful treatments (one or more) Exclusion: Schizophrenia spectrum diseases/psychoses, bipolar disorder (type I), neuropsychiatric disorder, severe substance abuse problem, axis I disorders diagnosed according to SCID-I, or presence of neuropsychiatric disorder	Mean (SD) age: G1: 32 (8.8) G2: 32 (8.3) % Female: 86 % Race/ethnicity: NR	Primary outcome: Borderline symmptoms by BPDSI-IV No difference between the groups on BPD outcomes Attrition: 26.8% (19/71) G1: 31.9% (15/47) G2: 16.7% (4/24)	High
Hilden et al. (2021)	Design: RCT Setting: Outpatient, single center Country: Finland Funding: Helsinki University Hospital	N=42 G1 (14): TAU: Psychiatrist visits and 45-minute therapy sessions of once monthly; both pharmacotherapy and some form of psychosocial support or psychotherapy for most patients G2 (28): SFT: 20 weekly 90-minute sessions 20 weeks	Inclusion: Adults; BPD with the DSM-IV SCID-II criteria (included those who had previously received treatment) Exclusion: Psychotic symptoms, suicide risk, principal diagnosis of uncontrollable SUD, or illness/symptoms affecting participation; those	Mean (SD) age: G1: 27 (3.7) G2: 31 (8.8) % Female: 83 % Race/ethnicity: NR	Primary outcome: Intra-individual change (expectedly decline) in borderline personality symptoms No difference between the groups on BPD outcomes Attrition: 16.7% (7/42) G1: 14.3% (2/14) G2: 17.97% (5/28)	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
			undergoing specific psychotherapy			

Abbreviations. BPD, borderline personality disorder; BPDSI-IV, Borderline Personality Disorder Severity Index-IV; BSI, Borderline Syndrome Index; DIB-R, Diagnostic Interview for Borderlines-Revised; DIPD-R, Diagnostic Interview for Personality Disorders-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; IQ, intelligence quotient; N, sample size; NIMH, National Institute of Mental Health; NR, not reported; SCID-I, Structured Clinical Interview for DSM-IV Axis I Disorders; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90, Symptom Checklist-90; SD, standard deviation; SFT, schema-focused therapy; SUD, substance use disorder; TAU, treatment as usual.

Schema-Focused Therapy vs. Schema-Focused Therapy With Extra Phone Support

Table G—4. Study characteristics and main results of schema-focused therapy compared with schema-focused therapy with extra phone support.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Nadort et al. (2009)	Design: RCT Setting: Outpatient, multicenter Country: Netherlands Funding: Other, public benefit organization	N=62 G1 (30): 45-minute sessions of SFT twice a week in year 1 and once a week in year 2 G2 (32): 45-minute sessions of SFT twice a week along with extra phone support outside office hours 18 months	Inclusion: Age 18-60 years; DSM-IV BPD diagnosis; BPDSI-IV score >20 Exclusion: Psychotic disorders, bipolar disorder, dissociative identity disorder, antisocial personality disorder, or ADHD; addiction of such severity that clinical detoxification was indicated; psychiatric disorders secondary to medical conditions	Mean (SD) age: G1: 32.13 (9.01) G2: 31.81 (9.24) % Female: 96.8 % Race/ethnicity: NR	Primary outcome: BPDSI-IV at 18 months No significant differences between G1 and G2 on BPD severity and burden as well as outcomes of global psychological problems, quality of life, and dysfunctional Incidence of AEs: NR Withdrawal due to AE: G1: 0 (0/30) G2: 3% (1/32) (suicide after treatment allocation but before treatment additional crisis support was provided) Attrition: 21% Differential attrition:≤10 percentage points	Moderate

Abbreviations. ADHD, attention-deficit/hyperactivity disorder; AE, adverse event; BPD, borderline personality disorder; BPDSI-IV, Borderline Personality Disorder Severity Index-IV; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; N, sample size; NR, not reported; RCT, randomized controlled trial; SFT, schema-focused therapy; SD, standard deviation.

Cognitive Rehabilitation vs. Psychoeducation

Table G—5. Study characteristics and main results of cognitive rehabilitation compared with psychoeducation.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Pascual et al. (2015)	Design: RCT Setting: Outpatient, multicenter Country: Spain Funding: Government, Centro de Investigación Biomédica en Red de Salud Mental, Instituto de Salud Carlos III, Fondo de Investigación Sanitaria	N=70 G1 (36): CR: twice weekly group sessions G2 (34): Psychoeducation: weekly group sessions 16 weeks	Inclusion: Age 18-45 years; outpatient; BPD diagnosis according to DSM-IV-TR and evaluated by SCID-II and DIB-R; CGI-BPD >4; GAF <65 Exclusion: Severe physical conditions that could affect neuropsychological performance; IQ <85; MDD or substance misuse within the last 6 months; schizophrenia, severe psychotic disorder, or bipolar disorder; previous participation in any psychoeducation or CR	Mean (SD) age: G1: 32 (6.04) G2: 33 (8.8) % Female: 74.3 % Race/ethnicity: NR	Primary outcome: Psychosocial functioning at 6 months No significant difference between G1 and G2 on psychosocial functioning including on BSL-23, FAST, BIS, Ham-A, MADRS at 6 months Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 40% Differential attrition: <10 percentage points	High

Abbreviations. AE, adverse event; BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; CGI-BPD, Clinical Global Impression Scale for Borderline Personality Disorder; CR, cognitive rehabilitation; BSL-23, Borderline Symptom List-23; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*; FAST, Functioning Assessment Scale Test; G1, Group 1; G2, Group 2; GAF, Global Assessment of Functioning; Ham-A, Hamilton Rating Scale for Anxiety; IQ, intelligence quotient; MADRS, Montgomery-Åsberg Depression Rating Scale; MDD, major depressive disorder; N, sample size; RCT, randomized controlled trial; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SD, standard deviation.

Cognitive Therapy vs. Rogerian Supportive Therapy

Table G—6. Study characteristics and main results of cognitive therapy compared with Rogerian supportive therapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Cottraux et al. (2009)	Design: RCT Setting: Outpatient, multicenter Country: France Funding: Other	N=65 G1 (32): Weekly individual RST for 6 months then biweekly individual RST for 6 months G2 (33): Weekly individual CT for 6 months then	Inclusion: DSM-IV BPD diagnosis (confirmed by DIB-R with a score of ≥8) Exclusion: Age <18 or >60 years; living too far from the centers; psychotic disorders with current delusions; significant drug or alcohol addiction in the foreground;	Mean (SD) age: G1: 32.6 (8.3) G2: 34.3 (10.2) % Female: 76.9 % Race/ethnicity: NR	Primary outcome: Combined response (score of ≤3 on CGI and Hopelessness score of <8) at 24 weeks No significant differences between G1 and G2 on the CGI-I, Hopelessness scale, Ham-D, or BAI at 24 weeks G2 significant improvement in BDI scores at 24 weeks (13.0 vs. 21.7, p=0.01)	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		biweekly individual RST for 6 months 1 year	antisocial behaviors; not following psychotherapy at the time of the study		Harms: NR Attrition: Week 24: 22% Week 104: 68% Differential attrition: ≤10 percentage points	

Abbreviations. BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; BPD, borderline personality disorder; CGI, Clinical Global Impression Scale; CGI-I, Clinical Global Impression-Improvement; CT, cognitive therapy; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; Ham-D, Hamilton Rating Scale for Depression; N, sample size; NR, not reported; RCT, randomized controlled trial; RST, rogerian supportive therapy; SD, standard deviation.

Motive-Oriented Therapeutic Relationship vs. General Psychiatric Management

Table G—7. Study characteristics and main results of motive-oriented therapeutic relationship compared with general psychiatric management.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Kramer et al. (2011)	Design: RCT Setting: Outpatient, single center Country: Switzerland Funding: NR	N=25 G1 (14): 10 session TAU with a manual-based psychiatric and psychotherapeutic approach G2 (11): 10 sessions of MOTR along with TAU 7 therapy sessions	Inclusion: Age 16-60 years; DSM-IV BPD diagnosis; speaks fluent French Exclusion: Organic disorder or persistent substance abuse/dependence which might affect brain function; psychotic disorder implying pronounced break in reality testing including schizophrenia, delusional disorder, and bipolar I disorder; acute risk of suicide; severe cognitive impairment	Mean (SD) age: 31 (10.59) % Female: 77 % Race/ethnicity: NR	Primary outcome: Psychotherapeutic results on the Outcome Questionnaire-45 after 7 therapy sessions No significant differences between G1 and G2 on psychotherapeutic results following 7 therapy sessions Attrition: 42% Differential attrition: G1: 57% (8/14) G2: 18% (2/11)	High
Kramer et al. (2014)	Design: RCT Setting: Outpatient, single center Country: Switzerland	N=85 G1 (43): 10 sessions of GPM	Inclusion: Age 18-65 years; DSM-IV BPD diagnosis Exclusion: DSM-IV psychotic disorders, intellectual disability, or substance abuse	Mean (SD) age: G1: 31 (11.00) G2: 35 (9.97) % Female: 68.9	Primary outcome: Psychotherapeutic results on the Outcome Questionnaire-45 at 3 months G2 significantly greater improvement on the Questionnaire-45 at 3 months (76.0 vs. 86.1, p<0.01)	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: Government, Swiss National Science Foundation	G2 (42): 10 sessions of GPM plus MOTR use of Plan Analysis 3 months		% Race/ethnicity: NR	No significant differences between G1 and G2 on the IIP and the BSL at 3 months Attrition: 29% Differential attrition: ≤10 percentage points	

Abbreviations. BPD, borderline personality disorder; BSL, Borderline Symptom List; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; G1, Group 1; G2, Group 2; GPM, general psychiatric management; IIP, Inventory of Interpersonal Problems; MOTR, motive-oriented therapeutic relationship; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation; TAU, treatment as usual.

Psychoanalytic-Interactional Therapy vs. Psychodynamic Therapy by experts

Table G—8. Study characteristics and main results of psychoanalytic-interactional therapy compared with psychodynamic therapy by experts.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Leichsenring et al. (2016)	Design: RCT Setting: Inpatient, single center Country: Germany Funding: Other	N=168 G1 (46): WL/TAU: 80% patients continued their usual treatment and the remaining did not receive any treatment during the WL period G2 (64): PIT: 1 or 2 weekly individual sessions and 3 weekly group therapy sessions; art therapy or body therapy and consultations with a social worker (on needs basis) G3 (58): E-PDT: 1 or 2 weekly sessions of non-manualized individual therapy and 3 weekly sessions of group therapy; art therapy or body therapy and consultations with a	Inclusions: Age 18-65 years; cluster B personality disorder diagnosis according to SCID-II (DSM-IV) Exclusions: Psychotic and acute substance-related disorders, acute (uncontrollable) risk of suicide, or organic mental disorders; severe medical conditions (according to ICD-10)	Mean (SD) age: G1: 31 (9.4) G2: 29 (8.7) G3: 30 (9.1) % Female: 69 % Race/ethnicity: NR	Primary outcome: BPI, GSI of the SCL-90-R at end of treatment (duration varies by treatment) G2 and G3 significantly more effective than G1 for improving in BPD outcomes (measured by BPI [G2 vs. G1: 18.76 vs. 26.39, p=0.004; G3 vs. G1: 19.41 vs. 26.39, p=0.0004]), depression (BDI [G2 vs. G1: 17.44 vs. 27.80, p=0.0001; G3 vs. G1: 15.20 vs. 27.80, p=0.0001]) and global functioning (measured by GSI of the SCL-90-R [G2 vs. G1: 0.99 vs. 1.65, p=0.0001; G3 vs. G1: 0.96 vs. 1.65, p=0.0001]) No significant differences between active arms (G2 and G3) and G1 for anxiety (BAI) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 18.0% (22/122) Differential attrition: ≤10 percentage point	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		social worker (on needs basis) G1: 89.69 days, mean (SD=105.31) G2: 106.7 days, mean (SD=41.71) G3: 76.78 days, mean (SD=21.07)				

Abbreviations. AE, adverse event; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; BPD, borderline personality disorder; BPI, Borderline Personality Inventory; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; E-PDT, Psychodynamic Therapy by Experts in Personality Disorders; G1, Group 1; G2, Group 2; G3, Group 3; GSI, Global Severity Index; ICD-10, *International Classification of Diseases, Tenth Revision;* N, sample size; NR, not reported; PIT, psychoanalytic-interactional therapy; RCT, randomized controlled trial; SCL-90-R, Symptom Checklist-90-Revised; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SD, standard deviation; TAU, treatment as usual; WL, wait-list.

Mechanism-Based Group Psychotherapy vs. Nonspecific Supportive Psychotherapy

Table G—9. Study characteristics and main results of mechanism-based group psychotherapy compared with nonspecific supportive psychotherapy.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Herpertz et al. (2020)	Design: RCT Setting: Outpatient, single center Country: Germany Funding: Other, German Research Foundation	N=59 G1 (29): Mechanism-based anti-aggression psychotherapy; a highly manualized program starting with 1 individual 1-hour session followed by 6 weeks of group therapy with 2 1.5-hour sessions per week (a total of 18 hours) G2 (30): Nonspecific supportive psychotherapy similar to DBT with same dosage as G1	Inclusion: Age 18-55 years; outpatients meeting ≥4 BPD criteria according to the International Personality Disorder Examination Exclusion: Additional non-study psychotherapy; pregnancy; epilepsy; bipolar I disorder, schizophrenia, or current substance abuse or addiction as well as change in medication within the last 3 weeks	Mean (SD) age: G1: 33 (8.8) G2: 30 (9.5) % Female: 64 % Race/ethnicity: NR	Primary outcome: MOAS at 6 months No difference between groups at end of treatment G2 significantly greater improvements in overt agression on the MOAS at 6 months (10.60 vs. 22.95, p=0.02) Incidence of AEs: G1: 6.9% (2/29) G2: 0% (0/30) Withdrawal due to AE: G1: 3.4% (1/29) G2: 0% (0/30) Attrition: 24%	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		6 months			Differential attrition: ≥10 percentage points G1: 31% (9/29) G2: 17% (5/30)	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; DBT, dialectical behavior therapy; G1, Group 1; G2, Group 2; MOAS, Modified Overt Aggression Scale; N, sample size; NR, not reported; RCT, randomized controlled trial; SD, standard deviation.

Other Psychotherapy vs. Treatment as Usual

Table G—10. Study characteristics and main results of other psychotherapy compared with treatment as usual.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Amianto et al. (2011)	Design: RCT Setting: Outpatient, single center Country: Italy Funding: Government, other	N=35 G1 (17): TAU: Supervised team management G2 (18): Supervised team management plus sequential brief Adlerian psychodynamic psychotherapy 12 months	Inclusion: Age 20-50 years; DSM-IV-TR BPD diagnosis; heavy use of MHS throughout prior year Exclusion: Acute comorbid Axis I disorder requiring hospitalization; current SUD; intellectual disability; previous psychotherapy interventions	Mean (SD) age: 40 (9.4) % Female: 49 % Race/ethnicity: NR	Primary outcome: High mental health use (more than 6 emergency interventions in the prior year) No significant differences between G2 and G1 in CGI, SCL-90, and GAF at 12 months Attrition: 5.7% (2/35) G1: 5.9% (1/17) G2: 5.6% (1/18)	Moderate
Gratz et al. (2014)	Design: RCT Setting: Outpatient, single center Country: NR Funding: Government, NIMH	N=61 G1 (30): TAU: Ongoing outpatient treatment with most participants (>70%) receiving supportive or dynamic individual therapy; others (19%) receiving CBT G2 (31): ERGT: Weekly 90-minute group sessions over 14 weeks (6 patients/group) 14 weeks	Inclusion: Women; age 18-60 years; threshold or subthreshold diagnosis of BPD; history of repeated deliberate self-harm, with ≥1 episode in the past 6 months; having an individual therapist, psychiatrist, or case manager; diagnostic interview for DSM-IV Exclusion: Diagnoses of a primary psychotic disorder, bipolar I disorder, or current (past month) SUD	Mean (SD) age: G1: 33 (0.9) G2: 33 (11.0) % Female: 100 % Race/ethnicity: Racial/ethnic minority: 21	Primary outcome: NR G2 significantly more effective than G1 for improving self-harm (measured by the SHI [16.05 vs. 29.40, p<0.05]), emotion dysregulation (measured by the DERS [95.27 vs. 113.62, p<0.05]), BPD severity (meaured by the ZAN-BPD [4.35 vs. 12.03, p<0.05]), and quality of life (measured by the QLI [0.31 vs. -0.50, p<0.05]) No significant differences between G2 and G1 for measures of BPD-related severity and symptoms (measured by	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
					the composite of the IIP, BEST, AAQ, BDI, and SDS) Attrition: 13.1% (8/61) G1: 10% (3/30) G2: 16.1% (5/31)	
Reneses et al. (2013)	Design: RCT Setting: Outpatient, multicenter Country: Spain Funding: Government, Ministry	N=53 G1 (28): TAU: Conventional treatment without specific additional psychotherapy for 6 months; psychopharmacological treatment in accordance with the standard applied in the hospital clinic G2 (25): PRFP along with conventional care: 20 face-to-face, 45-minute, consecutive weekly PRFP sessions plus conventional outpatient psychiatric treatment 12 months	Inclusion: Age 18-50 years; clinical diagnosis of BPD using the DSM-IV-TR and the SCID-II; clinical situation of outpatient treatment Exclusion: Active suicide risk symptoms, violent or unmanageable heteroaggressive behaviors; comorbidity with diagnosis of eating behavior disorder on Axis I, with toxic dependence disorder or current severe physical disease; interrupting patients' psychotherapy for >4 consecutive sessions without justification or for >6 sessions in any case	Mean (SD) age: 34 (7.5) % Female: 71 % Race/ethnicity: NR	Primary outcome: Severity of the general symptoms (GSI of SCL-90-R) and impulsivity (BIS, SASS) G2 significantly more effective than G1 for improving BPD severity (measured by ZAN-BPD [13.0 vs. 19.1, p<0.001]) and symptoms (SCL-90 [1.2 vs. 1.7, p<0.001], MADRS total [15.9 vs. 22.8, p<0.001], BIS score [52.5 vs. 68.2, p<0.01], and SASS score [35.4 vs. 27.6, p<0.001]) No significant differences between G2 and G1 for STAI state score or CGI Attrition: 13% (7/53) G1: 14% (4/28) G2: 12% (3/25)	High
Leichsenring et al. (2016)	Design: RCT Setting: Inpatient, single center Country: Germany Funding: Other	N=168 G1 (46): WL/TAU: 80% patients continued their usual treatment and the remaining did not receive any treatment during the WL period G2 (64): PIT: 1 or 2 weekly individual sessions and 3 weekly group therapy sessions; art therapy or body therapy and consultations	Inclusions: Age 18-65 years; cluster B personality disorder diagnosis according to SCID-II (DSM-IV) Exclusions: Psychotic and acute substance-related disorders, acute (uncontrollable) risk of suicide, or organic mental disorders; severe medical conditions (according to ICD-10)	Mean (SD) age: G1: 31 (9.4) G2: 29 (8.7) G3: 30 (9.1) % Female: 69 % Race/ethnicity: NR	Primary outcome: BPI, GSI of the SCL-90-R at end of treatment (duration varies by treatment) G2 and G3 significantly more effective than G1 for improving in BPD outcomes (measured by BPI [G2 vs. G1: 18.76 vs. 26.39, p=0.004; G3 vs. G1: 19.41 vs. 26.39, p=0.0004]), depression (BDI [G2 vs. G1: 17.44 vs. 27.80, p=0.0001; G3 vs. G1: 15.20 vs. 27.80, p=0.0001]) and global functioning (measured by GSI of the	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
		with a social worker (on needs basis) G3 (58): E-PDT: 1 or 2 weekly sessions of non-manualized individual therapy and 3 weekly sessions of group therapy; art therapy or body therapy and consultations with a social worker (on needs basis) G1: 89.69 days, mean (SD=105.31) G2: 106.7 days, mean (SD=41.71) G3: 76.78 days, mean (SD=21.07)			SCL-90-R [G2 vs. G1: 0.99 vs. 1.65, p=0.0001; G3 vs. G1: 0.96 vs. 1.65, p=0.0001]) No significant differences between active arms (G2 and G3) and G1 for anxiety (BAI) Incidence of AEs: NR Withdrawals due to AEs: NR Attrition: 18.0% (22/122) Differential attrition: ≤10 percentage point	

Abbreviations. AAQ, Acceptance and Action Questionnaire; AE, adverse event; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BIS, Barrat Impulsivity Scale; BPD, borderline personality disorder; BPI, Borderline Personality Inventory; CBT, cognitive-behavioral therapy; CGI, Clinical Global Impression; DERS, Difficulties in Emotion Regulation Scale; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*; ERGT, emotion regulation group therapy; E-PDT, psychodynamic therapy by experts in personality disorders; G1, Group 1; G2, Group 2; G3, Group 3; GAF, Global Assessment of Functioning; GSI, Global Severity Index; ICD-10, *International Classification of Diseases, Tenth Revision*; IIP, Inventory of Interpersonal Problems; MADRS, Montgomery-Åsberg Depression Rating Scale; MHS, mental health services; N, sample size; NIMH, National Institute of Mental Health; NR, not reported; PIT, psychoanalytic-interactional therapy; PRFP, psychic representation focused psychotherapy; QLI, Quality of Life Inventory; RCT, randomized controlled trial; SASS, Social Adaptation Self-evaluation Scale; SCID-II, Structured Clinical Interview for DSM-IV Axis II Disorders; SCL-90, Symptom Checklist-90; SCL-90-R, Symptom Checklist-90-Revised; SD, standard deviation; SDS, Sheehan Disability Scale; SHI, Self-Harm Inventory; STAI, State-Trait-Anxiety Inventory; SUD, substance use disorder; TAU, treatment as usual; WL, wait-list; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

Service Delivery Approaches

Table G—11. Study characteristics and main results of service delivery approaches.

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
Bartak et al. (2011)	Design: Prospective cohort study	N=245 G1 (59): Outpatient individual or group psychotherapy sessions for up to 2 sessions/week	Inclusion: Participants with cluster B personality disorders diagnosed with DSM-IV Personality	Based on N analyzed: Mean (SD) age: 31 (8.5)	Primary outcome: GSI at 18 months No significant differences in GSI and EQ-5-D Incidence of AEs: NR	Moderate

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Setting: University hospital and mental health care centers Country: Netherlands Funding: None	G2 (99): At least 1/week of psychotherapy in day-hospital but slept at home G3 (87): Stayed in the institution for 5 days/week and received different forms of psychotherapy 18 months	Exclusion: Organic cerebral impairment, intellectual disability, or schizophrenia	% Female: 71 % Race/ethnicity: NR 77% with BPD	Withdrawal due to AEs: NR Attrition:16% Differential attrition: ≤10 percentage points	
Laporte et al. (2018)	Design: Prospective cohort Setting: Outpatient, multicenter Country: Canada Funding: McGill University	N=681 G1 (479): 12 weekly sessions of individual therapy and 12 of group therapy G2 (138): Extended care clinic with weekly sessions of 2 types of group therapy, weekly sessions of individual therapy, and pharmacological management G1: 12 weeks G2: 6-24 months	Inclusion: DSM-5 BPD diagnosis; ≥8 on the DIB-R for current BPD Exclusion: NR	Mean (SD) age: G1: 27 (7.8) G2: 36 (10.4) % Female: 93% % Race/ethnicity: NR	Primary outcome: NR Significant reductions in both groups but no reporting on between group comparisons Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 32% Differential attrition: G1: 29% (137/479) G2: 43% (59/138)	High
Sinnaeve et al. (2018)	Design: RCT Setting: Community mental health centers Country: Netherlands Funding: GGZ Rivierduinen	N=84 G1 (42): Standard, outpatient DBT G2 (42): Step-down DBT consisting of 3 months of residential DBT plus 6 months of outpatient DBT G1: 12 months G2: 9 months	Inclusion: DSM-IV BPD diagnosis; age 18-45 years; ≥24 on the BPDSI-IV and ≥1 episode of self-injurious behavior Exclusion: Chouronic psychotic disorder, bipolar I disorder, intellectual disability, or SUD requiring detoxification; involuntary psychiatric treatment	Mean (SD) age: G1: 26 (7.5) G2: 26 (6.2) % Female: 95 % Race/ethnicity: NR	Primary outcome: NR Significant reductions in both groups but no reporting on between group comparisons Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 44% Differential attrition: ≤10 percentage points	High
Smits et al. (2020; 2022)	Design: RCT Setting: Outpatient, multicenter Country: Netherlands	N=114 G1 (70): MBT, day-hospital setting	Inclusion: BPD diagnosis; age ≥18 years Exclusion: Autism spectrum disorder, chouronic	Mean (SD) age: G1: 31 (10.6) G2: 30 (9.2) % Female: 83	Primary outcome: GSI of BSI at 18 months Significant improvements on all outcomes (GSI, SSHI, PAI-BOR, EQ-	High

Author (year) Trial name	Study characteristics	N of participants Interventions Duration	Study population including main inclusion and exclusion criteria	Sample demographics	Primary outcome; main results; attrition	Risk of bias
	Funding: ZonMw	G2 (44): MBT, intensive outpatient setting 18 months	psychotic disorder, or organic brain disorder, intellectual disability (IQ <80), or antisocial personality disorder with a history of physical violence	% Race/ethnicity: NR	5D, IIP, SIPP) at 18 months and no significant between group difference except on IIP and SIPP No significant differences between groups at 36 months Incidence of AEs: NR Withdrawal due to AEs: NR Attrition: 78% at 18 months Differential attrition: 18%	

Abbreviations. AE, adverse event; BPD, borderline personality disorder; BPDSI-IV, Borderline Personality Disorder Severity Index-IV; BSI, Brief Symptom Inventory; DBT, dialectical behavior therapy; DIB-R, Diagnostic Interview for Borderlines-Revised; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; DSM-5, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; EQ-5D, European Quality of Life–5 Dimension; G1, Group 1; G2, Group 2; G3, Group 3; GSI, Global Severity Index; IIP, Inventory of Interpersonal Problems; IQ, intelligence quotient; MBT, mentalization-based treatment; N, sample size; NR, not reported; PAI-BOR, Personality Assessment Inventory-Borderline Features Scale; RCT, randomized controlled trial; SD, standard deviation; SIPP, Severity Indices of Personality Problems; SSHI, Suicide and Self-Harm Inventory; SUD, substance use disorder.

Appendix H. Assessments

Table H—1. Summary of Outcome Measures for Borderline Personality Disorder[1].

Measure	Full name	Description	Minimally important difference
ALS	Affective Lability Scale	Items: 54 item self-report measure of lability of anger Scale: 0 to 3 (greater affective lability) Scoring: Patients rate different features of mood instability on a 4-point Likert scale from 0 (very uncharacteristic) to 3 (very characteristic); the total score is the mean of all item responses divided by the number of responses	NR
BIS-11	Barratt Impulsiveness Scale	Items: 30-item self-report questionnaire designed to measure impulsivity, items describe common impulsive or nonimpulsive behaviors and preferences Scale: 30 to 120 (greater impulsivity) Scoring: Each item is rated on a 4-point Likert scale from 1 (rarely/never) to 4 (almost always/always); overall score is calculated from the sum of the 30 items	NR
BAI	Beck Anxiety Inventory	Items: 21-item self-report measure of anxiety items Scale: 0 (low anxiety) to 63 (score of ≥ 36=potentially concerning levels of anxiety) Scoring: Each item is rated on a 4-point Likert scale from 0 (not at all bothered) to 3 (severely bothered); total score is calculated by finding the sum of the 21 items	NR
BDI	Beck Depression Inventory	Items: 21-item self-report inventory that measures characteristic attitudes and symptoms of depression Scale: 0 to 63 (minimal to severe depression) Scoring: Each item is rated on a 4-point Likert scale from 0 (mild) to 3 (severe); total score is calculated by finding the sum of the 21 items	MCID=5
BHS	Beck Hopelessness Scale	Items: 20-item checklist that assesses negative attitudes about the future Scale: 0 to 20 (scores of 9 or more are associated with an 11 times higher suicide rate than score of 8 or below) Scoring: Each item is rated true or false; total score is calculated by finding the sum of endorsed pessimistic statements and denied optimistic statements	NR
BSS	Beck Scale for Suicide Ideation	Items: 21-item self-report instrument evaluating the current intensity of suicidality in the past week Scale: 0 to 38	NR

Measure	Full name	Description	Minimally important difference
		Scoring: Each item consists of 3 options graded according to suicidal intensity on a 3-point scale ranging from 0 to 2; ratings for the first 19 items are summed to yield total score	
BEST	Borderline Evaluation of Severity Over Time	Items: 15-item self-report questionnaire designed to assess change in the severity of BPD during the prior month Scale: 12 (best) to 72 (worst) Scoring: Each item is rated on a 5-point Likert scale from 1 (none/never) to 5 (extreme/almost always); items are divided among 3 subscales (A, B, C); total score is calculated by adding together the scores of subscales A and B then subtracting the total from subscale C and adding a correction factor of 15	NR
BPDSI	Borderline Personality Disorder Severity Index	Items: 70-item semi-structured clinical interview measure assessing frequency and severity of BPD-related symptoms among nine symptom areas corresponding to DSM-IV criteria Scale: 0 to 90 (scores above 15 signify BPD pathology) Scoring: Each item is rated on an 11-point scale from 0 (never) to 10 (daily); for each DSM criterion an average score is derived (range=0-10) with the sum of these 9 scores providing the total score	NR
BSL-23	Borderline Symptom List-23	Items: 23-item self-report scale to assess borderline typical symptomatology Scale: 0 (none or low) to 4 (extremely high) Scoring: Each item is rated on a 5-point Likert scale from 0 (not at all) to 4 (very strong); total score is calculated as the sum of item response ratings divided by the total number of responses	NR
BSI	Brief Symptom Inventory	Items: 53-item self-report scale derived from SCL-90-R to identify clinically relevant psychological symptoms Scale: NR Scoring: Each item is rated on a 5-point Likert scale from 0 (not at all) to 4 (extremely); Global Severity Index (GSI) is calculated by using the sums for the 9 symptom dimensions plus 4 additional items and dividing by the total number of item responses, providing the mean score	NR
CGI-I	Clinical Global Impression-Improvement	Items: 1-item clinician-rated instrument to conduct global assessment of illness improvement Scale:1 to 7 Scoring: A clinician rates patient's mental illness on a scale from 1 (very much improved) to 7 (very much worse)	NR

Measure	Full name	Description	Minimally important difference
CGI-S	Clinical Global Impression-Severity	Items: 1 item clinician-rated instrument to conduct global assessment of illness severity Scale: 0 to 7 Scoring: A clinician rates patient's mental illness on a 7-point scale: 1 (normal, not at all ill), 2 (borderline mentally ill), 3 (mildly ill), 4 (moderately ill), 5 (markedly ill), 6 (severely ill), 7 (among the most extremely ill patients); the score should reflect the average severity level across the past 7 days	NR
CUXOS	Clinically Useful Anxiety Outcome Scale	Items: 20-item self-report measure designed to assess the severity of anxiety symptoms in adults with a diagnosed anxiety disorder or depression Scale: 0 to 80 (<10 nonanxious; 11-20 minimal anxiety; 21-30 mild anxiety; 31-40 moderate anxiety; 41+ severe anxiety) Scoring: There are two subscales: psychic anxiety and somatic anxiety; each item is rated on a 5-point Likert scale from 0 (not at all) to 4 (almost always); total score is the sum of all items	NR
CUDOS	Clinically Useful Depression Outcome Scale	Items: 18-item self-report scale to identify depression symptoms and impact Scale: 0 to 72 (nondepressed 0 to 10; minimal depression, 11 to 20; mild depression, 21 to 30; moderate depression, 31 to 45; and severe depression, 46 and above) Scoring: Each item is rated on a 5-point Likert scale from 0 (not at all) to 4 (almost always); total score is the sum of all items	NR
DSHI	Deliberate Self-Harm Inventory	Items: 17-item self-report measure that assesses the method, frequency and medical severity of deliberate self-harm without suicidal intent Scale: 0 to 17 Scoring: Each item is answered yes or no; total score is sum of "yes" answers	NR
DASS	Depression, Anxiety, Stress Scale	Items: 42-item self-report questionnaire that measures depression, anxiety, and stress Scale: 0 to 126 (suggested cutoffs for normal, mild, moderate, severe, extremely severe for depression are [9, 13, 20, 27, 42], for anxiety [7, 9, 14, 19, 42], and for stress [14, 18, 25, 33, 42]) Scoring: Each item is rated on a 4-point Likert scale ranging from 0 to 3 for how often the item has been experienced in the past week; total score is calculated by summing all items	NR

Measure	Full name	Description	Minimally important difference
DERS	Difficulties in Emotion Regulation Scale	Items: 36-item self-report measure of 6 facets of emotion regulation Scale: 36 to 180 (higher scores indicate greater degree of emotion dysregulation) Scoring: Each item is rated on a 5-point Likert scale from 1 (almost never) to 5 (almost always); total score is calculated by summing all items	NR
DES	Dissociative Experiences Scale	Items: 28-item self-report scale to measure a variety of types of dissociation Scale: 0 to 100 (higher scores indicate greater likelihood of dissociative disorder; suggested cutoff score is 45) Scoring: Each item is rated from 0% of time experiencing the item to 100% time, increasing in 10% increments; the mean score is used as the total	NR
EQ-5D	European Quality of Life–5 Dimension	Items: 5-item instrument to measure health-related quality of life in Europe Scale: 0 (worst) to 100 (best) Scoring: Each item can be rated at one of 3 response levels: “slight problems,” “moderate problems,” “extreme problems”	NR
GAF	Global Assessment of Functioning	Items: 100-item clinician-rated instrument indicating overall psychosocial functioning during a specified period on a continuum from psychological sickness to health Scale: 0 to 100 (severely impaired to extremely high functioning) Scoring: GAF rating can be based on many things, including: an interview or questionnaire, medical records, information from medical providers, caregivers, or relatives, police or court records about violent or illegal behavior; the summary score reflects the level of an individual’s overall functioning	NR
GAS	Global Assessment Scale	Items: 1-item clinician-rated instrument evaluating overall functioning during a specified period on a continuum from psychological sickness to health Scale: 1 (hypothetically sickest) to 100 (hypothetically healthiest); the scale is divided into 10 equal intervals Scoring: In making a rating, the lowest interval that describes the subject’s functioning during the preceding week is selected; information needed to make the rating can come from the patient, a reliable informant, or case record	NR

Measure	Full name	Description	Minimally important difference
Ham-A	Hamilton Rating Scale for Anxiety	Items: 14-item questionnaire used to assess patients' anxiety Scale: 0 to 56 (<17=mild severity, 18-24=mild to moderate severity, >25=severe) Scoring: Each item is rated on a 5-point Likert scale from 0 (not present) to 4 (most severe); the sum of the score indicates the severity of anxiety	NR
Ham-D	Hamilton Rating Scale for Depression	Items: 17 or more item questionnaire used to assess patients' depression Scale: 0 to 53 (0-7 considered normal and >20 considered moderate severity) Scoring: Each item is rated on a 3- or 5-point Likert scale from 0 to 2 or 0 to 4; the sum of the score indicates the severity of depression	NR
IIP	Inventory of Interpersonal Problems	Items: 64-item self-report measure of interpersonal distress Scale: 0 to 64 (higher scores indicate more interpersonal distress) Scoring: Each item is rated on a 5-point Likert scale from 0 (not at all) to 4 (extremely) on how much difficulty/distress it causes participants; items are grouped into 8 subscales	NR
LSASI	Lifetime Suicide Attempt Self-Injury Interview	Items: 20-item clinician-administered structured, face to face interview for assessing information regarding participant's first, most recent, and most severe episodes of self-injury Scoring: Assessors code suicide and self-injury behaviors according to method, lethality, intent to die, and level of medical treatment received	NR
MOAS	Modified Overt Aggression Scale	Items: 20-item clinician-administered, semi-structured interview designed to assess various manifestations of aggressive behavior in outpatients Scale: 0 to 100 (no symptoms to severe) Scoring: 4 subcomponent types of aggression are scored between 0 (no aggression) and 4 with a potential cumulative score of 10 for each subcomponent with each subcomponent is weighted differently; total score is calculated by multiplying sum score of each subcomponent by the weight for that category, then summing the weighted scores	NR
MADRS	Montgomery-Åsberg Depression Rating Scale	Items: 10-item clinician-rated measure of severity of ten depressive symptoms Scale: 0 to 60 (0-6 is defined as symptom absent and >34 is defined as severe depression)	NR

Measure	Full name	Description	Minimally important difference
		Scoring: Each item is rated on a scale from 0 to 6, with 6 as the most severe description of the symptom; total score is the sum of scores for each item	
OAS-M	Overt Aggression Scale—Modified	Items: 20-item clinician-administered, semi-structured interview designed to assess various manifestations of aggressive behavior in outpatients Scoring: Manifestations of aggression from the preceding week are scored between 0 (no events within that category) and 5 (most severe form of aggression within that category), frequency of events is then multiplied by a weighted severity level for that category (0 to 5) to produce a raw score for each subscale; each subscale is also weighted (1 to 3x) and total score is calculated by summing weighted scores from each subscale	NR
PAI	Personality Assessment Inventory	Items: 344-item self-report instrument of 22 nonoverlapping scales to assess personality and psychopathology Scoring: Each item is rated from 0 (false) to 4 (very true) on a 4-point Likert scale	NR
QoL	Quality of Life Index	Items: 10-item self-report instrument measuring 10 dimensions of health-related quality of life Scale: 0 to 100 Scoring: Each item is rated from 1 (poor) to 10 (excellent), total score is summed total from each item	NR
SHI	Self-Harm Inventory	Items: 22-item self-report instrument that explores respondents' histories of self-harm Scale: 0 to 22 Scoring: Each item is answered yes or no, total score is summed by counting the number of endorsed self-harm behaviors	NR
SDS	Sheehan Disability Scale	Items: 5-item self-rated instrument used to measure the effect of the individual's symptoms on three areas Scale: 0 to 30 (no symptoms to severe) Scoring: Each of 3 areas is scored according to how much it was disrupted by symptoms (0=not at all to 10=very severely)	NR
SFQ	Social Functioning Questionnaire	Items: 8-item self-report scale to assess perceived social function Scale: 0 to 24 (score of >10 indicates poor social functioning)	NR

Measure	Full name	Description	Minimally important difference
		Scoring: Each item is scored on a 4-point scale from 0 (no/never) to 3 (severely/always), total score is the sum of all items	
STAXI	State-Trait Anger Expression Inventory	Items: 69-item self-report questionnaire that focuses on anger expression Scoring: Each item is rated on a 4-point scale for frequency of exhibiting behavior (almost always, often, sometimes, almost never)	NR
STAXI-II	State-Trait Anger Expression Inventory-II	Items: 57-item self-report questionnaire that focuses on anger expression; updated version of STAXI Scoring: Each item is rated on a 4-point scale for frequency of exhibiting behavior (almost always, often, sometimes, almost never)	NR
SBQ	Suicidal Behaviors Questionnaire	Items: 4-item self-reported measure of suicidal thoughts and behaviors Scale: 5 to 19 Scoring: Each item is rated on a Likert scale from 1 to 3, 5, or 6; total score is the sum of all items	NR
SASII	Suicide Attempt Self-Injury Interview	Items: 40-item semi-structured interview measures frequency, intent, and medical severity of suicide attempts and NSSI acts Scale: Nonsuicidal self-injury, ambivalent suicide attempt, nonambivalent suicide attempt, failed suicide Scoring: Assessors use 6 screening items, 9 open-ended questions, and scores from 6 scales to categorize episodes	NR
SRS	Suicide Risk Scale	Items: 26-item scale to measure risk of suicide Scale: 0 to 26 Scoring: Each item is answered yes or no; the number of positive responses can be summed for a total score	NR
SCL-90-R	Symptom Checklist-90-Revised	Items: 90-item self-report screening measure of general psychiatric symptomatology along nine symptom constructs Scale: 0 to 4 Scoring: Each item is scored on a 5-point Likert scale from 0 (not at all bothered) to 4 (extremely bothered); Global Severity Index (GSI) can be calculated as the average score of the 90 items in the questionnaire	NR
WHOQOL	World Health Organization	Items: 100-item self-report questionnaire assessing quality of life through 6 domains	NR

Measure	Full name	Description	Minimally important difference
	Quality of Life Scale	Scale: 0 to 100 (higher scores denote higher quality of life) Scoring: Each item is rated on a 5-point Likert scale from 1 (not at all) to 5 (extremely); the scale has 24 facets divided inequally among 6 domains, each domain has a unique method of calculating mean score; domain and facet scores can be transformed to a 100-point scale using this formula: TRANSFORMED SCORE=(SCORE-4)x(100/16)	
ZAN–BPD	Zanarini Rating Scale for Borderline Personality Disorder	Items: 9-item semi structured interview Scale: 0 to 36 (no symptoms to severe) Scoring: Each item is rated on a 5-point Likert scale from 0 (no symptoms) to 4 (severe symptoms) on each of nine items corresponding to the nine DSM-IV criteria for BPD, total score is the sum of all items	NR

Note. [1] Additional rating scales that can be used in adolescents include the Beck Depression Inventory for Youth, the Borderline Personality Features Scale for Children, the Children's Global Assessment of Functioning Scale, the Millon Adolescent Clinical Inventory, the Youth Quality of Life Research Version, and the Youth Self-Report Scale (Jørgensen et al. 2021).

Abbreviations. BPD, BPD, borderline personality disorder; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; GAF, Global Assessment of Functioning; GSI, Global Severity Index; MCID, minimal clinically important difference; NR, not reported; SCL-90, Symptom Checklist-90; STAXI, State-Trait Anger Expression Inventory.

Appendix I. Excluded Studies

List of Exclusion Codes:

X1: Ineligible population

X2: Ineligible intervention

X3: Ineligible comparator

X4: Ineligible outcome

X5: Ineligible timing

X6: Ineligible study design

X7: Duplicate or superseded by paper publication

X8: Non-English full text

X9: Ineligible country

X10: Not primary research

1. Antonsen BT, Kvarstein EH, Urnes Ø, et al. Favourable outcome of long-term combined psychotherapy for patients with borderline personality disorder: six-year follow-up of a randomized study. Psychother Res. 2017 Jan;27(1):51-63. doi: 10.1080/10503307.2015.1072283. PMID: 26261865. Exclusion Code: X1.

2. Arnevik E, Wilberg T, Urnes O, et al. Psychotherapy for personality disorders: short-term day hospital psychotherapy versus outpatient individual therapy - a randomized controlled study. Eur Psychiatry. 2009 Mar;24(2):71-8. doi: 10.1016/j.eurpsy.2008.09.004. PMID: 19097870. Exclusion Code: X1.

3. Artusio E. Acceptance and commitment therapy for individuals with emotion regulation problems: A pilot study: ProQuest Information and Learning; 2021. Exclusion Code: X6.

4. Bales DL, Verheul R, Hutsebaut J. Barriers and facilitators to the implementation of mentalization-based treatment (MBT) for borderline personality disorder. Personality and Mental Health. 2017;11(2):118-31. doi: 10.1002/pmh.1368. Exclusion Code: X4.

5. Barnicot K, Crawford M. Conclusions and questions from a non-randomised comparison of routine clinical services implementing different treatment models for borderline personality disorder. Psychol Med. 2019 Dec;49(16):2812-4. doi: 10.1017/s0033291719002447. PMID: 31551098. Exclusion Code: X6.

6. Bateman A, Fonagy P. Impact of clinical severity on outcomes of mentalisation-based treatment for borderline personality disorder. Br J Psychiatry. 2013 Sep;203(3):221-7. doi: 10.1192/bjp.bp.112.121129. PMID: 23887998. Exclusion Code: X6.

7. Bateman A, O'Connell J, Lorenzini N, et al. A randomised controlled trial of mentalization-based treatment versus structured clinical management for patients with comorbid borderline personality disorder and antisocial personality disorder. BMC Psychiatry. 2016 Aug 30;16(1):304. doi: 10.1186/s12888-016-1000-9. PMID: 27577562. Exclusion Code: X6.

8. Beck E, Bo S, Jørgensen MS, et al. Mentalization-based treatment in groups for adolescents with borderline personality disorder: a randomized controlled trial. J Child Psych Psychiatry. 2020 May 61(5):594-604. doi: 10.1111/jcpp.13152. PMID: 2313759274. Exclusion Code: X7.

9. Bellino S, Bozzatello P, Rocca G, et al. Efficacy of omega-3 fatty acids in the treatment of borderline personality disorder: a study of the association with valproic acid. J Psychopharmacol. 2014 Feb;28(2):125-32. doi: 10.1177/0269881113510072. PMID: 24196948. Exclusion Code: X2.

10. Bernard L, Walburg V. Efficacy of a brief cognitive-emotional group intervention for patients with borderline personality disorder. Psychologie Française. 2020;65(3):185-96. doi: 10.1016/j.psfr.2019.11.001. PMID: 2021-11287-002. Exclusion Code: X6.

11. Berthoud L, Pascual-Leone A, Caspar F, et al. Leaving distress behind: a randomized controlled study on change in emotional processing in borderline personality disorder. Psychiatry. 2017 Summer;80(2):139-54. doi: 10.1080/00332747.2016.1220230. PMID: 28767333. Exclusion Code: X6.

12. Boritz T, Barnhart R, McMain SF. The influence of posttraumatic stress disorder on treatment outcomes of patients with borderline personality disorder. J Pers Disord. 2016 Jun;30(3):395-407. doi: 10.1521/pedi_2015_29_207. PMID: 26305394. Exclusion Code: X6.

13. Bozzatello P, Rocca P, Bellino S. Combination of omega-3 fatty acids and valproic acid in treatment of borderline personality disorder: a follow-up study. Clin Drug Investig. 2018 Apr;38(4):367-72. doi: 10.1007/s40261-017-0617-x. PMID: 29302857. Exclusion Code: X2.

14. Buchheim A, Hörz-Sagstetter S, Doering S, et al. Change of unresolved attachment in borderline personality disorder: RCT study of transference-focused psychotherapy. Psychother Psychosom. 2017;86(5):314-6. doi: 10.1159/000460257. Exclusion Code: X6.

15. Chapman AL, Rosenthal MZ, Dixon-Gordon KL, et al. borderline personality disorder and the effects of instructed emotional avoidance or acceptance in daily life. J Pers Disord. 2017;31(4):483-502. doi: 10.1521/pedi_2016_30_264. Exclusion Code: X6.

16. Coyle TN, Shaver JA, Linehan MM. On the potential for iatrogenic effects of psychiatric crisis services: the example of dialectical behavior therapy for adult women with borderline personality disorder. J Consult Clin Psychol. 2018 Feb;86(2):116-24. doi: 10.1037/ccp0000275. PMID: 29369662. Exclusion Code: X6.

17. Davidson KM, Tyrer P, Norrie J, et al. Cognitive therapy v. usual treatment for borderline personality disorder: prospective 6-year follow-up. Br J Psychiatry. 2010 Dec;197(6):456-62. doi: 10.1192/bjp.bp.109.074286. PMID: 21119151. Exclusion Code: X6.

18. Edel MA, Raaff V, Dimaggio G, et al. Exploring the effectiveness of combined mentalization-based group therapy and dialectical behaviour therapy for inpatients with borderline personality disorder - a pilot study. Br J Clin Psychol. 2017 Mar;56(1):1-15. doi: 10.1111/bjc.12123. PMID: 27897326. Exclusion Code: X2.

19. Einy S, Narimani M, Atadokht A, et al. Effectiveness of mentalization based therapy and cognitive-analytical therapy on improved object relationship of people with borderline personality disorder: a comparison. Payesh Health Monitor. 2018;17(3):98-110. PMID: CN-02114870. Exclusion Code: X8.

20. Euler S, Stalujanis E, Allenbach G, et al. Dialectical behavior therapy skills training affects defense mechanisms in borderline personality disorder: an integrative approach of mechanisms in psychotherapy. Psychother Res. 2019 Nov;29(8):1074-85. doi: 10.1080/10503307.2018.1497214. PMID: 30005584. Exclusion Code: X4.

21. Farrell JM, Shaw IA, Webber MA. Corrigendum to "A schema-focused approach to group psychotherapy for outpatients with borderline personality disorder: a randomized controlled trial" [(J. Behav. Ther. Exp. Psychiatr.) 40(2) (June 2009) 317-328]. J Behav Ther Exp Psychiatry. 2018 Sep;60:111. doi: 10.1016/j.jbtep.2018.04.001. PMID: 29680679. Exclusion Code: X7.

22. Flynn D, Kells M, Joyce M, et al. Dialectical behaviour therapy for treating adults and adolescents with emotional and behavioural dysregulation: study protocol of a coordinated implementation in a publicly funded health service. BMC Psychiatry. 2018 Feb 26;18(1):51. doi: 10.1186/s12888-018-1627-9. PMID: 29482538. Exclusion Code: X10.

23. Gleeson JF, Chanen A, Cotton SM, et al. Treating co-occurring first-episode psychosis and borderline personality: a pilot randomized controlled trial. Early Interv Psychiatry. 2012 Feb;6(1):21-9. doi: 10.1111/j.1751-7893.2011.00306.x. PMID: 22379625. Exclusion Code: X2.

24. Gray AS, Townsend ML, Bourke ME, et al. Effectiveness of a brief parenting intervention for people with borderline personality disorder: A 12-month follow-up study of clinician implementation in practice. Advances in Mental Health. 2018 Apr doi: 10.1080/18387357.2018.1464887. PMID: 2032685803. Exclusion Code: X1.

25. Griffiths H, Duffy F, Duffy L, et al. Efficacy of mentalization-based group therapy for adolescents: the results of a pilot randomised controlled trial. BMC Psychiatry. 2019;19(1) doi: 10.1186/s12888-019-2158-8. Exclusion Code: X1.

26. Harned MS, Gallop RJ, Valenstein-Mah HR. What changes when? The course of improvement during a stage-based treatment for suicidal and self-injuring women with borderline personality disorder

and PTSD. Psychother Res. 2018 Sep; 28(5):761-75. doi: 10.1080/10503307.2016.1252865. PMID: 27808001. Exclusion Code: X6.

27. Heerebrand SL, Bray J, Ulbrich C, et al. Effectiveness of dialectical behavior therapy skills training group for adults with borderline personality disorder. J Clin Psychol. 2021 Jul;77(7):1573-90. doi: 10.1002/jclp.23134. PMID: 33821506. Exclusion Code: X6.

28. Hollander E, Swann AC, Coccaro EF, et al. Impact of trait impulsivity and state aggression on divalproex versus placebo response in borderline personality disorder. Am J Psychiatry. 2005 Mar;162(3):621-4. doi: 10.1176/appi.ajp.162.3.621. PMID: 15741486. Exclusion Code: X6.

29. Jariani M, Saaki M, Nazari H, et al. The effect of olanzapine and sertraline on personality disorder in patients with methadone maintenance therapy. Psychiatr Danub. 2010 Dec;22(4):544-7. PMID: 21169896. Exclusion Code: X9.

30. Jørgensen CR, Bøye R, Andersen D, et al. Eighteen months post-treatment naturalistic follow-up study of mentalization-based therapy and supportive group treatment of borderline personality disorder: clinical outcomes and functioning. Nord Psychol. 2014;66(4):254-73. Exclusion Code: X6.

31. Juul S, Simonsen S, Poulsen S, et al. Detailed statistical analysis plan for the short-term versus long-term mentalisation-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder randomised clinical trial (MBT-RCT). Trials. 2021 Jul 28;22(1):497. doi: 10.1186/s13063-021-05450-y. PMID: 34321051. Exclusion Code: X6.

32. Keefe JR, Kim TT, DeRubeis RJ, et al. Treatment selection in borderline personality disorder between dialectical behavior therapy and psychodynamic psychiatric management. Psychol Med. 2021 Aug;51(11):1829-37. doi: 10.1017/s0033291720000550. PMID: 32204742. Exclusion Code: X6.

33. Keller S, Stelmaszczyk K, Kolly S, et al. Change in biased thinking in a treatment based on the motive-oriented therapeutic relationship for borderline personality disorder. J Pers Disord. 2018 Jan;32(Suppl):75-92. doi: 10.1521/pedi.2018.32.supp.75. PMID: 29388899. Exclusion Code: X6.

34. Kellett S, Gausden J, Gaskell C. The effectiveness of cognitive analytic therapy for borderline personality disorder: utilizing a withdrawal experimental design to improve sensitivity to abandonment. Psychol Psychother. 2021 Mar;94 Suppl 1:96-119. doi: 10.1111/papt.12278. PMID: 32396677. Exclusion Code: X6.

35. Keng SL, Lee CSL, Eisenlohr-Moul TA. Effects of brief daily mindfulness practice on affective outcomes and correlates in a high BPD trait sample. Psychiatry Res. 2019 Oct;280:112485. doi: 10.1016/j.psychres.2019.112485. PMID: 31408773. Exclusion Code: X6.

36. Keng SL, Tan HH. Effects of brief mindfulness and loving-kindness meditation inductions on emotional and behavioral responses to social rejection among individuals with high borderline personality traits. Behav Res Ther. 2018 Jan;100:44-53. doi: 10.1016/j.brat.2017.11.005. PMID: 29179024. Exclusion Code: X10.

37. Keng SL, Tan JX. Effects of brief mindful breathing and loving-kindness meditation on shame and social problem solving abilities among individuals with high borderline personality traits. Behav Res Ther. 2017;97:43-51. doi: 10.1016/j.brat.2017.07.004. Exclusion Code: X1.

38. Khalid-Khan S, Segal SC, Jopling EN, et al. Effectiveness of a modified dialectical behaviour therapy for adolescents within a stepped-care model. Int J Adolesc Med Health. 2018 Apr;30(2):1-7. PMID: 2207631172. Exclusion Code: X6.

39. Klein JP, Hauer-von Mauschwitz A, Berger T, et al. Effectiveness and safety of the adjunctive use of an internet-based self-management intervention for borderline personality disorder in addition to care as usual: results from a randomised controlled trial. BMJ Open. 2021 Sep 8;11(9):e047771. doi: 10.1136/bmjopen-2020-047771. PMID: 34497078. Exclusion Code: X2.

40. Knoblich N, Gundel F, Bruckmann C, et al. DNA methylation of APBA3 and MCF2 in borderline personality disorder: potential biomarkers for response to psychotherapy. Eur Neuropsychopharmacol. 2018 Feb;28(2):252-63. doi: 10.1016/j.euroneuro.2017.12.010. PMID: 29274998. Exclusion Code: X4.

41. Kramer U, Pascual-Leone A, Berthoud L, et al. Assertive anger mediates effects of dialectical behaviour-informed skills training for borderline personality disorder: a randomized controlled trial. Clin Psychol Psychother. 2016 May;23(3):189-202. doi: 10.1002/cpp.1956. PMID: 25864773. Exclusion Code: X4.

42. Krantz LH, McMain S, Kuo JR. The unique contribution of acceptance without judgment in predicting nonsuicidal self-injury after 20-weeks of dialectical behaviour therapy group skills training. Behav Res Ther. 2018 May;104:44-50. doi: 10.1016/j.brat.2018.02.006. PMID: 29529508. Exclusion Code: X6.

43. Krause-Utz A, Walther JC, Schweizer S, et al. Effectiveness of an emotional working memory training in borderline personality disorder: a proof-of-principle study. Psychother Psychosom. 2020;89(2):122-4. doi: 10.1159/000504454. PMID: 31901902. Exclusion Code: X2.

44. Krawitz R, Miga EM. Cost-effectiveness of dialectical behaviour therapy for borderline personality disorder. In: Swales MA, ed. The Oxford handbook of dialectical behaviour therapy. New York, NY, USA: Oxford University Press; 2019:497-513. Exclusion Code: X6.

45. Kvarstein EH, Pedersen G, Folmo E, et al. Mentalization-based treatment or psychodynamic treatment programmes for patients with borderline personality disorder – the impact of clinical severity. Psychology and Psychotherapy: Theory, Research and Practice. 2019 Mar ;92(1):91-111. doi: 10.1111/papt.12179. PMID: 2019666106. Exclusion Code: X7.

46. Lin TJ, Ko HC, Wu JY, et al. The effectiveness of dialectical behavior therapy skills training group vs. cognitive therapy group on reducing depression and suicide attempts for borderline personality disorder in Taiwan. Arch Suicide Res. 2019 Jan-Mar;23(1):82-99. doi: 10.1080/13811118.2018.1436104. PMID: 29528807. Exclusion Code: X1.

47. Linehan MM, Schmidt H, 3rd, Dimeff LA, et al. Dialectical behavior therapy for patients with borderline personality disorder and drug-dependence. Am J Addict. 1999 Fall;8(4):279-92. doi: 10.1080/105504999305686. PMID: 10598211. Exclusion Code: X1.

48. Marziali E, Munroe-Blum H, McCleary L. The contribution of group cohesion and group alliance to the outcome of group psychotherapy. Int J Group Psychother. 1997 Oct;47(4):475-97. doi: 10.1080/00207284.1997.11490846. PMID: 9314699. Exclusion Code: X6.

49. McMain SF, Fitzpatrick S, Boritz T, et al. Outcome trajectories and prognostic factors for suicide and self-harm behaviors in patients with borderline personality disorder following one year of outpatient psychotherapy. J Pers Disord. 2018 Aug;32(4):497-512. doi: 10.1521/pedi_2017_31_309. PMID: 28910214. Exclusion Code: X6.

50. Mehlum L, Tørmoen AJ, Ramberg M, et al. Dialectical behavior therapy for adolescents with repeated suicidal and self-harming behavior: a randomized trial. J Am Acad Child Adolesc Psychiatry. 2014 Oct;53(10):1082-91. doi: 10.1016/j.jaac.2014.07.003. PMID: 25245352. Exclusion Code: X1.

51. Mendo-Cullell M, Arenas-Pijoan L, Forné C, et al. A pilot study of the efficacy of an adventure therapy programme on borderline personality disorder: a pragmatic controlled clinical trial. Personal Ment Health. 2021 Aug;15(3):159-172. doi: 10.1002/pmh.1505. PMID: 33569869. Exclusion Code: X2.

52. Molavi P, Aziziaram S, Basharpoor S, et al. Repeated transcranial direct current stimulation of dorsolateral-prefrontal cortex improves executive functions, cognitive reappraisal emotion regulation, and control over emotional processing in borderline personality disorder: a randomized, sham-controlled, parallel-group study. J Affect Disord. 2020 May 20;274:93-102. doi: 10.1016/j.jad.2020.05.007. PMID: 32469838. Exclusion Code: X5.

53. Morey LC, Lowmaster SE, Hopwood CJ. A pilot study of manual-assisted cognitive therapy with a therapeutic assessment augmentation for borderline personality disorder. Psychiatry Res. 2010 Aug 15;178(3):531-5. doi: 10.1016/j.psychres.2010.04.055. PMID: 20537722. Exclusion Code: X2.

54. Munroe-Blum, H., and Marziali, E. A controlled trial of short-term group treatment for borderline personality disorder. J Pers Disord. 1995 9(3):190-198. Exclusion Code: X1.

55. Neacsiu AD, Lungu A, Harned MS, et al. Impact of dialectical behavior therapy versus community treatment by experts on emotional experience, expression, and acceptance in borderline personality disorder. Behav Res Ther. 2014 Feb;53:47-54. doi: 10.1016/j.brat.2013.12.004. PMID: 24418652. Exclusion Code: X4.

56. Neacsiu AD, Rizvi SL, Linehan MM. Dialectical behavior therapy skills use as a mediator and outcome of treatment for borderline personality disorder. Behav Res Ther. 2010 Sep;48(9):832-9. doi: 10.1016/j.brat.2010.05.017. PMID: 20579633. Exclusion Code: X4.

57. Norrie J, Davidson K, Tata P, et al. Influence of therapist competence and quantity of cognitive behavioural therapy on suicidal behaviour and inpatient hospitalisation in a randomised controlled trial

in borderline personality disorder: further analyses of treatment effects in the BOSCOT study. Psychol Psychother. 2013 Sep;86(3):280-93. doi: 10.1111/papt.12004. PMID: 23420622. Exclusion Code: X6.

58. O'Dwyer N, Rickwood D, Buckmaster D, et al. Therapeutic interventions in Australian primary care, youth mental health settings for young people with borderline personality disorder or borderline traits. Borderline Personal Disord Emot Dysregul. 2020;7:23. doi: 10.1186/s40479-020-00138-2. PMID: 33042549. Exclusion Code: X6.

59. Pahwa M, Nuñez NA, Joseph B, et al. Efficacy and tolerability of lamotrigine in borderline personality disorder: a systematic review and meta-analysis. Psychopharmacol Bull. 2020 Sep 14;50(4):118-36. PMID: 33012875. Exclusion Code: X6.

60. Paret C, Niedtfeld I, Lotter T, et al. Single-dose effects of Citalopram on neural responses to affective stimuli in borderline personality disorder: a randomized clinical trial. Biol Psychiatry Cogn Neurosci Neuroimaging. 2021 Aug;6(8):837-845. doi: 10.1016/j.bpsc.2021.02.002. PMID: 33607327. Exclusion Code: X5.

61. Priebe S, Bhatti N, Barnicot K, et al. Effectiveness and cost-effectiveness of dialectical behaviour therapy for self-harming patients with personality disorder: a pragmatic randomised controlled trial. Psychother Psychosom. 2012;81(6):356-65. doi: 10.1159/000338897. PMID: 22964561. Exclusion Code: X1.

62. Ramseyer F, Ebert A, Roser P, et al. Exploring nonverbal synchrony in borderline personality disorder: a double-blind placebo-controlled study using oxytocin. Br J Clin Psychol. 2020 Jun;59(2):186-207. doi: 10.1111/bjc.12240. PMID: 31774581. Exclusion Code: X1.

63. Rathus JH, Miller AL. Dialectical behavior therapy adapted for suicidal adolescents. Suicide Life Threat Behav. 2002 Summer;32(2):146-57. doi: 10.1521/suli.32.2.146.24399. PMID: 12079031. Exclusion Code: X1.

64. Reyes-López J, Ricardo-Garcell J, Armas-Castañeda G, et al. Clinical improvement in patients with borderline personality disorder after treatment with repetitive transcranial magnetic stimulation: preliminary results. Braz J Psychiatry. 2018 Jan-Mar;40(1):97-104. doi: 10.1590/1516-4446-2016-2112. PMID: 28614492. Exclusion Code: X9.

65. Reyes-Ortega MA, Miranda EM, Fresán A, et al. Clinical efficacy of a combined acceptance and commitment therapy, dialectical behavioural therapy, and functional analytic psychotherapy intervention in patients with borderline personality disorder. Psychol Psychother. 2020 Sep;93(3):474-89. doi: 10.1111/papt.12240. PMID: 31246370. Exclusion Code: X9.

66. Ridolfi ME, Rossi R, Occhialini G, et al. A clinical trial of a psychoeducation group intervention for patients with borderline personality disorder. J Clin Psychiatry. 2019 Dec 31;81(1). doi: 10.4088/JCP.19m12753. PMID: 31917907. Exclusion Code: X5.

67. Ridolfi ME, Rossi R, Occhialini G, et al. A randomized controlled study of a psychoeducation group intervention for patients with borderline personality disorder. J Clin Psychiatry. 2020;81(1). doi: 10.4088/JCP.19m12753. Exclusion Code: X5.

68. Rinne T, van den Brink W, Wouters L, et al. SSRI treatment of borderline personality disorder: a randomized, placebo-controlled clinical trial for female patients with borderline personality disorder. Am J Psychiatry. 2002 Dec;159(12):2048-54. doi: 10.1176/appi.ajp.159.12.2048. PMID: 12450955. Exclusion Code: X2.

69. Rohde C, Polcwiartek C, Correll CU, et al. Real-world effectiveness of clozapine for borderline personality disorder: results from a 2-year mirror-image study. J Pers Disord. 2018 Dec;32(6):823-37. doi: 10.1521/pedi_2017_31_328. PMID: 29120277. Exclusion Code: X6.

70. Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. J Am Acad Child Adolesc Psychiatry. 2012 Dec;51(12):1304-13.e3. doi: 10.1016/j.jaac.2012.09.018. PMID: 23200287. Exclusion Code: X1.

71. Ryberg W, Diep LM, Landrø NI, et al. Effects of the Collaborative Assessment and Management of Suicidality (CAMS) model: a secondary analysis of moderation and influencing factors. Arch Suicide Res. 2020 24(4):589-608. doi: 10.1080/13811118.2019.1650143. PMID: 2288878579. Exclusion Code: X10.

72. Sahin Z, Vinnars B, Gorman BS, et al. Clinical severity as a moderator of outcome in psychodynamic and dialectical behavior therapies for borderline personality disorder. Personal Disord. 2018 Sep;9(5):437-46. doi: 10.1037/per0000276. PMID: 29239627. Exclusion Code: X6.

73. Salamin V, Kratzel A, Gothuey I, et al. Compliant patients with borderline personality disorder non-responsive to one-year dialectical behavior therapy: outcomes of a second year. J Behav Cogn Ther. 2021;31(2):115-24. doi: 10.1016/j.jbct.2021.01.001. PMID: 2021-54842-005. Exclusion Code: X3.

74. Santamarina Perez P, Romero Cela S, Mendez Blanco I, et al. Efficacy of dialectical behavior therapy compared to supportive therapy in adolescents with suicidal behavior. Eur Neuropsychopharmacol. 2017;27:S853-S4. PMID: CN-01428527. Exclusion Code: X1.

75. Schilling L, Moritz S, Kriston L, et al. Efficacy of metacognitive training for patients with borderline personality disorder: preliminary results. Psychiatry Res. 2018 Apr;262:459-64. doi: 10.1016/j.psychres.2017.09.024. PMID: 28927866. Exclusion Code: X2.

76. Shafti SS, Shahveisi B. Olanzapine versus haloperidol in the management of borderline personality disorder: a randomized double-blind trial. J Clin Psychopharmacol. 2010 Feb;30(1):44-7. doi: 10.1097/JCP.0b013e3181c826ff. PMID: 20075647. Exclusion Code: X9.

77. Signer S, Estermann Jansen R, Sachse R, et al. Social interaction patterns, therapist responsiveness, and outcome in treatments for borderline personality disorder. Psychol Psychother. 2019 Oct 4:e12254. doi: 10.1111/papt.12254. PMID: 31583805. Exclusion Code: X6.

78. Sinyor M, Williams M, Mitchell R, et al. Cognitive behavioral therapy for suicide prevention in youth admitted to hospital following an episode of self-harm: a pilot randomized controlled trial. J Affect Disord. 2020;266:686-94. doi: 10.1016/j.jad.2020.01.178. Exclusion Code: X1.

79. Smith EG. An opportunity to report closer-to-efficacy findings in a study of lamotrigine for borderline personality disorder. Am J Psychiatry. 2018 Dec 1;175(12):1265. doi: 10.1176/appi.ajp.2018.18080991. PMID: 30501419. Exclusion Code: X6.

80. Snoek A, Beekman ATF, Dekker J, et al. A randomized controlled trial comparing the clinical efficacy and cost-effectiveness of eye movement desensitization and reprocessing (EMDR) and integrated EMDR-Dialectical Behavioural Therapy (DBT) in the treatment of patients with post-traumatic stress disorder and comorbid (sub)clinical borderline personality disorder: study design. BMC Psychiatry. 2020 Aug 6;20(1):396. doi: 10.1186/s12888-020-02713-x. PMID: 32762677. Exclusion Code: X6.

81. Soler J, Elices M, Pascual JC, et al. Effects of mindfulness training on different components of impulsivity in borderline personality disorder: results from a pilot randomized study. Borderline Personal Disord Emot Dysregul. 2016;3:1. doi: 10.1186/s40479-015-0035-8. PMID: 26759718. Exclusion Code: X6.

82. St-Amour S, Cailhol L, Ruocco AC, et al. Could physical exercise be an effective treatment for adults with borderline personality disorder? Psychiatry Res. 2021 Jan;295:113625. doi: 10.1016/j.psychres.2020.113625. PMID: 33302133. Exclusion Code: X2.

83. Thompson-Brenner H, Shingleton RM, Thompson DR, et al. Focused vs. broad enhanced cognitive behavioral therapy for bulimia nervosa with comorbid borderline personality: a randomized controlled trial. Int J Eat Disord. 2016;49(1):36-49. doi: 10.1002/eat.22468. PMID: CN-02109829. Exclusion Code: X1.

84. Tong K, Costello S, McCabe E, et al. Mentalization-based treatment in a naturalistic setting in Ireland: a cohort study. Psychiatr Serv. 2021 Jun 9:appips202000810. doi: 10.1176/appi.ps.202000810. PMID: 34106745. Exclusion Code: X4.

85. Turner RM. Naturalistic evaluation of dialectical behavior therapy-oriented treatment for borderline personality disorder. Cogn Behav Pract. 2000;7(4):413-9. Exclusion Code: X2.

86. Vita A, Deste G, Barlati S, et al. Feasibility and effectiveness of cognitive remediation in the treatment of borderline personality disorder. Neuropsychol Rehabil. 2018 Apr;28(3):416-28. doi: 10.1080/09602011.2016.1148054. PMID: 26872501. Exclusion Code: X2.

87. Waltz J, Dimeff LA, Koerner K, et al. Feasibility of using video to teach a dialectical behavior therapy skill to clients with borderline personality disorder. Cogn Behav Pract. 2009;16(2):214-22. doi: 10.1016/j.cbpra.2008.08.004. PMID: CN-01742938. Exclusion Code: X4.

88. Wolkenstein L, Rombold-Bruehl F, Bingmann T, et al. Challenging control over emotions in borderline personality disorder—a tDCS study. Neuropsychologia. 2021;156. doi: 10.1016/j.neuropsychologia.2021.107850. PMID: 2021-43448-001. Exclusion Code: X4.

89. Wong MT, Tye C. Low hospital inpatient readmission rate in patients with borderline personality disorder: a naturalistic study at Southern Health, Victoria, Australia. Aust N Z J Psychiatry. 2005 Jul;39(7):607-11. doi: 10.1080/j.1440-1614.2005.01633.x. PMID: 15996142. Exclusion Code: X4.

90. Zanarini MC, Schulz SC, Detke H, et al. Open-label treatment with olanzapine for patients with borderline personality disorder. J Clin Psychopharmacol. 2012 Jun;32(3):398-402. doi: 10.1097/JCP.0b013e3182524293. PMID: 22544004. Exclusion Code: X3.

91. Zeifman RJ, Boritz T, Barnhart R, et al. The independent roles of mindfulness and distress tolerance in treatment outcomes in dialectical behavior therapy skills training. Personal Disord. 2020 May;11(3):181-90. doi: 10.1037/per0000368. PMID: 31647267. Exclusion Code: X6.

92. Ziegenhorn AA, Roepke S, Schommer NC, et al. Clonidine improves hyperarousal in borderline personality disorder with or without comorbid posttraumatic stress disorder: a randomized, double-blind, placebo-controlled trial. J Clin Psychopharmacol. 2009 Apr;29(2):170-3. doi: 10.1097/JCP.0b013e31819a4bae. PMID: 19512980. Exclusion Code: X2.