

DET SAMFUNDSVIDENSKABELIGE FAKULTET
KØBENHAVNS UNIVERSITET



Duration of psychotherapy for borderline personality disorder

Empirical, clinical, and methodological advancements

Doctoral dissertation in Psychology
Sophie Merrild Juul, PhD

Author

Sophie Merrild Juul, PhD

Departments

¹ Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital, Rigshospitalet

² Department of Psychology, University of Copenhagen

³ Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark

Funding

This research was funded by the Research Fund of the Mental Health Services in the Capital Region of Denmark with two separate grants of DKK 1.590.000 and DKK 660.000 and by Trygfonden A/S with DKK 2.670.000 (Grant no. 123488), as well as salaries from Copenhagen Trial Unit and University of Copenhagen.



Mental Health
Services

Med støtte fra

TrygFonden

CTU Copenhagen Trial Unit
Centre for Clinical Intervention Research



UNIVERSITY OF
COPENHAGEN

CONTENT

LIST OF ABBREVIATIONS AND DEFINITIONS.....	7
PREFACE	8
ACKNOWLEDGEMENTS	9
LIST OF PAPERS.....	10
Differences between my PhD thesis and doctoral dissertation	13
ABSTRACTS	14
English abstract	14
Dansk resumé	16
BACKGROUND.....	19
Borderline personality disorder	19
The new personality disorder classification in ICD-11	21
Psychotherapy for borderline personality disorder.....	22
Concurrent pharmacological interventions for borderline personality disorder.....	25
Duration of psychotherapy for borderline personality disorder	26
Patient and therapist perspectives on short-term psychotherapy.....	28
The dose-response and the good-enough-level paradigms in psychotherapy research.....	30
Evidence-based medicine	31
<i>The evidence hierarchy</i>	32
<i>Systematic errors</i>	36
<i>Random errors</i>	37
<i>Design errors</i>	38
<i>Heterogeneity</i>	38
<i>Certainty of evidence</i>	39
Duration effects of psychotherapy within an evidence-based medicine paradigm	40
OBJECTIVES.....	41
SECTION 1 - EMPIRICAL PAPERS.....	43
SYSTEMATIC REVIEW WITH META-ANALYSIS OF SHORTER VERSUS LONGER-TERM PSYCHOTHERAPY FOR ADULT PSYCHIATRIC DISORDERS.....	44

Methods	44
Results	46
Conclusion.....	47
 RANDOMISED CLINICAL TRIAL OF SHORT-TERM VERSUS LONG-TERM MENTALISATION-BASED THERAPY FOR BORDERLINE PERSONALITY DISORDER – PRIMARY RESULTS AND LONG-TERM FOLLOW-UP OF THE MBT-RCT	48
Methods	48
Results	49
<i>16 months follow-up</i>	50
<i>24 months follow-up</i>	50
Conclusion.....	51
Aftermath.....	51
 INDIVIDUAL PATIENT DATA POOLED ANALYSYS OF TWO TRIALS ASSESSING SHORT-TERM VERSUS LONG-TERM PSYCHOTHERAPY FOR BODERLINE PERSONALITY DISORDER (BPD-IPD)	54
Methods	54
Results	55
Conclusion.....	56
 PHARMACOLOGICAL INTERVENTIONS FOR CO-OCCURING PSYCHOPATHOLOGY IN PEOPLE WITH BORDERLINE PERSONALITY DISORDER: A SYSTEMATIC REVIEW.....	57
Methods	57
Results	58
Conclusion.....	58
 SECTION 2 - CLINICAL PAPERS	59
 THE CAPACITY TO END: TERMINATION OF MENTALISATION-BASED THERAPY FOR BORDERLINE PERSONALITY DISORDER	60
Methods	60
Results	60
Conclusion.....	62
 A CLINICAL ILLUSTRATION OF SHORT-TERM MENTALISATION-BASED THERAPY FOR BORDERLINE PERSONALITY DISORDER	63

Methods	63
Results	63
Conclusion.....	64
ADAPTING SHORT-TERM MENTALISATION-BASED THERAPY TO ICD-11 PERSONALITY DISORDER IN ADOLESCENTS.....	65
Methods	65
Results	65
Conclusion.....	66
SECTION 3 - METHODOLOGICAL PAPERS	67
EXTERNAL VALIDITY OF RANDOMISED CLINICAL TRIALS INCLUDING OUTPATIENTS WITH BORDERLINE PERSONALITY DISORDER	68
Methods	68
Results	68
Conclusion.....	69
MISSING OUTCOME DATA IN RANDOMISED CLINICAL TRIALS OF PSYCHOLOGICAL INTERVENTIONS: A REVIEW OF PUBLISHED TRIAL REPORTS IN MAJOR PSYCHIATRY JOURNALS	70
Methods	70
Results	70
Conclusion.....	72
IMPROVING THE METHODOLOGICAL QUALITY OF RANDOMISED CLINICAL TRIALS ASSESSING PSYCHOTHERAPY FOR BORDERLINE PERSONALITY DISORDER: RECOMMENDATIONS FOR THE FUTURE	73
Methods	73
Results	73
Conclusion.....	77
DISCUSSION.....	78
Strengths and limitations of the findings.....	78
Psychotherapy in the era of evidence-based medicine: friends or foes?	81
Selecting the optimal psychotherapy duration for the individual patient.....	83
Psychotherapy for borderline personality disorder – towards a stepped care approach?.....	85
Outcome selection and measurement in borderline personality disorder trials.....	87

<i>Excessive reliance on symptom scales.....</i>	<i>87</i>
<i>Heterogeneity in outcome selection across trials.....</i>	<i>88</i>
<i>Hard outcomes.....</i>	<i>89</i>
<i>The complexity of outcome selection in a heterogeneous group of patients</i>	<i>91</i>
CONCLUSION	93
REFERENCES	94
Appendices	113
Paper 1	113
Paper 2	113
Paper 3	113
Paper 4	214
Paper 5	226
Paper 6	235
Paper 7	363
Paper 8	416
Paper 9	419
Paper 10	113
Paper 11	476
Paper 12	506
Paper 13	519
Paper 14	530
Paper 15	537
Paper 16	564

LIST OF ABBREVIATIONS AND DEFINITIONS

BPD: Borderline personality disorder

CONSORT: Consolidated Standards of Reporting Trials

DBT: Dialectical behavioural therapy

DSM-5: Diagnostic and Statistical Manual of Mental Disorders – 5th Edition

GRADE: Grading of Recommendations, Assessment, Development and Evaluation

ICD-11: International Classification of Diseases – 11th Edition

MCID: Minimal clinically important difference

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PROSPERO: International prospective register of systematic reviews

ROB1: Cochrane Risk of Bias tool – version 1

ROB2: Cochrane Risk of Bias tool – version 2

SMD: Standardised mean difference

TAU: Treatment as usual

PREFACE

This doctoral dissertation is a culmination of my past eight years research, which I have carried out while simultaneously being affiliated with three different research institutions.

A large part of the research presented in this doctoral dissertation was carried out as part of my employment at Stolpegaard Psychotherapy Centre within the Mental Health Services of the Capital Region of Denmark. The randomised clinical trial titled MBT-RCT was implemented at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre. Once the trial was running, I became employed at the Copenhagen Trial Unit, Centre for Clinical Intervention Research at Copenhagen University Hospital, Rigshospitalet, which already served as a key collaborator on the MBT-RCT. Much of the research I have done since the finalisation of the MBT-RCT has been done within my employment at the Copenhagen Trial Unit, allowing me to specialise in clinical research methodology and evidence-based medicine.

During my PhD, I was formally enrolled at the Department of Psychology at the University of Copenhagen, where I am now permanently employed as a part time tenure-track assistant professor in clinical psychology and clinical research methodology.

To this date, I am still affiliated with all three institutions.

ACKNOWLEDGEMENTS

I am deeply grateful to my family, especially my husband Simon, for encouraging and supporting me in writing this doctoral dissertation. Without him, I simply would not have written it. I would also like to thank my 3-year-old daughter Alma and currently unborn son, not because they contributed with anything (rather the contrary), but because they motivate me to become someone they can look up to.

I have two mentors, primary collaborators, and friends, whom I want to specifically acknowledge. I want to thank Sebastian Simonsen from Stolpegaard Psychotherapy Centre for believing in me, supporting me, and entrusting me with the great responsibility of doing this research. Sebastian continues to be a safe haven for me when navigating both academia and clinical practice. I also want to thank Professor Janus Christian Jakobsen from the Copenhagen Trial Unit. Meeting him changed my academic life and was why I entered the science of clinical research methodology. Janus has always supported me with great enthusiasm, inspiration, and unlimited (and sometimes overwhelming) ambition.

Both the Stolpegaard Psychotherapy Centre and the Copenhagen Trial Unit have kindly provided means to enable the completion of this thesis, including an inspiring work environment with incredibly dedicated colleagues. I would like to thank all the people who work or have worked at Stolpegaard Psychotherapy Centre and the Copenhagen Trial Unit and all my co-authors and collaborators during the last eight years. Especially, I would like to thank Caroline Barkholt Kamp for many hours of bi-lateral supervision, Johanne Juul Petersen, Pascal Faltermeier, Johanne Ribeiro, Per Sørensen, Ole Jakob Storebø, Sune Bo, Anthony Bateman, Markus Harboe Olsen, Stig Poulsen, and Christian Gluud, for excellent cooperation and hard work.

LIST OF PAPERS

This doctoral dissertation is based on results reported in the following 16 papers. The 16 papers consist of 12 first author papers, two last author papers, one sole author paper, and one second author paper.

Paper 1

***Juul, S.**, Poulsen, S., Lunn, S., Sørensen, P., Jakobsen, J. C., & Simonsen, S. (2019). Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Systematic Reviews*. 8(1):169. [1]

Paper 2

***Juul, S.**, Jakobsen, J. C., Jørgensen, C. K., Poulsen, S., Sørensen, P., & Simonsen, S. (2023). The difference between shorter-versus longer-term psychotherapy for adult mental health disorders: a systematic review with meta-analysis. *BMC Psychiatry*, 23(1), 438. [2]

Paper 3

***Juul, S.**, Lunn, S., Poulsen, S., Sørensen, P., Salimi, M., Jakobsen, J. C., ... & Simonsen, S. (2019). Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 20(1):196. [3]

Paper 4

***Juul, S.**, Simonsen, S., Poulsen, S., Lunn, S., Sørensen, P., Bateman, A., & Jakobsen, J. C. (2021). 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*, 22(1):497. [4]

Paper 5

***Juul, S.**, Simonsen, S., & Bateman, A. (2020). The capacity to end: termination of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy*. 50:331-338. [5]

Paper 6

Juul, S., Jakobsen, J. C., Hestbaek, E., Jørgensen, C. K., Olsen, M. H., Rishede, M., ... & Simonsen, S. (2023). Short-term versus long-term mentalization-based therapy for borderline personality disorder: a randomized clinical trial (MBT-RCT). *Psychotherapy and Psychosomatics*, 92(5), 329-339. [6]

Paper 7

Juul, S., Jakobsen, J. C., Hestbaek, E., Kamp, C. B., Olsen, M. H., Rishede, M., ... & Simonsen, S. (2025). Short-term versus long-term mentalization-based therapy for borderline personality disorder: 24 months follow-up of a randomized clinical trial (MBT-RCT). *Psychotherapy and Psychosomatics* [Published online ahead of print]. DOI: 10.1159/000544934. [7]

Paper 8

Juul, S., Jakobsen, J. C., Hestbaek, E., Kamp, C. B., Olsen, M. H., Rishede, M., ... & Simonsen, S. (2024). Response to Philips et al.: Several fundamental misconceptions about evidence-based practice. *Psychotherapy and Psychosomatics*, 93(5), 348-349. [8]

Paper 9

Juul, S., McMain, S., Olsen, M. H., Chapman, A., Ribeiro, J. P., Storebø, O. J., ... Jakobsen, J. C., (2025). Short-term versus long-term psychotherapy for borderline personality disorder: protocol for an individual patient data meta-analysis [In peer review in BMJ Open]

Paper 10

Juul, S., McMain, S., Olsen, M. H., Chapman, A., Ribeiro, J. P., Storebø, O. J., ... Jakobsen, J. C., (2025). Short-term versus long-term psychotherapy for borderline personality

disorder: an individual patient data pooled analysis of two randomized clinical trials
[Submitted to Lancet Psychiatry]

Paper 11

Ribeiro, J. P., **Juul, S.**, Kongerslev, M. T., Jørgensen, M. S., Völlm, B. A., ... & Storebø, O. J. (2025). Pharmacological interventions for co-occurring psychopathology in people with borderline personality disorder: secondary analysis of the Cochrane systematic review with meta-analyses. *British Journal of Psychiatry*. 226(4), 226-237. [9]

Paper 12

Juul, S., Frandsen, F. W., Bo Hansen, S., Sørensen, P., Bateman, A., & Simonsen, S. (2022). A clinical illustration of short-term mentalization-based therapy for borderline personality disorder. *Journal of Clinical Psychology*, 78(8), 1567-1578. [10]

Paper 13

Simonsen, S., Hestbæk, E., & **Juul, S.** (2023). Adapting Short-Term Mentalization-Based Therapy to ICD-11 Personality Disorder in Adolescents. *Children*, 10(1), 93. [11]

Paper 14

Kamp, C. B., Simonsen, S., **Juul, S.** (2025). External validity of randomised clinical trials including outpatients with borderline personality disorder [Submitted to JAMA Psychiatry]

Paper 15

Juul, S., Faltermeier, P., Petersen, J. J., Olsen, M. H., Andersen, R. K., Kamp, C. B., ... & Jakobsen, J. C. (2024). Missing outcome data in randomised clinical trials of psychological interventions: a review of published trial reports in major psychiatry journals. *BMC Psychiatry*, 24(1), 798 [12]

Paper 16

Juul, S. (2022). Improving the methodological quality of randomized clinical trials assessing psychotherapy for borderline personality disorder: Recommendations for the future. *Frontiers in Psychiatry*, 13, 1053844. [13]

* Papers 1-5 have previously been submitted for evaluation as part of my PhD thesis [14]. Papers 6-16 have never previously been submitted with the intention to acquire an academic degree. For a more detailed description, see below.

All papers are available in the appendices.

Differences between my PhD thesis and doctoral dissertation

On April 20, 2021 I defended my PhD thesis at the Department of Psychology, University of Copenhagen titled “*Which length for whom? A systematic review, meta-analysis, and a randomised clinical trial of short-term versus long-term psychotherapy for adult psychiatric disorders*” [14]. The thesis included Papers 1-5 of the present doctoral dissertation, namely 1) a protocol for a systematic review, 2) the systematic review, 3) the protocol for the MBT-RCT trial, 4) the statistical analysis plan for the MBT-RCT trial, and 5) a clinical paper about termination in mentalisation-based therapy. Hence, all the results from the MBT-RCT trial including the primary results and follow-up results (Papers 6-8), the pooled analysis of individual patient data from two trials of short-term versus long-term psychotherapy for borderline personality disorder (BPD) (Papers 9 and 10), the systematic review about pharmacological interventions for BPD (Paper 11), two additional clinical papers (Papers 12 and 13), and the three methodological papers reported in this doctoral dissertation (Papers 14-16) have never been evaluated as part of acquiring an academic degree. As a whole, these additional papers and projects constitute the most significant part of this doctoral dissertation.

ABSTRACTS

English abstract

The heart of the discipline of evidence-based medicine lies in the randomised clinical trial design, with systematic reviews of these trials regarded as the highest level of evidence assessing the effects of any healthcare intervention. Finding the optimal duration of psychotherapy is a clinically important question, which should be answered using such research designs. Patients with borderline personality disorder (BPD) have historically been treated with long-term psychotherapy, often lasting a year or even longer in clinical practice, but the evidence for long-term durations of psychotherapy is unclear. The present doctoral dissertation investigates and discusses the optimal duration of psychotherapy for BPD through several empirical, clinical, and methodological papers.

The first section of this dissertation presents empirical papers assessing the beneficial and harmful effects of short-term versus long-term psychotherapy for all adult psychiatric disorders in general in a systematic review with meta-analysis, and specifically the effects of short-term versus long-term mentalisation-based therapy for adult outpatients with BPD in a randomised clinical trial assessing the effects at both end of treatment (16 months) and long-term follow-up (24 months) in 166 outpatients treated at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre (the MBT-RCT trial). Additionally, the thesis explores subgroup effects of short-term versus long-term psychotherapy (both mentalisation-based therapy and dialectical behavioural therapy) in a comprehensive, pooled analysis of individual patient data from two comparable trials of 406 outpatients with BPD. Finally, the empirical section of the thesis explores the effects of pharmacological interventions for co-occurring psychopathology in adults with BPD, as a

large proportion of patients are treated simultaneously with both psychotherapy and psychopharmacology, and the effects of such co-interventions need to be better understood.

The second section of this dissertation presents three clinical papers discussing clinical potential and pitfalls when delivering short-term psychotherapy for BPD. Particularly, the theme of treatment termination is discussed, which is a potentially potent period of treatment that can be challenging both for patients, therapists, and the therapeutic relationship. Further, the clinical application of the new short-term mentalisation based therapy approach to treating BPD proposed as part of the MBT-RCT trial, is discussed through a patient case study of a young woman diagnosed with BPD with prominent symptoms such as affective dysregulation, unstable relationships, and intense abandonment anxiety. Furthermore, the aim was to discuss whether the short-term mentalisation-based therapy programme from the MBT-RCT trial could be a suitable treatment option for adolescents with BPD, particularly considering the new ICD-11 personality disorder classification.

The third section of this dissertation presents three methodological papers discussing the potential methodological limitations when assessing the optimal duration of psychotherapy for BPD. It discusses the external validity of randomised clinical trials of BPD by exploring the possible limitations of implementing screening procedures once the patient has already entered the outpatient setting in contrast to possible inclusion already at the referral stage. Furthermore, it discusses the potential bias risk associated with missing outcome data in psychological research in general, which is a specific problem in personality disorder research and also apparent in the MBT-RCT trial. Finally, it proposes some methodological recommendations for future randomised clinical trials assessing psychotherapy for BPD.

Dansk resumé

Kernen i evidensbaseret medicin ligger i det randomiserede kliniske forsøg, hvor systematiske gennemgange af disse forsøg betragtes som det højeste evidensniveau til at vurdere effekterne af enhver form for sundhedsfaglig intervention. Spørgsmålet om den optimale varighed af psykoterapi til psykiatriske patienter er et klinisk vigtigt spørgsmål, som bør besvares ved hjælp af sådanne forskningsdesigns. Patienter med borderline personlighedsforstyrrelse har historisk set været behandlet med langvarig psykoterapi, ofte et år eller af længere varighed i klinisk praksis, men evidensen for den optimale varighed af psykoterapi er uklar. Denne doktorafhandling undersøger og diskuterer den optimale varighed af psykoterapi med særlig fokus på patienter med borderline personlighedsforstyrrelse, gennem flere empiriske, kliniske og metodologiske artikler.

I den første del af afhandlingen præsenteres empiriske artikler, der vurderer de gavnlige og skadelige effekter af kortvarig versus langvarig psykoterapi for alle voksenpsykiatriske lidelser generelt i en systematisk gennemgang med meta-analyse, med særlig fokus på de specifikke effekter af kortvarig versus langvarig mentaliseringsbaseret terapi for voksne ambulante patienter med borderline personlighedsforstyrrelse i et randomiseret klinisk forsøg. I forsøget vurderes effekterne både ved behandlingsafslutning (16 måneder) og ved langtidsopfølgning (24 måneder) hos 166 ambulante patienter fra Ambulatoriet for Personlighedsforstyrrelser ved Psykoterapeutisk Center Stolpegård i Region Hovedstadens Psykiatri (MBT-RCT-forsøget). Desuden udforsker afhandlingen subgrupp-effekter af kortvarig versus langvarig psykoterapi (både mentaliseringsbaseret terapi og dialektisk adfærdsterapi) i en omfattende analyse af individuelle patientdata fra 406 danske og canadiske ambulante patienter med borderline personlighedsforstyrrelse. Endeligt

undersøger den empiriske sektion af afhandlingen effekterne af farmakologiske interventioner for samtidig psykopatologi hos voksne med borderline personlighedsforstyrrelse, idet en stor andel af patienterne behandles samtidig med både psykoterapi og psykofarmakologi, og effekterne af sådanne co-interventioner er nødvendige at forstå, når effekterne af psykoterapi skal estimeres.

I den anden del af afhandlingen præsenteres tre kliniske artikler, der diskuterer kliniske potentialer og udfordringer ved kortvarig psykoterapi til patienter med borderline personlighedsforstyrrelse. Særligt diskuteres behandlingsafslutning som en central og klinisk betydende komponent i behandlingen, som stiller store krav til patienter, klinikere og den terapeutiske relation. Desuden diskuteres den kliniske anvendelse af den nye kortvarige udgave af mentaliseringsbaseret terapi til behandling af borderline personlighedsforstyrrelse, som foreslået i MBT-RCT-forsøget, gennem et casestudie af en ung kvinde diagnosticeret med borderline personlighedsforstyrrelse, hvis fremtrædende symptomer var emotionel dysregulering, ustabile relationer og intens forladthedsangst. Til slut diskuteres det, hvorledes det kortvarige mentaliseringbaserede terapiprogram fra MBT-RCT-forsøget kunne være en passende behandlingsmulighed for unge med borderline personlighedsforstyrrelse, især i betragtning af den nye ICD-11 klassifikation af personlighedsforstyrrelser.

I den tredje del af afhandlingen præsenteres tre metodologiske artikler, der diskuterer de potentielle metodologiske begrænsninger ved vurderingen af den optimale varighed af psykoterapi for borderline personlighedsforstyrrelse. Det diskuteres, hvordan den eksterne validitet af randomiserede kliniske forsøg vedrørende borderline personlighedsforstyrrelse

kan påvirkes af eventuelle begrænsninger ved implementeringen af screeningsprocedurer. Desuden diskuteres den potentielle bias risiko (risiko for systematiske fejl) forbundet med manglende data i psykologisk forskning generelt, hvilket er et specifikt problem i forskning i personlighedsforstyrrelser, og som også var tydeligt i MBT-RCT-forsøget. Endelig fremsættes metodologiske anbefalinger til fremtidige randomiserede kliniske forsøg, der vurderer psykoterapi for borderline personlighedsforstyrrelse.

BACKGROUND

In the first part of this doctoral dissertation, I will introduce key theoretical and conceptual understandings from the included papers. I will introduce the borderline personality disorder (BPD) diagnosis, the new personality disorder classification in ICD-11, psychotherapy for BPD, concurrent pharmacological interventions when delivering psychotherapy, duration of psychotherapy for BPD, patient and therapist perspectives on short-term psychotherapy, the dose-response and the good-enough-level paradigms in psychotherapy research, clinical research methodology, and finally, I will introduce the optimal research design to assess duration effects of psychotherapy.

Borderline personality disorder

The classification of BPD has evolved historically. What would later be recognised as BPD can be traced back to clinical observations in the 1930s and 1940s, and many prominent clinical researchers have since debated its ontology. One of them is the American psychiatrist and psychoanalyst Adolph Stern, who first described patients who did not fit into existing diagnostic categories and used the term “borderline patients” to describe patients who would neither fall into the category of neurosis nor psychosis [15]. Later, the American psychiatrist Otto Kernberg played a significant role in defining and describing borderline traits during the 1960’s and 1970’s, particularly with the introduction of the term “borderline personality organization”, referring to both a descriptive classification of borderline pathology between neurosis and psychosis, and also to a structural analysis of the aetiology behind the disorder [16].

Building on these early theoretical and conceptual contributions about the ontogeny of the diagnosis, the American psychiatrist and psychoanalyst John Gunderson proposed a revised definition, developed a structured clinical interview and, based on these, established the discriminating features and diagnostic threshold of the disorder [17]. These criteria later formed the basis for the disorder's entry into the official classification system in 1980, the American Psychological Association Diagnostic and Statistical Manual for Mental Disorders – 3rd Edition (DSM-3) largely similar to the updated version in DSM-5 [18].

Today, BPD is considered a distinct psychiatric disorder categorised in the DSM-5 with nine symptoms including 1) frantic efforts to avoid real or imagined abandonment, 2) intense and unstable interpersonal relationships, 3) identity diffusion, 4) impulsivity, 5) recurrent suicidal behaviour, 6) affective instability, 7) chronic feelings of emptiness, 8) intense anger, and 9) transient, stress-related paranoia or severe dissociation [19]. Five out of the nine diagnostic criteria will constitute a diagnosis of BPD [19]. Additionally, people with BPD typically suffer from one or multiple co-occurring psychiatric disorders, including mood disorders, anxiety disorders, eating disorders, posttraumatic stress disorder, other personality disorders, and substance use disorders [20]. BPD is characterized by severe functional impairments, a high risk of suicide-related mortality, a negative effect on the course of depressive disorders, extensive use of health care services, and high societal costs [21, 22]. BPD is a prevalent psychiatric diagnosis, which is found in approximately 1.7% of the general population but in 15% to 28% of patients in psychiatric clinics or hospitals and in a large proportion of individuals seeking help for psychological problems in secondary care [23].

The clinical presentation of BPD can vary substantially between patients [24] and within the same patient across time [25]. The five out of nine criteria needed to reach a full diagnosis in DSM-5 [19] results in 256 possible combinations that can lead to the diagnosis [26]. Hence, two patients diagnosed with BPD may not overlap in more than one criterion [26]. Additionally, patients meeting seven or more BPD criteria have been shown to present with greater rates of substance use disorders and co-occurring disorders, as well as a greater number of suicidal behaviours compared with patients meeting less than seven BPD criteria [27]. Furthermore, a longitudinal study of the five-year clinical course of BPD patients showed that adolescent BPD patients were later also diagnosed with anxiety disorders (37%), depressive disorders (32%), PTSD or complex PTSD (20%), schizophrenia (16%), and eating disorders (13%) [28]. Only 16% did not meet criteria for any mental disorder. Hence, there is often a continued clinical burden proceeding after initial diagnosis with large degrees of co-morbidities and, in some cases, a change of primary diagnosis - e.g. from personality disorder to schizophrenia, maybe because of misdiagnosis or later onset of psychotic symptoms.

The new personality disorder classification in ICD-11

While the DSM-5 is mostly used in research, the World Health Organization (WHO) International Classification of Diseases (ICD) system is mainly used as a diagnostic system in clinical practice, particularly in Europe. The 11th revision of the ICD system (ICD-11) [29] includes a fundamentally new approach to diagnosing personality disorders [30]. In short, the new ICD-11 personality disorder classification system abandons the categorical personality disorder diagnoses from the DSM-5 and ICD-10 (with their 8 to 12 distinct personality disorders) with a shift towards a dimensional approach focusing on a severity

classification of global and shared features that apply to all personality disorders [30]. Additionally, the ICD-11 includes trait domain specifiers (negative affectivity, detachment, disinhibition, dissociality, anankastia), and an additional borderline pattern specifier comparable to the categorial DSM-5 diagnosis of BPD. Hence, the personality disorder model proposed in the ICD-11 addresses the issue of co-occurrence or comorbidity of multiple personality disorders by focusing on the severity of the personality disorder itself, rather than on heterogeneous categories [30].

Psychotherapy for borderline personality disorder

Over the past decades, a variety of different psychotherapies have been developed specifically aimed at treating patients with BPD. The most common psychotherapies, both in research and clinical practice, include dialectical behavioural therapy, mentalisation-based therapy, schema-focused therapy, transference-focused psychotherapy, and long-term psychodynamic/psychoanalytical psychotherapy. Below I will give a brief introduction to dialectical behavioural therapy and mentalisation-based therapy, as these psychotherapies will be the focus of the included papers of this dissertation.

Dialectical behavioural therapy, originally developed by Marsha Linehan, is a third-wave cognitive behavioural therapy comprised of a broad range of cognitive, behavioural, and problem-solving therapeutic interventions, including skills training and in-session coaching, chain analysis, contingency management, and exposure [31, 32]. These interventions are often paired with mindfulness and other Zen Buddhism practices to decouple extreme emotional arousal from maladaptive behavioural responses like self-harm or alcohol- or drug use [33]. The ‘dialectic’ in the treatment refers to the continued therapeutic shifts

between validation and the strive towards change [31]. Traditional dialectical behavioural therapy includes both individual sessions and group-based skills training modules focused on mindfulness, interpersonal effectiveness, distress tolerance, and emotion regulation.

Mentalisation-based therapy, originally developed by Peter Fonagy and Anthony Bateman, is a psychodynamic therapy rooted in attachment theory [34, 35]. Mentalisation refers to the capacity to understand own and others internal mental states such as beliefs and intentions [36-38]. The goal of mentalisation-based therapy is to support patients in enhancing their capacity to mentalise, particularly in situations where they experience emotional distress. According to the mentalisation theory, patients with BPD experience so-called non-mentalising modes more frequently than others, where they temporarily lose the capacity to balance several perspectives at once [36]. Therapeutic techniques to help the patient shift back into a mentalising mode include, e.g. stop-and-rewind, empathic validation, focus on emotional states, and mentalising the relationship and the counter-relationship [37].

The existing psychotherapies developed for BPD vary in their therapeutic focus and therapeutic stance. However, they also share some common features including focusing on emotional regulation, increasing distress tolerance, and reducing maladaptive behaviour patterns using various therapeutic techniques. Furthermore, starting each psychotherapy with psychoeducation followed by developing a case formulation is a concept shared by many psychotherapies [39].

Even though several psychotherapies have been developed, there is still a scarcity of randomised clinical trials assessing the effects of each specific psychotherapy type. The

most recent systematic review synthesising the evidence for specialised psychotherapy for BPD is a Cochrane systematic review published in 2020 [40]. In short, the Cochrane review identified 75 randomised clinical trials randomising a total of 4507 participants, predominantly involving females with a mean age range from 14.8 to 45.7 years. More than 16 different types of psychotherapy were included, but dialectical behaviour therapy and mentalisation-based therapy were the psychotherapy types with the most primary trials. The comparator interventions included treatment-as-usual (TAU), waiting list, and other active treatments. Duration of psychotherapy ranged from one to 36 months in the included trials.

Overall, the authors of the Cochrane review found beneficial effects on all primary outcomes in favour of BPD-tailored psychotherapy compared with TAU [40]. Specifically, they found that dialectical behavioural therapy was more effective compared with treatment as usual in reducing BPD severity (SMD -0.60 , 95% CI -1.05 to -0.14 ; 3 trials, 149 participants), self-harm (SMD -0.28 , 95% CI -0.48 to -0.07 ; 7 trials, 376 participants) and improving psychosocial functioning (SMD -0.36 , 95% CI -0.69 to -0.03 ; 6 trials, 225 participants) [40]. Furthermore, they found that mentalisation-based therapy was more effective than treatment as usual at reducing self-harm (RR 0.62 , 95% CI 0.49 to 0.80 ; 3 trials, 252 participants), suicidality (RR 0.10 , 95% CI 0.04 , 0.30 , 3 trials, 218 participants) and depression (SMD -0.58 , 95% CI -1.22 to 0.05 , 4 trials, 333 participants). However, all results were based on low certainty of evidence, meaning that the reported effects may be correct or may not be correct [40].

Concurrent pharmacological interventions for borderline personality disorder

According to clinical guidelines on the management of BPD, psychotherapy is considered the primary treatment of choice [41]. However, many patients with BPD receive concurrent pharmacotherapy including a large variety of substance classes and oftentimes many drugs concurrently [42]. It is important to know the effects of such co-interventions, as they can potentially complicate the interpretation of psychotherapy trial results, particularly in unblinded trials where an unbalanced administration pattern between groups can seriously compromise the trial results, e.g. if participants randomised to short-term psychotherapy are given additional pharmacological co-interventions compared with participants randomised to long-term psychotherapy.

Reasons for the prescriptions of pharmacological interventions for patients with BPD typically include an absence of psychotherapy at the clinics [43], co-occurring psychopathology as an indication for pharmacological interventions, or the addition of new substances in times of crisis, which, in the absence of regular medication reviews, simply accumulate over time [44]. However, the evidence for the effects of pharmacological interventions for BPD patients is scarce and of low to very low certainty. A Cochrane systematic review was recently updated and published, assessing the beneficial and harmful effects of pharmacological interventions for people with BPD [44-46]. It included 46 randomised clinical trials randomising a total of 2769 participants with BPD to either pharmacotherapy (e.g., antipsychotics, antidepressants, or anticonvulsants) or placebo. No restrictions were applied concerning co-occurring psychopathology. Primary outcomes were BPD symptom severity, self-harm, suicide-related outcomes and psychosocial functioning. The review found no evidence of a difference between pharmacotherapy versus

placebo on any of the primary outcomes [44-46]. Hence, while pharmacotherapy is often used to treat co-morbidities related to BPD, including depression and anxiety [47, 48], there is still no convincing evidence that pharmacotherapy is the most suitable treatment option for BPD [47, 48], and it is therefore often assigned an adjudicative role.

Beyond core BPD symptoms, many patients with the disorder also suffer from co-occurring psychopathologies, including e.g., major depressive disorder, anxiety disorders, psychotic disorders, and substance use disorders[20]. These co-occurring symptoms can worsen functional impairment and complicate treatment planning. While psychotherapeutic approaches remain the primary recommended treatment for BPD, pharmacological interventions are frequently prescribed in clinical practice to target co-occurring symptomatology. However, despite their widespread use, robust evidence to guide such off-label pharmacological interventions remains sparse to non-existing.

Duration of psychotherapy for borderline personality disorder

Traditionally, patients with BPD have been given long-term psychotherapy as a standard treatment due to the complexity and long-term nature of the disorder and the assumed necessity of an equally complex and long-term treatment. There is no generally recognised standard definition of ‘long-term psychotherapy’. For example, long-term psychodynamic psychotherapy has been defined by experts in the field to range from a minimum of three months to a maximum of 20 years [49]. Different long-term psychotherapies have been developed or adapted to treat patients with BPD over the past decades. These include long-term psychodynamic therapy (usually delivered for complex psychiatric disorders including BPD) lasting one year or more [49], dialectical behavioural therapy for up to one year [50],

transference-focused psychotherapy up to three years [51], schema-focused therapy up to three years [51], and mentalisation-based therapy, which was originally manualised and tested for 18 months [36, 52] but is being delivered in longer durations in some clinical settings, e.g. up to three years of outpatient treatment in Norway [53].

However, there is a considerable demand versus supply gap of evidence-based psychotherapies for BPD, which indicates the need for briefer alternatives to long-term psychotherapies or to pay more attention to modular steps of care to tailor such steps for individual patients' needs during the trajectory of the clinical course of their psychiatric disorder [54]. Hence, many clinicians and researchers have begun to look for shorter-term alternatives to the existing treatments for BPD [55].

In the past 10 to 20 years, it has been increasingly recognised that short-term psychotherapy should be developed and adapted as a possible treatment option for patients with BPD to meet the increasing demands of the public mental health system. These short-term treatment options now include, but are not limited to, brief dialectical behaviour skills therapy for 20 weeks [56], emotion regulation group therapy for 14 weeks [57], systems training for emotional predictability and problem solving for 20 weeks [58], and short-term mentalisation-based therapy (which will be presented in this thesis).

In 2016, Stolpegaard Psychotherapy Centre, a psychiatric outpatient hospital for non-psychotic disorders within the Mental Health Services of the Capital Region of Denmark, experienced increased waiting lists for outpatient treatment of patients with personality disorders. As a solution to the increasing problem, the leadership, together with the staff,

decided to design and implement short-term alternatives to the existing long-term mentalisation-based programme, which had originally been implemented at the clinic since 2010-2011. The short-term mentalisation-based therapy programme was designed as a 20 weeks programme including five weeks of psychoeducation followed by 15 sessions of mentalisation-based group therapy combined with 10 to 12 individual sessions with one of the two group therapists. Seven to nine patients could start and end the group together, meaning that the groups were closed. From 2017-2018, four pilot groups were started, the process was evaluated internally at the clinic, and at the end of 2018, the MBT-RCT trial recruited the first participant to a randomised clinical trial comparing the effects of short-term (20 weeks; experimental intervention) versus long-term (14 months; control intervention) of mentalisation-based therapy for outpatients with BPD.

Patient and therapist perspectives on short-term psychotherapy

After the pilot and implementation phases of the MBT-RCT trial, we performed two qualitative studies about the patient [59] and therapist [60] perspectives on receiving and delivering short-term mentalisation-based therapy for BPD (not included in this doctoral dissertation). In our qualitative study about patient perspectives on short-term mentalisation-based therapy, we performed semi-structured qualitative interviews with 12 BPD patients from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, who attended short-term MBT for five months. We used verbatim transcription of the interviews which were analysed using thematic analysis with double coding. Four themes emerged from the analyses: (1) Treatment duration – too short or appropriately short?, (2) The group as a “safe space,” (3) Bad experiences impacted treatment negatively, and (4) My life has changed for the better [59]. Our results suggest

that most of the participants were satisfied with receiving short-term MBT. However, a subgroup of participants desired to have their treatment extended after the short-term programme as they did not find the program to be sufficient, which they discussed with their therapists in the termination phase of treatment. Hence, our study highlights important strengths and limitations of receiving short-term mentalisation-based therapy for BPD as experienced by the patients.

In our qualitative study about therapist perspectives on short-term mentalisation-based therapy, we performed semi-structured qualitative interviews with seven therapists from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre about their experiences with short-term mentalisation-based therapy after a one-year pilot phase. Again, we used verbatim transcription of the interviews, which were analysed using thematic analysis with double coding. Four major themes were found, which are summarised in **Figure 1** including two overarching themes. In short, we found that most therapists were overall reluctant to change from long-term to short-term MBT. Most of them had experiences with delivering long-term psychotherapy for many years, and they pointed to some clinical challenges in having to shorten the treatment, including problems with difficult treatment termination, and the worry that some patients may not benefit sufficiently from a short-term treatment, which may be considered inadequate or superficial. These therapists' experiences could inform the implementation of short-term MBT in future clinical management of outpatients with BPD [60].

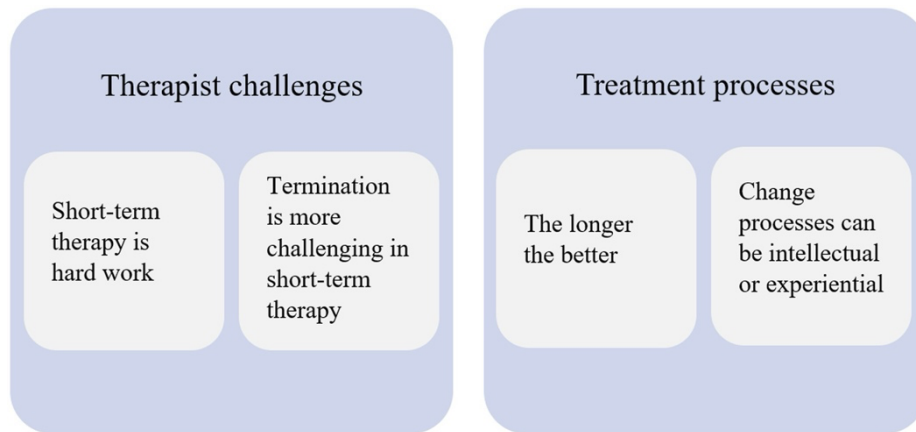


Figure 1. Therapist attitudes to short-term MBT are illustrated by four themes and their relation to two overarching themes [60].

The dose-response and the good-enough-level paradigms in psychotherapy research

In clinical practice, the duration of psychotherapy varies considerably across patient populations and psychotherapy types. Deciding how long psychotherapy should last and when the individual patient outcomes have reached an optimal level is therefore a key challenge for clinicians [61]. The question of the optimal duration of psychotherapy has been debated in the literature for several decades given its clinical, ethical, and economic implications [62, 63].

Two prominent paradigms have emerged to answer the question of the optimal duration of psychotherapy: the dose-response paradigm [64] and the good-enough-level paradigm [65]. According to the dose-response paradigm, the relationship between treatment duration (typically measured in the number of sessions) and outcomes (typically assessed with symptom scales) is typically characterised by a negatively accelerating curve, where symptomatic improvement mostly occurs in the early phases of psychotherapy and tends to lessen thereafter. Key assumptions of the dose-response paradigm are that most patients

tend to follow this negatively accelerating curve and that the duration or “dose” of psychotherapy causes changes to occur, but this effect tends to lessen over time [64]. In the good-enough-level paradigm, the duration of psychotherapy has been argued to result from responsive regulation by patients and clinicians, where psychotherapy continues until a good-enough level of improvement is attained [65]. According to the good-enough level paradigm, psychotherapy duration is not necessarily the cause of improvement, but rather a consequence of patients’ responsivity (or lack of responsivity) to psychotherapy [61]. Hence, some patients characterised as non-responders may require long-term psychotherapy to reach a good-enough level.

Both the dose-response and the good-enough-level paradigms are based on results from non-controlled observational studies [66, 67]. While several non-controlled observational studies indicate that there is a linear relationship between the number of psychotherapy sessions and outcomes for most psychiatric disorders, there is still conflicting evidence as to which psychotherapy duration is the most optimal for most psychiatric disorders [61]. The inconclusiveness of the existing research and the methodological problems associated with non-controlled observational studies [68, 69] indicate the need for research on duration effects using the principles of evidence-based medicine.

Evidence-based medicine

Evidence-based medicine is “the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients” [70]. Evidence-based medicine started as a movement in the early 1990s with the primary focus on educating clinicians in the understanding and use of published research to guide clinical decision-

making [71]. The primary goal of the evidence-based medicine movement was to replace an approach to healthcare based on authority and unsystematic clinical experience – so-called 'expert-based medicine' – with an approach based on the best currently available research-based evidence. Evidence-based medicine progressed to recognise the limitations of the published research and has increasingly stressed the need to combine critical appraisal of the evidence with patient's values and preferences through shared decision making [71]. Hence, applying an evidence-based medicine approach to health care is a three-legged stool involving 1) research evidence, 2) clinical experience, and 3) values, preferences, and characteristics of service users [70, 72]. However, when selecting the 'current best evidence', a crucial part of applying an evidence-based medicine approach is the notion of *the evidence hierarchy*.

The evidence hierarchy

In the evidence hierarchy, the systematic reviews with meta-analysis of randomised clinical trials, are regarded as the highest level of evidence when assessing the effects of any healthcare intervention followed by the single randomised clinical trial [73]. The benefit of the randomised clinical trial design is that all known and unknown variables that may potentially influence the intervention effect will be balanced between the groups (assuming the number of randomised participants is large enough), allowing researchers to draw causal conclusions about intervention effects without the risk of misinterpreting the results due to confounding [74]. Within many medical specialties, numerous randomised clinical trials have been published testing the same hypothesis. Meta-analysis of such trials can then be used to generate a summary of the results of these different trials often expressed by a point estimate, a confidence interval, and a p-value. Hence, with a meta-analysis it is possible to

synthesise the results of several similar trials to provide an estimate representing an overall direction and size of the intervention effect for a specific outcome of interest. A good reason for conducting a meta-analysis is the expected gain in statistical power. Combining trials in a meta-analysis often results in a higher level of power compared to the individual trials. This is particularly important if the individual trials for example used a continuous outcome (e.g. a symptom scale) as the primary outcome, which often results in a relatively low required sample size, but the outcome of the systematic review is binary (e.g. proportion of participants with a suicide-attempt) [75]. In such a case, the required sample size for the systematic review will be substantially larger than for the individual trials. Heterogeneity between the included trials can further enhance the required sample size (see Heterogeneity-section below).

An additional advantage of combining several trials in a systematic review with meta-analysis is the possibility of providing the most up-to-date summary of the individual trials that address a similar question to provide a more reliable estimate of treatment effects for a given patient population [76]. However, the pooled results of many smaller trials incorporate biases from the original trials and embody new sources of bias, mostly because of the selection of trials and the inevitable heterogeneity among them [77]. There is a continuing debate about which evidence is the most reliable: the pooled meta-analysis results of smaller trials from a systematic review or the results of one large randomised clinical trial [78-81]. In the evaluation of evidence from different sources, one must consider the impacts of systematic errors, random errors, and certainty of the evidence (GRADE) (please see below).

Data used for meta-analysis comes in two forms: aggregate data or individual patient data. Aggregate data provides data at a trial level, such as point estimates, standard deviations, and the number of analysed participants. Aggregate data can often be collected from the published trial reports, or by contacting the trialists if the aggregate data is not available in the trial reports. Individual patient data is the raw patient data on each participant individually and can sometimes be obtained from public repositories in a completely anonymised format, or by contacting the trialists of the original trial and requesting individual patient data [82]. One of the advantages of performing individual patient data meta-analysis is the possibility to perform more detailed subgroup analyses, because all the individual trial data is available for new analyses.

Observational studies using non-randomised analyses of typically large electronic patient databases are often being promoted as valid alternatives to randomised clinical trials as a source of “real-world evidence” about the beneficial and harmful effects of new and existing treatments [83]. However, large observational studies may yield misleading associations of interventions with health outcomes that are statistically significant when assessed as pre-post data but noncausal, or that are mistaken. According to the evidence hierarchy, observational studies (for example naturalistic, pre-post studies with repeated measurements of a population without a control group) are viewed as having less internal validity because they are likely to overestimate intervention effects, as it is difficult to distinguish an effect of the intervention from spontaneous change from baseline (i.e. regression towards the mean) and other confounding [84, 85]. Expert theories and clinical expertise or opinions constitute the bottom of the hierarchy of evidence [86].

Using clinical research evidence with the best quality is essential for choosing the right treatment for patients. When applying an evidence-based medicine approach, the treatment of an individual patient should generally not be chosen based only on evidence from expert-opinion, case-control studies, observational studies, or single randomised clinical trials, unless these designs constitute the best available evidence within a given clinical area. Rather, systematic reviews with meta-analysis of all identifiable randomised clinical trials taking into account the risk of systematic and random errors and the certainty of the evidence represent the highest level of evidence [87].

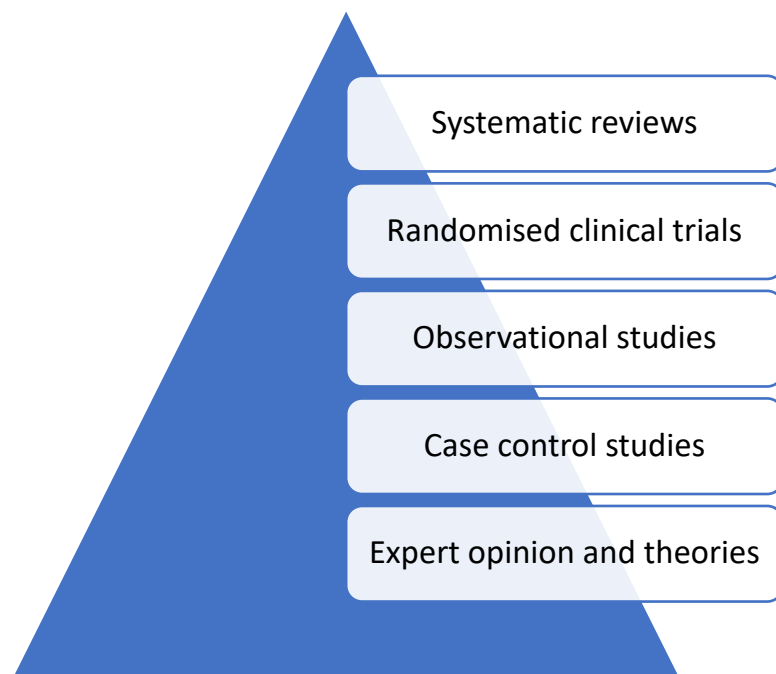


Figure 2. The evidence hierarchy

All levels of the hierarchy of evidence may be threatened by systematic errors, random errors, and design errors [68, 76].

Systematic errors

Systematic errors (or bias) in randomised clinical trials included in systematic reviews may impact the effect estimates of the assessed interventions, which in turn may impact the effect estimates of the meta-analyses, and ultimately the clinical recommendations proposed based on the results. Such systematic errors do not spuriously skew the results. On the contrary, meta-epidemiological studies have shown that randomised clinical trials at unclear or high risk of bias have a tendency to overestimate the beneficial effects and underestimate the harmful effects of the experimental intervention [88-95]. Trials at high risk of bias may therefore ultimately mislead clinical decision-making.

Until recently, Cochrane encouraged the use of the Cochrane Risk of Bias tool (ROB 1) originally published in the Cochrane Handbook for Systematic Review of Interventions in 2008 and updated in 2011 [96]. Review authors can use this tool to assess bias in randomised clinical trials in several bias risk domains [97]. Risk of bias in each domain is assessed by two independent review authors as low, unclear or high. Subsequently, these judgements lead to a bias risk assessment on an outcome level as well as on a trial level. Cochrane has recently updated the assessment tool to Cochrane Risk of Bias tool – version 2 (ROB 2) [98], particularly by adding several signalling questions for a more precise estimate of each bias risk domain. However, the ROB2 tool has also been criticised for being too difficult and resource-intensive to employ [99]. For an overview of the bias risk domains in both tools, see **Table 1**.

Table 1. Bias risk domains in Cochrane Risk of Bias tool versions 1 and 2

Cochrane Risk of Bias tool – version 1 (ROB 1) [97]	Cochrane Risk of Bias tool – version 2 (ROB 2) [98]
Generation of the allocation sequence	Risk of bias arising from the randomisation process
Allocation concealment	
Blinding of participants and treatment providers	Risk of bias arising from deviations from the intended interventions
Incomplete outcome data	Risk of bias arising from missing outcome data
Blinding of the outcome assessor	Risk of bias in measurement of the outcome
Selective outcome reporting	Risk of bias in selection of the reported result
Other bias	

Random errors

Meta-analyses of randomised clinical trials frequently include too few participants to achieve sufficient statistical power, making it challenging to differentiate genuine negative findings from false negatives when evaluating expected intervention effects [100]. Inadequately powered meta-analyses statistically resemble interim analyses in randomised clinical trials, essentially providing an 'early look' at data. Repeated evaluations of accumulating data increase the likelihood of erroneously concluding an effect when repetitive testing of accumulating data is conducted [101, 102]. Consequently, meta-analyses with limited data are more susceptible to incorrect conclusions, such as false positives (type I errors) or false negatives (type II errors), due to random error (play of chance) [101, 102].

Trial Sequential Analysis (TSA) is a statistical method to mitigate these risks [103-107]. Using the Lan-DeMets-O'Brien-Fleming boundaries [106, 108], TSA maintains type I error control regardless of sequential testing. TSA ensures that a meta-analytic sample size (required information size) is calculated, which (for dichotomous outcomes) is based on the proportion of events in the control group, the estimated intervention effect in the experimental group, predefined type I and II error risks, and the heterogeneity [103]. TSA uses sequential monitoring boundaries that adjust the significance threshold dynamically according to the accumulating patient data. Stricter monitoring boundaries necessitate lower P-values for statistical significance, reducing the likelihood of spurious findings [102].

Design errors

When a randomised clinical trial has a satisfactory degree of internal validity, i.e. low risks of systematic errors and random errors, it becomes relevant to study the degree of design errors (external validity) [109]. Design errors typically refer to trials having a wrong design or the wrong context to answer the question posed [109]. Different design errors exist, but typical examples in psychological intervention research include screening and selecting participants who are not comparable to the background clinical population, choosing the wrong comparator group, assessing participants at one time-point in the experimental group and at another time-point in the comparator group, and relying on outcomes with limited importance for patients.

Heterogeneity

The combination of multiple randomised clinical trials in a meta-analysis has an inherent challenge of heterogeneity, which can obscure the estimation of a 'true' effect and

complicate the interpretation of results. Heterogeneity can be caused by many factors related to e.g. heterogeneous participants, intervention groups, control groups, and outcomes. Substantial heterogeneity can undermine the reliability of the meta-analytic results by producing diverging results and misleading conclusions [110, 111]. Therefore, clinical and statistical heterogeneity should always be considered when planning meta-analyses [76]. Heterogeneity can be identified by visually inspecting the forest plot of a meta-analysis and by calculating the statistical heterogeneity (for example by inconsistency (I^2) or diversity (D^2)). If large clinical or statistical heterogeneity exists, a meta-analysis should be avoided, and the results should be narratively reported [76].

Certainty of evidence

When grading certainty of the accumulated body of evidence in systematic reviews with meta-analysis, the Grading Recommendations, Assessment, Development, and Evaluation (GRADE) approach should be used [112-115]. Assessments of certainty of evidence are performed at the last stages of a systematic review when all evidence has been analysed and methodologically evaluated. Certainty of evidence for each systematic review outcome is appraised and presented in a Summary of Findings table. According to the GRADE approach, randomised clinical trials begin as high-quality evidence, and are subsequently downgraded (if relevant) according to five factors: 1) risk of bias in the included trials; 2) indirectness of the evidence; 3) heterogeneity; 4) imprecision of effect estimates; and 5) risk of publication bias. Ultimately, the certainty of evidence for each outcome falls into one of four categories: high, moderate, low, or very low certainty of evidence [112-115].

Duration effects of psychotherapy within an evidence-based medicine paradigm

When applying an evidence-based medicine approach to assessing the duration effects of psychotherapy, the appropriate research design to answer the question “which duration of psychotherapy is better for patients with X psychiatric disorder?” would be to randomise participants to different durations of treatment with the same psychotherapy type. Only by applying such a design is it possible to assess the true difference between durations thus eliminating confounding, as the randomisation process, if implemented adequately, will result in identical baseline characteristics of both known and unknown confounding variables in both groups.

The empirical papers presented in this doctoral dissertation have applied such a design using different methodologies, i.e. 1) a systematic review with meta-analysis; 2) a randomised clinical trial; and 3) a pooled analysis of individual patient data from two comparable trials. Additionally, one empirical paper assessing the effects of pharmacotherapy for co-occurring psychopathology in people with BPD has used a traditional placebo-controlled design.

OBJECTIVES

Empirical papers

The aim of the empirical papers reported in this doctoral dissertation was to assess the beneficial and harmful effects of short-term versus long-term psychotherapy for all adult psychiatric disorders in a comprehensive systematic review with meta-analysis. Additionally, the aim was to specifically assess the beneficial and harmful effects of short-term versus long-term mentalisation-based therapy for adult outpatients with BPD in a randomised clinical trial evaluating the effects at both end of treatment (16 months) and long-term follow-up (24 months) in 166 outpatients treated at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre (the MBT-RCT trial). Furthermore, the aim was to explore specific subgroup effects of short-term versus long-term psychotherapy (both mentalisation-based therapy and dialectical behavioural therapy) in a comprehensive, pooled analysis of individual patient data from the MBT-RCT and FASTER trials. Finally, the aim was to explore the effects of pharmacological interventions for co-occurring psychopathology in adults with BPD, as a large proportion of patients are treated simultaneously with both psychotherapy and psychopharmacology, and the effects of such co-interventions need to be better understood.

Clinical papers

The aim of the clinical papers reported in this doctoral dissertation was to discuss the crucial phase of termination of psychotherapy, as this is a particularly vulnerable phase for patients with BPD, with great potential to influence the duration of treatment. Additionally, we aimed to discuss the clinical application of the new short-term mentalisation based therapy approach to treating BPD proposed as part of the MBT-RCT trial through a patient case

study of a young woman diagnosed with BPD and with a history of treatment dropouts and with prominent symptoms such as affective dysregulation, unstable relationships, and intense abandonment anxiety. Furthermore, the aim was to discuss whether the short-term mentalisation-based therapy programme from the MBT-RCT trial could be a suitable treatment option for adolescents with BPD, particularly considering the new ICD-11 personality disorder classification.

Methodological papers

The aim of the methodological papers reported in this doctoral dissertation was to discuss the potential methodological limitations when assessing the optimal duration of psychotherapy for BPD. The first aim was to examine the external validity of randomised clinical trials of BPD by exploring the possible limitations of implementing inclusion procedures once the patient has already entered the outpatient setting in contrast to possible inclusion already at the referral stage. Furthermore, the aim was to discuss the potential bias risk associated with missing outcome data in psychological research in general, which is a specific problem in personality disorder research and also apparent in the MBT-RCT trial. Finally, the aim was to propose some methodological recommendations for future randomised clinical trials assessing psychotherapy for BPD.

SECTION 1 - EMPIRICAL PAPERS

SYSTEMATIC REVIEW WITH META-ANALYSIS OF SHORTER VERSUS LONGER-TERM PSYCHOTHERAPY FOR ADULT PSYCHIATRIC DISORDERS

The results of this project are reported in the following dissertation papers: **Paper I** and **Paper 2**.

Methods

In this project, we protocolised [1] and subsequently performed a systematic review with meta-analysis of the comparative effects of shorter versus longer-term psychotherapy for all adult psychiatric disorders [2] aiming to form the basis for evidence-based recommendations taking into account both beneficial and harmful effects, systematic errors (bias), random errors (play of chance), and the certainty of the evidence into consideration. The systematic review was performed according to the Cochrane Handbook [76] and was reported according to the PRISMA guideline [116], and was registered in the PROSPERO database (CRD42019128535) prior to commencement of the systematic literature search.

An experienced information specialist from the Copenhagen Trial Unit performed the electronic searches based on a pre-specified search strategy outlined in our protocol [1]. Our systematic search included the following databases: Cochrane Central Register of Controlled Trials (CENTRAL), Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica database (EMBASE), Latin American and Caribbean Health Sciences Literature (LILACS), PsycINFO, Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index—Science (CPCI-S), and Conference Proceedings Citation Index—Social Science &

Humanities (CPCI-SSH). Eligible trials were included regardless of publication type, publication status, and language.

We assessed the following three primary outcomes:

- Quality of life
- Serious adverse events (as defined by the ICH-GCP guidelines) [117]
- Symptom severity (assessed separately for each disorder)

Additionally, we assessed the following three secondary outcomes:

- Suicide or suicide attempts (dichotomous data)
- Self-harm (dichotomous data)
- Level of functioning.

We primarily assessed the outcomes at the end of treatment. For all outcomes, we used the trial results reported at the time point closest to the end of treatment in the long-term treatment group. We secondarily assessed the outcome at maximum follow-up.

We systematically evaluated risk of bias using the original Cochrane risk of bias tool [97], planned to conduct TSA on all outcomes [103-107] (and in case it was not possible, we performed simple power calculations), and followed the eight-step procedure to assess statistical and clinical significance thresholds [118]. The overall quality of the evidence was evaluated using the GRADE approach [119].

Results

On June 27, 2022 our literature search identified a total of 31,669 records after removal of duplicates. We included 19 randomised clinical trials enrolling a total of 3,447 participants randomised to either a shorter-term or a longer-term version of the same psychotherapy type. All trials and outcomes were assessed as high risk of bias.

Three individual trials met the required information size needed to confirm or reject realistic intervention effects. One individual trial showed no evidence of a difference between 6 versus 12 months of dialectical behavioural therapy for borderline personality when assessing quality of life, symptom severity, and level of functioning (the FASTER trial, which will be included later in the IPD of this doctoral dissertation). Another individual trial showed evidence of a beneficial effect of adding booster sessions to 8 and 12 weeks of internet-based cognitive behavioural therapy for depression and anxiety when assessing symptom severity and level of functioning. A third individual trial showed no evidence of a difference between 20 weeks versus 3 years of psychodynamic psychotherapy for mood- or anxiety disorders when assessing symptom severity and level of functioning. It was only possible to conduct two pre-planned meta-analyses. Meta-analysis showed no evidence of a difference between shorter- and longer-term cognitive behavioural therapy for anxiety disorders on anxiety symptoms at the end of treatment (SMD: 0.08; 95% CI: -0.47 to 0.63; $p = 0.77$; $I^2 = 73\%$; four trials; very low certainty). Additionally, meta-analysis showed no evidence of a difference between shorter and longer-term psychodynamic psychotherapy for mood- and anxiety disorders on level of functioning (SMD 0.16; 95% CI -0.08 to 0.40; $p = 0.20$; $I^2 = 21\%$; two trials; very low certainty) [2].

Conclusion

The evidence for shorter versus longer-term psychotherapy for most adult psychiatric disorders is currently unclear. We only identified 19 randomised clinical trials across all adult psychiatric disorders with heterogeneous comparisons in terms of participants, durations, psychotherapy types, and outcomes. More trials at low risk of bias and at low risk of random errors, assessing participants at different levels of psychopathological severity, are needed.

RANDOMISED CLINICAL TRIAL OF SHORT-TERM VERSUS LONG-TERM MENTALISATION-BASED THERAPY FOR BORDERLINE PERSONALITY DISORDER – PRIMARY RESULTS AND LONG-TERM FOLLOW-UP OF THE MBT-RCT

The results of this project are reported in the following dissertation papers: **Paper 3, Paper 4, Paper 6, Paper 7 and Paper 8.**

Methods

We performed a parallel-group, single-centre, randomised clinical superiority trial of short-term (5 months; experimental) versus long-term (14 months) mentalisation-based therapy for adult outpatients with BPD recruited from Stolpegaard Psychotherapy Centre in the Mental Health Services of the Capital Region of Denmark. The trial was pre-registered in clinicaltrials.gov (NCT03677037) prior to randomisation of the first participant, the general methodology was pre-defined in a published protocol [3], and the statistical analyses were pre-defined in a published detailed statistical analysis plan prior to unblinding of the trial results [4].

Adult outpatients (≥ 18 years) with subthreshold or diagnosed BPD were randomly assigned to short-term or long-term mentalisation-based therapy. The primary outcome was BPD symptoms assessed with the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD). Secondary outcomes were functional impairment (assessed with the Work and Social Adjustment Scale; WSAS), quality of life (assessed with the Short Form Health Survey; SF-36), global functioning (assessed with the Global Assessment of Functioning; GAF), and severe self-harm. All outcomes were primarily assessed at 16 months after

randomisation and secondarily assessed at 24 months after randomisation. Our sample size calculation showed that we needed 166 participants to confirm or reject intervention effects on the primary outcome, the ZAN-BPD, with a 3.5-point superiority margin considered as the minimal clinically relevant difference. (MCID).

Short-term mentalisation-based therapy was designed as a 20 weeks therapy programme in closed groups, beginning with five weekly psychoeducative sessions of introduction to mentalisation-based therapy followed by 15 weekly sessions of group mentalisation-based therapy. In addition to group therapy, participants were given individual sessions every second week with one of the two group therapists.

Long-term mentalisation-based therapy was designed as a 14-month programme beginning with a 6-week psychoeducative introduction to mentalisation-based therapy. Subsequently, the participants were allocated to one of eight slow-open mentalisation-based therapy groups. Long-term mentalisation-based therapy was then organised as 12 months of weekly group therapy sessions combined with individual therapy every second week with a third therapist not being one of the group therapists. Both the short-term and the long-term mentalisation-based therapy group and individual therapies were manualized by Bateman and Fonagy [36].

Results

Between October 4, 2018, and December 3, 2020, we randomly assigned 166 participants to short-term ($n = 84$) or long-term mentalisation-based therapy ($n = 82$).

16 months follow-up

After 16 months, regression analyses showed no evidence of a difference between short-term versus long-term mentalisation-based therapy on BPD symptoms (MD 0.99; 95% CI: -1.06 to 3.03; $p = 0.341$), level of functioning (MD 1.44; 95% CI: -1.43 to 4.32; $p = 0.321$), quality of life (MD -0.91; 95% CI: -4.62 to 2.79; $p = 0.626$), global functioning (MD -2.25; 95% CI: -6.70 to 2.20; $p = 0.318$), or severe self-harm (RR 1.37; 95% CI: 0.70–2.84; $p = 0.335$). More participants in the long-term mentalisation-based therapy group tended to experience a serious adverse event compared with short-term mentalisation-based therapy (RR 1.63; 95% CI: 0.94–3.07; $p = 0.088$), primarily driven by a difference in psychiatric hospitalisations (RR 2.03; 95% CI: 0.99–5.09; $p = 0.056$) [6].

24 months follow-up

After 24 months, regression analyses showed no evidence of a difference on BPD symptoms (ZAN-BPD MD -0.56; 95% CI -2.67 to 1.54; $p = 0.598$), level of functioning (WSAS MD -1.42; 95% CI -5.04 to 2.21; $p = 0.440$), global functioning (GAF MD 2.51 95% CI: -1.65 to 6.67; $p = 0.234$), or severe self-harm (RR 1.38; 95% CI 0.88 to 2.21; $p = 0.149$). At 24 months, regression analyses showed evidence of a beneficial effect of long-term mentalisation-based therapy when assessing quality of life (SF-36 MD 5.52; 95% CI 1.72 to 9.32; $p = 0.005$), but the result was potentially biased by an exceptionally large proportion of missing data (40%). We found no evidence of a difference between short-term and long-term mentalisation-based therapy on serious adverse events at 24 months (RR 1.38; 95% CI 0.88 to 2.21; $p = 0.149$). The serious adverse event result had no missing data.

Conclusion

At 16 months after randomisation, long-term mentalisation-based therapy did not lead to lower levels of BPD symptoms, nor did it influence any of the secondary outcomes compared with short-term mentalisation-based therapy for outpatients with borderline personality disorder. At 24 months after randomisation, long-term mentalisation-based therapy did not lead to lower levels of BPD symptoms compared with short-term mentalisation-based therapy. Considering scarce treatment resources for outpatients with BPD in many public health services worldwide, briefer psychotherapy (5 months instead of 14 months) may be regarded as a desirable option. Long-term mentalisation-based therapy led to more serious adverse events, primarily driven by psychiatric hospitalisations, at 16 months after randomisation, but this difference was balanced out at 24 months after randomisation.

Aftermath

After we published the primary results of the MBT-RCT trial (the 16 month time-point), a critique of the trial was published by Philips and colleagues in the same journal, where our original trial was published [120]. Their critique was built on three main arguments: 1) what we defined as “long-term” i.e. 14 months of mentalisation-based therapy, was not equivalent to what they considered long-term. For example, three years of psychotherapy is offered for outpatient personality disorder treatment services in Norway [121]. Second, the participants in our trial did not improve very much from baseline on the primary outcome in both groups as compared to an observational study employed by their group, in which they found greater symptom reductions over the course of three years of outpatient mentalisation-based

therapy. Third, they disagreed with our choice of designing the trial as a superiority trial and suggested that we should have used a non-inferiority design instead.

First, long-term mentalisation-based therapy lasted 14 months in our trial, compared to the original 18 months proposed in the mentalisation-based therapy manual by Bateman & Fonagy [122]. We had to shorten the long-term intervention to comply with local resources at the clinic. While Phillips et al. may argue that MBT is only effective when delivered in this exact 18-month format, or even longer, we find their belief theoretical and unsupported by evidence (at least from randomised clinical trials). Additionally, we argue that if long-term psychotherapy is indeed important, we should observe a relative difference between our short- and long-term programmes.

Second, relying solely on ‘change data’ from baseline to an end-of-treatment time-point for patients receiving a psychological intervention would, in a research setting, correspond to relying on observational evidence. Observational studies (e.g. studies with repeated measurements of a population without a control group) are viewed as having less internal validity, because they are likely to overestimate treatment effects, and because it is difficult to distinguish an effect of the treatment from spontaneous change over time (i.e. regression to the mean) (see this thesis section *Clinical research methodology*) [84, 85]. Hence, comparing observational data across different patients in different health care systems in different countries using different assessment time-points is difficult, as the possible confounding will not be balanced between the populations due to the lack of randomisation.

Third, we deliberately chose a superiority design, which is a two-sided test, allowing us to conclude that short-term mentalisation-based therapy is both 'not better' and 'not worse' than long-term mentalisation-based therapy. When a randomised clinical trial reaches a pre-defined estimated sample size based on an anticipated intervention effect and finds no difference, then the anticipated intervention effect can be rejected. Since we did not find a difference between short-term and long-term mentalisation-based therapy on the primary outcome, the ZAN-BPD, of 3.5 points, which we pre-defined as the anticipated intervention effect, we could reject this effect with a superiority design. Furthermore, employing a non-inferiority design would mean that the effects of the control intervention (i.e. long-term mentalisation-based therapy) were already established, if we were to make a new trial, where the experimental intervention was completely unlikely to be better than the control. However, we do not believe this is the case for long-term mentalisation-based therapy, for which the evidence is still considered of low to very low certainty evidence [40].

INDIVIDUAL PATIENT DATA POOLED ANALYSIS OF TWO TRIALS ASSESSING SHORT-TERM VERSUS LONG-TERM PSYCHOTHERAPY FOR BORDERLINE PERSONALITY DISORDER (BPD-IPD)

The results of this project are reported in the following dissertation papers: **Paper 9** and **Paper 10**. We have previously performed a systematic review of randomised clinical trials assessing different durations of the same psychotherapy type before June 27, 2022 [2]. Although employing different forms of psychotherapy, we identified two comparable trials for borderline personality disorder; the FASTER trial and our own MBT-RCT trial.

Methods

In this project, we pre-defined and subsequently published an individual patient data pooled analysis of short-term versus long-term psychotherapy for BPD with data from the FASTER trial and the MBT-RCT trial [6, 123]. The project was pre-registered in the PROSPERO database before data analysis (CRD42024612840). The design and results of the FASTER and MBT-RCT trials were published previously [3, 6, 123, 124].

The populations in both trials were adult outpatients with BPD. In this pooled analysis of individual patient data, we pruned the short-term groups and the long-term groups of the two trials. Hence, the intervention group was 6 months of dialectical behaviour therapy in the FASTER trial and 5 months of mentalisation-based therapy in the MBT-RCT trial. The control group was 12 months of dialectical behaviour therapy in the FASTER trial and 14 months of mentalisation-based therapy in the MBT-RCT trial.

The primary outcome was a composite of the proportion of participants with a suicide, a suicide attempt, or a psychiatric hospitalisation. The secondary outcome was the proportion of participants with self-harm. Exploratory outcomes were BPD symptoms, symptom distress, level of functioning, and quality of life. We primarily assessed outcomes at 15-months after randomization for the FASTER trial and at 16 months after randomization for the MBT-RCT trial. Predefined subgroups based on the design variables in the original trials were tested for interaction with the intervention as follows: trial, sex (male compared to female), age (below or at 30 years compared to above 30 years), and baseline level of functioning (Global Assessment of Functioning baseline score at 0-49 compared to 50-100).

Results

The primary analysis included 406 participants with 204 assigned to short-term psychotherapy and 202 assigned to long-term psychotherapy. The proportion of participants with either a suicide, a suicide-attempt, or a psychiatric hospitalisation was 11 of 204 participants (5.4%) in the short-term group, and 27 of 202 participants (13.4%) in the long-term group (relative risk, 0.41; 95% (CI): 0.21 to 0.79; P=0.008) primarily driven by a difference in psychiatric hospitalisations. We found no evidence of a difference between short-term and long-term psychotherapy on deliberate self-harm (relative risk 0.8; 95% CI: 0.53 to 1.21, P=0.297), BPD symptoms (mean difference [MD] -2.93; 95% CI: -7.24 to 1.38; P=0.184), symptom distress (MD -0.01; 95% CI:-0.16 to 0.13; P=0.854), level of functioning (MD -0.15; 95% CI:-2.90 to 2.60; P=0.916), and quality of life (MD: 0.45 95% CI: -2.21 to 3.10; P=0.740). We found no evidence of subgroup effects based on the design variables in the original trials or any patient characteristics.

Conclusion

Based on a pooled analysis of individual patient data from two comparable trials, long-term psychotherapy increases the risk of psychiatric hospitalisations compared with short-term psychotherapy for outpatients with BPD at 15-16 months after randomisation. There seems to be no evidence of a difference between short-term and long-term psychotherapy for BPD on any secondary or exploratory outcomes.

PHARMACOLOGICAL INTERVENTIONS FOR CO-OCCURRING PSYCHOPATHOLOGY IN PEOPLE WITH BORDERLINE PERSONALITY DISORDER: A SYSTEMATIC REVIEW

The result of this project is reported in the following dissertation paper: **Paper 11**.

Methods

We conducted a secondary analysis of a recent Cochrane systematic review [46] of all randomised clinical trials comparing pharmacological interventions (antipsychotics, antidepressants, or anticonvulsants) versus placebo in adults meeting diagnostic criteria for BPD. This secondary analysis aimed to assess all identified placebo-controlled trials *post hoc* for reporting of non-BPD specific ('co-occurring') psychopathology (e.g. depression, anxiety, substance use, post-traumatic stress disorder, obsessive-compulsive disorder, schizotypal symptoms, psychotic/dissociative symptoms), and explored treatment effects in subgroups of samples with and without defined co-occurring disorders. We assessed the certainty of evidence using GRADE.

We searched 21 databases and trial registries until February 21, 2022, and subsequently updated the search on March 6, 2024, in Medline, Cochrane Central, EMBASE, PsycINFO, CINAHL, Web of Science, ERIC, LILACS, ProQuest Dissertations and Sociological Abstracts. Meta-analyses were conducted for each pharmacological drug class (antipsychotics, antidepressants, and anticonvulsants) using random-effects models, and effect sizes were calculated using SMDs with 95% CIs. The GRADE framework was employed to assess the certainty of the evidence.

Results

22 trials were available for quantitative analyses. For antipsychotics, we found evidence of a beneficial effect compared with placebo on depressive symptoms (SMD -0.22 , 95% CI -0.42 to -0.01 , $P = 0.04$, 12 trials, $I^2 = 59\%$, very low certainty) and psychotic–dissociative symptoms (SMD -0.28 , 95% CI -0.49 to -0.08 , $I^2 = 49\%$, nine trials, $P = 0.007$, 936 participants, low certainty). For anticonvulsants, we found evidence of a beneficial effect compared with placebo on depressive symptoms (SMD -0.44 , 95% CI -0.80 to -0.08 , $P = 0.02$, six trials, $I^2 = 46\%$, low certainty) and anxious symptoms (SMD -1.11 , 95% CI -1.60 to -0.62 , $I^2 = 24\%$, $P < 0.00001$, three trials, 104 participants, very low certainty). For antidepressants, we found no evidence of a difference compared with placebo on any of the outcomes.

Exploratory subgroup analyses indicated a greater effect of antipsychotics in samples including participants with co-occurring substance use disorders on psychotic–dissociative symptoms ($P = 0.001$).

Conclusion

Our results do not support the use of pharmacological interventions in patients with BPD to target co-occurring psychopathology. Hence, the current evidence does not support differential treatment effects in patients with versus without defined comorbidities. Pharmacological interventions should be used cautiously to target co-occurring psychopathology in patients with BPD. However, the results are based on low to very low certainty evidence alone, and more trials are needed to confirm or reject these results.

SECTION 2 - CLINICAL PAPERS

THE CAPACITY TO END: TERMINATION OF MENTALISATION-BASED THERAPY FOR BORDERLINE PERSONALITY DISORDER

The result of this project is reported in the following dissertation paper: **Paper 5**.

Methods

This is a clinical discussion paper, in which we discuss the clinical challenges associated with terminating mentalisation-based therapy for patients diagnosed with BPD [5]. The paper is an expansion of the theory and application of mentalisation-based therapy for BPD, as termination challenges until recently had only been rather superficially described in the literature on mentalisation-based therapy. The clinical challenges and proposed solutions were developed in the author group, which includes one of the developers of mentalisation-based therapy, Anthony Bateman, through several meetings and discussions both physically and online over a period of two years.

Results

Termination is a critical phase of psychotherapy that has received relatively little attention despite its clinical significance. While termination challenges can occur in any type or duration of psychotherapy, it is often evident in short-term psychotherapies, where time is an important factor. Given the attachment-based difficulties which are typical for patients with BPD [125], terminating therapy can evoke intense emotions, making this phase particularly challenging for both patients and therapists. We proposed a mentalisation-based approach to better understand and manage these challenges, which can be categorised into three main areas: patient factors, therapist factors, and therapeutic relationship factors.

Patient factors influencing termination difficulties include disrupted mentalizing modes such as psychic equivalence mode (where emotions are perceived as absolute reality), teleological mode (where only tangible actions are perceived as meaningful), and pretend mode (where emotions and cognition are disconnected). These modes can manifest as heightened arousal, dependency, or premature detachment during therapy termination. Additionally, epistemic mistrust or vigilance—a reluctance to accept new knowledge as relevant—can lead patients to doubt their ability to function without the therapist's support. Therapist factors also play a crucial role in the termination phase. Therapists may experience countertransference reactions such as feelings of guilt, overprotection, or helplessness when terminating therapy. These reactions can lead to an avoidance of necessary therapeutic confrontations or an undue extension of treatment. The difficulty therapists have in letting go of patients can also be influenced by their own epistemic trust in the patient's capacity for independence.

Therapeutic relationship factors include the emotional bond formed between therapist and patient, which, if not properly mentalized, can create resistance to termination. Overly supportive but insufficiently challenging therapeutic relationships may fail to address core interpersonal issues until termination, at which point they may surface abruptly. Poorly structured treatment plans with unclear goals can further exacerbate termination difficulties. To address these challenges, we proposed several strategies. First, we emphasised the importance of explicitly mentalizing the loss of the therapeutic relationship, validating the patient's emotions while encouraging the exploration of alternative perspectives on separation. Second, we recommended a structured approach to therapy, including regular reminders of treatment boundaries and reflections on session progress. Finally, we

introduced the concept of a termination formulation, an extension of the case formulation that helps patients and therapists collaboratively review progress, address unresolved interpersonal issues, and develop a plan for maintaining therapeutic gains post-treatment.

Conclusion

The approach to termination of mentalisation-based therapy presented in this paper aims to enhance mentalizing capacities surrounding the termination process, thereby reducing attachment-related distress and fostering more secure relational patterns. Although further empirical research is needed, we believe that providing both therapists and patients with a structured framework for termination can significantly improve outcomes in the treatment of BPD.

A CLINICAL ILLUSTRATION OF SHORT-TERM MENTALISATION-BASED THERAPY FOR BORDERLINE PERSONALITY DISORDER

The result of this project is reported in the following dissertation paper: **Paper 12**.

Methods

In this paper, we presented a case study of one patient illustrating the application of short-term mentalisation-based therapy for BPD.

Results

The case study focused on Sandra, a 28-year-old woman with BPD, characterised by affective dysregulation, unstable relationships, and intense abandonment anxiety. Despite multiple past treatment attempts, she struggled with emotional regulation and interpersonal difficulties. The short-term mentalisation-based therapy programme aimed to enhance her mentalizing capacity and improve personality functioning by addressing her difficulties with emotional reactivity and interpersonal conflict.

The structure of short-term mentalisation-based therapy differs from the traditional long-term format in some ways. Short-term mentalisation-based therapy includes five initial psychoeducational sessions (MBT-I), followed by 15 group therapy sessions (MBT-G), supplemented by individual therapy every two weeks. Unlike long-term mentalisation-based therapy, where different therapists are responsible for the individual- and group therapy, the short-term model employs the same therapists for both formats, with the aim of enhancing group cohesion and continuity.

Sandra's treatment followed the standard mentalisation-based therapy framework, beginning with a collaborative case formulation that identified her primary challenges and treatment goals. She initially exhibited avoidance and detachment in group therapy but gradually became more engaged. Over time, she learned to regulate her emotional responses, particularly in situations that triggered her abandonment anxiety. Through a process of guided reflection and therapist intervention, she developed greater autonomy, agency, and a more nuanced understanding of her emotions and relationships.

A critical aspect of Sandra's treatment was managing termination, a phase often challenging for patients with BPD (see Paper 5 of this dissertation). The structured approach included explicit discussions about separation, the use of a termination formulation to review treatment outcomes, and strategies for maintaining gains post-therapy. Despite initial anxiety, Sandra successfully navigated the end of therapy, demonstrating improved emotional regulation and interpersonal functioning.

Conclusion

This individual patient case study suggests that short-term mentalisation-based therapy can be a viable alternative for patients with moderate personality dysfunction. It provides a structured, time-limited intervention that fosters meaningful improvements.

ADAPTING SHORT-TERM MENTALISATION-BASED THERAPY TO ICD-11 PERSONALITY DISORDER IN ADOLESCENTS

The result of this project is reported in the following dissertation paper: **Paper 13.**

Methods

In this conceptual opinion paper, we argue that the short-term mentalisation-based therapy programme, initially designed for adults with BPD and proposed as part of the MBT-RCT trial may offer notable clinical benefits for adolescents with ICD-11 personality disorder at mild to moderate severity levels.

Results

In this paper, we propose that the short-term mentalisation-based therapy programme developed for the MBT-RCT trial could offer several key benefits for treating adolescents with BPD. First, the inclusion of individual therapy throughout the treatment process may be a crucial element in fostering mentalizing, irrespective of the patient's initial level of reflective functioning. For those adolescents with lower reflective functioning at baseline, it could potentially prevent early termination of therapy, while for patients with higher levels of reflective functioning at baseline, the short-term mentalisation-based therapy programme can offer continued support to address sensitive issues with their individual therapist before discussing them in the group setting. Second, having one of the group therapists also act as the individual therapist may contribute to a greater sense of security. This would mean that the adolescent only needs to establish a therapeutic bond with two primary therapists during treatment, and the individual therapist can help create a safe environment within the group.

However, this approach may also have drawbacks, as the adolescent might worry that information shared in individual therapy could be used in the group setting. Third, unlike previous mentalisation-based therapy formats, the short-term version involves closed groups, which may promote group cohesion and a sense of belonging, an aspect that could be especially important for younger patients. Fourth, the shorter duration of treatment could be beneficial from both an administrative and clinical perspective. From an administrative standpoint, more patients can be treated, reducing waiting-lists. Clinically, the shorter timeframe may feel less overwhelming for young people, as five months may seem less intrusive in their daily lives compared to a longer programme.

Conclusion

Given the significant role of social environments and family in adolescent development and personality functioning, we argue that a crucial addition to psychotherapy is to increase the involvement of key individuals in the adolescent's life throughout treatment. Examples of such approaches could be offering family therapy as a co-intervention, as seen in the Rossouw and Fonagy [126] trial on mentalisation-based therapy for self-harming adolescents, or integrating one or more network-based approaches as co-interventions to the short-term MBT programme, as outlined by Bo and colleagues [127] in the Adolescent Mentalisation-Based Integrative Therapy (AMBIT) for hard-to-reach adolescents concept [128].

SECTION 3 - METHODOLOGICAL PAPERS

EXTERNAL VALIDITY OF RANDOMISED CLINICAL TRIALS INCLUDING OUTPATIENTS WITH BORDERLINE PERSONALITY DISORDER

The result of this project is reported in the following dissertation paper: **Paper 14**.

While the assessment of internal validity of randomised clinical trials is often prioritised when evaluating the methodological rigor of trials, the external validity, or generalisability, of these trials also warrants careful scrutiny. One often overlooked factor is the pre-screening phase: which patients are actually offered participation in randomised clinical trials?

Methods

We reviewed the most recent Cochrane systematic review of psychological interventions for BPD [40] and assessed whether trialists report the time from referral to recruitment in trials. Additionally, we used administrative data from the Mental Health Services of the Capital Region of Denmark (2017–2023) investigating the care trajectory of referred BPD patients, who were offered an initial consultation, who attended, and who initiated treatment. This allowed us to estimate the proportion of potentially eligible patients who never reach the stage of outpatient services, where trial inclusion often occurs.

Results

Recruitment to outpatient BPD trials typically occur at the initial clinical consultation or before treatment commencement at the clinic. However, of the 3,751 patients offered an initial consultation for outpatient treatment of borderline personality disorder at Danish

outpatient clinics at the Capital Region of Denmark during the period from 2017-2023, 11% did not attend the initial consultation. Only 65% of the patients offered initial consultation were also offered treatment, while 56% initiated treatment. Consequently, we can assume that between 11% and 44% of potentially eligible patients were not informed about the opportunity to participate in a randomised clinical trial, if such trial was implemented at the clinic. Some of these patients may have been rightfully excluded because of the clinicians' failure to confirm the referral diagnosis. However, as the patients were all offered an initial consultation, we must assume that a large proportion of these patients were potentially eligible.

Conclusion

Our findings highlight potential problems with external validity of outpatient BPD trials, as many outpatient trials may inadvertently recruit a potentially non-representative subgroup of patients, omitting those with potential different clinical characteristics (e.g., instability or impulsivity) who drop-out before being offered participation in a trial.

One proposed solution is to conduct trial screening and consent procedures at the point of referral or immediately after diagnosis confirmation, rather than waiting for treatment initiation. Such adjustments could improve our understanding of the evidence base for outpatient interventions across the full spectrum of clinical presentations of BPD, ultimately supporting more equitable and informed clinical decision-making.

MISSING OUTCOME DATA IN RANDOMISED CLINICAL TRIALS OF PSYCHOLOGICAL INTERVENTIONS: A REVIEW OF PUBLISHED TRIAL REPORTS IN MAJOR PSYCHIATRY JOURNALS

The results of this project are reported in the following dissertation paper: **Paper 15**

Methods

In this study, we conducted a review of reports of randomised clinical trials published in leading psychiatry journals (*The American Journal of Psychiatry*, *JAMA Psychiatry*, *The Lancet Psychiatry*, *The British Journal of Psychiatry*, and *Psychological Medicine*) between 2007-2022. We aimed to assess the extent of missing outcome data in psychological intervention trials for psychiatric disorders. Additionally, we assessed the potential risk of bias introduced by such missingness and explored whether specific outcomes, intervention lengths, diagnostic categories, and psychological intervention types were associated with increased risk of missing outcome data.

Results

We identified a total of 182 published randomised clinical trials. From these trials, we extracted data on 233 primary outcomes and calculated the proportion of missing data for each outcome. Across the 233 primary outcomes, we found that the average proportion of missing data was 18.3% (SD = 17.6%). A substantial majority, 88.4% of the primary outcomes, were assessed as high risk of bias because of missing outcome data.

When assessing missingness stratified by types of outcomes, we found that trials using continuous symptom severity scales (e.g. depression or anxiety scores) had significantly more missing data (mean 18.9%) than those using hard binary outcomes such as hospitalisation, completed suicide, psychiatric hospitalisations, or all-cause mortality (mean 1.8%). Furthermore, when assessing missingness stratified by types of psychiatric diagnoses, we found the highest rates of missingness in trials involving participants with personality disorders, especially BPD. Specifically, for the subgroup of trials targeting BPD, the average proportion of missing outcome data was 33.1%, thus far exceeding the mean across all trials. This result represents nearly twice the overall average and underscores the unique methodological challenges associated with conducting psychotherapy research in BPD populations. Fisher's exact test demonstrated that intervention lengths and psychological intervention types were associated with the proportion of missing outcome data ($p < 0.001$), but there were no clear patterns.

Multiple imputation and regression-based methods were the most frequently used strategies to statistically handle missing outcome data, though some trials used single imputation approaches such as last observation carried forward (LOCF). Lastly, we found that only half of all trials (49.5%) mentioned missing data in the discussion section, even though missing outcome data was a considerable problem in most trials.

Together, these results indicate that missing data remain a pervasive and under-addressed issue in psychotherapy trials, with implications for both internal validity and confidence in intervention effect estimates in psychotherapy trials. In the article, we propose a simple sensitivity analysis that trialists could implement to investigate the mechanism of

missingness in their dataset, i.e. if the missingness is caused by the missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) assumptions. The proposed sensitivity analyses involve creating two additional baseline tables: a baseline table that only includes baseline data from participants with available outcome data, and another that only includes baseline data from participants with missing outcome data. Visual inspection of the two baseline tables may aid the evaluation of whether the two tables seem similar, i.e. if data are MCAR or MAR/MNAR.

Conclusion

Missing outcome data is a considerable problem in randomised clinical trials of psychological interventions, particularly in trials including participants with BPD. The issue of missing outcome data was also a considerable methodological problem in the MBT-RCT trial included in this thesis, approaching 40% at the 24 months follow-up. In the follow-up paper, we employed the method suggested in **Paper 15** and found baseline differences between participants with complete data and missing follow-up data. Hence, we could conclude that the data was likely missing under the MAR/MNAR assumptions.

IMPROVING THE METHODOLOGICAL QUALITY OF RANDOMISED CLINICAL TRIALS ASSESSING PSYCHOTHERAPY FOR BORDERLINE PERSONALITY DISORDER: RECOMMENDATIONS FOR THE FUTURE

The results of this project are reported in the following dissertation paper: **Paper 16**

Methods

In this conceptual article I offer a critical examination of recurrent methodological limitations in randomised clinical trials of psychotherapy for BPD. Additionally, I present some practical recommendations aimed at improving the methodological rigor of future trials. I draw upon prior empirical reviews, including the Cochrane systematic review on psychotherapy for BPD [40] published in 2020, along with my own experience as a researcher in psychotherapy and BPD, to inform these recommendations.

Results

Published randomised clinical trials evaluating the effects of psychotherapies for BPD have generated promising findings [40], but their methodological limitations risk undermining the credibility, comparability, and clinical utility of the evidence base. In the paper, I propose four central methodological themes in psychotherapy trials for BPD, which should be further developed in the future: 1) balancing beneficial and harmful effects; 2) minimising systematic errors (bias); 3) standardisation of participants, interventions, comparators, and outcomes (PICOs); 4) improving conflict of interest disclosures and retrieval of unpublished data.

First, although pharmacological trials are routinely required to monitor and report adverse events, psychotherapy trials have not consistently implemented systematic assessments of potential harms. This discrepancy is noteworthy in trials involving participants with BPD, which is a population with increased risk of psychological distress, suicidality, and potentially adverse reactions to therapeutic processes. Possible explanations for this reporting gap include assumptions about the benign nature of psychotherapy and challenges in defining and measuring psychological harm.

There are several recommendations for how adverse effects should be reported in randomised clinical trials and systematic reviews. The Consolidated Standards of Reporting Trials (CONSORT) [129] and Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) [130] have published extensions with guidance on how to report harms, emphasising the need to evaluate potential adverse effects. Additionally, the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use - Good Clinical Practice (ICH-GCP) encourages specific definitions when collecting and reporting adverse effects in clinical trials [131, 132]. These include adverse events, adverse drug reactions, serious adverse events, serious adverse drug reactions, and suspected unexpected serious adverse reactions [131, 132]. A serious adverse event is defined by ICH-GCP as “any untoward medical occurrence that results in death, is life-threatening, requires inpatient hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect” [131, 132]. In this paper, I propose that future trials should implement the ICH-GCP definition of a serious adverse event and assess these systematically by blinded research personnel, for example, through patient medical records. Implementing such

procedures would not only align psychotherapy trials with broader biomedical research standards but also provide essential information for weighing beneficial and harmful effects in psychotherapy trials. This procedure was implemented in the MBT-RCT trial presented in this dissertation.

Second, the paper highlights some common bias risks in psychotherapy trials for BPD and discusses possible ways to mitigate these risks. The most biased domains in previous trials are bias associated with blinding of outcome assessors, incomplete outcome data, selective outcome reporting, and other bias (e.g., allegiance bias). Due to the interpersonal nature of psychotherapy, full blinding of participants and therapists is extremely difficult [133]. However, I argue that outcome assessors and data analysts can and should be blinded whenever possible. Moreover, the use of objective hard outcomes, such as hospitalisation or self-harm events drawn from registries by blinded personnel, can supplement a more subjective self-report tool. Additionally, participant dropout from the trial is another critical bias risk, particularly pronounced in BPD trials (see Paper 15 in this doctoral dissertation). High risk of missing outcome data not only reduces statistical power but also increases the risk of biased estimates, especially when data are not missing completely at random. I call for the adoption of proactive strategies to maximise data acquisition (again hard outcomes can be a desirable and patient-important solution), and I propose the use of statistical methods that appropriately handle incomplete data, including multiple imputation and sensitivity analyses.

Third, the paper highlights that heterogeneity across trials in terms of participants, interventions, comparators, and outcomes (the PICO framework) poses challenges for

evidence synthesis and limits the generalisability of findings. For example, BPD trials differ in diagnostic inclusion criteria, comorbidity thresholds, interventions, comparators, and primary outcome selection. This variability complicates meta-analytic integration and increases heterogeneity.

To promote comparability, the article proposes consensus-driven standardisation in trial design. This could include common inclusion criteria based on standardised diagnostic interviews, detailed intervention manuals to ensure replicability, and consensus-based core outcome sets, for example following the COMET Initiative [134].

Fourth, another source of potential bias is allegiance bias, which occurs when researchers' personal investment in a particular psychotherapy influences trial conduct or interpretation. This is particularly salient in psychotherapy research, where investigators are often also clinicians or treatment developers. Such bias can influence multiple aspects of a trial—from how it is designed and who delivers the therapy to how findings are framed and interpreted. The article urges transparent and detailed disclosure of all potential conflicts of interest, including intellectual and financial ties to the intervention under investigation. I encourage journals and funding bodies to require such declarations and recommends the inclusion of independent investigators—uninvolved with the therapy's development—on study teams or in data interpretation processes. These safeguards can mitigate problems associated with allegiance bias and could improve the credibility of trial results.

Another aspect of allegiance bias is the potential for publication bias, where negative or inconclusive results sometimes go unpublished, leading to potentially distorted meta-

analytic results. Systematic reviews relying only on published trials therefore risk biased conclusions [135]. In the paper, I recommend that systematic review authors make considerable efforts to find unpublished trial data. This could involve reviewing clinical trial registries, contacting principal investigators, and examining grey literature sources. Furthermore, preregistration of trial protocols, complete with predefined outcomes and statistical analysis plans, should be mandated and enforced. These measures would increase transparency, discourage selective reporting, and ensure that the full spectrum of available evidence is considered in systematic reviews.

Conclusion

While we have already advanced substantially in both designing and implementing structured, manualised psychotherapies for BPD, and in assessing their effects in randomised clinical trials, the current evidence is still restricted by substantial methodological limitations. Improving these methodological limitations can result in potentially identifying more evidence-based psychotherapies or strengthening the evidence-base for existing ones, which may ultimately result in better care for patients with BPD.

DISCUSSION

The heart of the discipline of evidence-based medicine lies in the randomised clinical trial design, with systematic reviews of these trials regarded as the highest level of evidence assessing the beneficial and harmful effects of any healthcare intervention [70, 76, 87]. Finding the optimal duration of psychotherapy is a clinically important question that should be answered using such research designs. Patients with BPD have historically been treated with long-term psychotherapy, often lasting more than a year in clinical practice, but the evidence for the optimal duration of psychotherapy is unclear. The papers presented in this doctoral dissertation have investigated the optimal duration of psychotherapy for BPD through several empirical, clinical, and methodological papers.

In this final part of the doctoral dissertation, I will discuss the strengths and limitations of the findings, whether psychotherapy research can be integrated into the evidence-based medicine paradigm, how to select the optimal psychotherapy duration for the individual patient, a stepped-care approach to psychotherapy for BPD as a potential next step forward, and the challenges and future directions in outcome selection and measurement in BPD trials.

Strengths and limitations of the findings

The strengths and limitations of each paper included in this thesis have been described previously (see papers 1-16 in the appendix). In general, the empirical, clinical, and methodological papers have some strengths and limitations.

The strengths of the empirical papers includes their strict methodologies following the recommended guidelines for such research, including the CONSORT recommendations [136], Cochrane Handbook [137], the PRISMA guideline [116], and GRADE assessments [119], a five-step procedure for assessing statistical and clinical significance in randomised clinical trials [138], and an eight-step procedure for systematic reviews [118]. Hence, the empirical papers both consider the risk of systematic errors, the risk of random errors, and the certainty of evidence when evaluating the duration effects of psychotherapy for BPD and the role of co-interventions such as pharmacology. The empirical papers also have limitations. Most notably, the papers are limited by relying on evidence from trials at relatively small sample sizes. This accounts for both the trials included in the systematic reviews as well as the MBT-RCT trial presented in the doctoral dissertation. There is a continuing debate about which evidence is the most reliable: the pooled meta-analysis results of smaller trials from a systematic review or the results of one large, randomised trial [78-81]. Large randomised clinical trials are often pragmatically defined as trials randomising more than 1000 participants [139-141]. This definition is convenient for pragmatic reasons, but even trials of more moderate sample sizes may be large enough if they have sufficient power to detect or reject a postulated treatment effect [140]. The MBT-RCT trial was indeed powered to assess the postulated intervention effect, but this sample size estimation was based on an unclear MCID on the primary outcome, and had large proportions of missing data, which again questions whether the trial was in fact adequately powered. Furthermore, the trial was not powered to assess hard outcomes such as serious adverse events, and while the added value of combining the data with the FASTER trial in the pooled individual patient data analysis was profound, there is still a need for more trials comparing different durations of the same psychotherapy type for BPD to reach the required

information size to confirm or reject realistic intervention effects. Additionally, it is a limitation that the long-term intervention in the MBT-RCT trial was only 14 months of psychotherapy with individual therapy only given every second week, while the original evidence-based long-term intervention is 18 months of psychotherapy with weekly individual therapy. We have discussed this limitation more extensively in Paper 8 [8].

Furthermore, in the review of shorter versus longer-term psychotherapy for adult psychiatric disorders in the review on pharmacological interventions, a consistent limitation across the included trials was the overall low methodological quality. Many trials were assessed at high risk of bias due to lack of blinding, incomplete outcome data, small sample sizes, and selective reporting. As reflected in the low to very low GRADE ratings, such methodological shortcomings undermine the strength and interpretability of the evidence.

The clinical papers in the present dissertation have the following strengths and limitations. They were all developed based on extensive clinical experience in the author groups with both short-term and long-term mentalisation-based therapy for BPD, and they were all developed in close collaborations with leading experts in the field. However, as with any such theoretical opinion-based articles, they must be considered of potentially great clinical value, but they are methodologically placed at the bottom of the evidence hierarchy.

The methodological papers in the present dissertation have the following strengths and limitations. They were all developed based on my increasing experience with being both a trialist and a systematic review author within the clinical field of personality disorders, and the methodological topics and recommendations were selected based on this experience.

This is a strength, as my experience as a trialist and systematic review author allows me to have experienced first-hand the struggles of conducting clinical research within the BPD field. However, the methodological papers are also limited by being selective in their focus, as (unfortunately) more methodological problems could be highlighted, for example, the problem with publication bias or allegiance bias in psychotherapy trials, and the potential problem with selective outcome reporting in trials assessing psychotherapy for BPD [40, 142].

Psychotherapy in the era of evidence-based medicine: friends or foes?

Psychotherapy, regardless of the theoretical orientation, involves a theoretical framework that is applied by one or more clinicians often aiming to tailor the intervention to the specific patient (to various extents) while still adhering to a set of standardized criteria. Hence, psychotherapy can be described as a ‘complex intervention’, which are health interventions that have many ‘active ingredients’ [143]. A complex intervention combines different components in a whole that is more than the sum of its parts [144]. The UK Medical Research Council (MRC) suggests that the complexity resides in the number of interacting components, the number and difficulty of behaviours required by those delivering or receiving the intervention, the number of groups or organizational levels targeted by the intervention, the number and variability of outcomes, and the degree of flexibility or tailoring of the intervention permitted [145]. Thus, complex interventions can be defined as non-standard while still conforming to specific, context-dependent, theory driven processes [144].

Regardless of the complexity of a psychotherapeutic intervention, randomised clinical trials are considered the cornerstone of evidence-based medicine when the beneficial and harmful effects of new interventions should be assessed [73, 76]. However, because of their multifaceted nature and dependence on social and cultural context, complex interventions such as psychotherapy pose methodological challenges to the design, conduct, analysis, and interpretation of randomised clinical trials. Examples of such methodological challenges include: how to implement standardisation and assess treatment fidelity without sacrificing the possibility of adaptation of the psychotherapy to the needs of the individual patient, staffing issues, including unique patient/therapist interactions, and controlling for contextual characteristics to increase generalisability [146]. Still, despite the significant amount of methodological challenges in complex intervention trials, there is still a broad consensus among clinical trial specialists, that the randomised clinical trial is the most optimal design to assess intervention effects for complex intervention trials [76]. However, the opposite perspective also exists within psychotherapy research. Some argue that psychotherapy can only survive within the realm of evidence-based medicine by formulating its own criteria to assess efficacy. These could, for example, involve demanding that psychotherapy should be studied in naturalistic settings using psychotherapy treatments of a variable duration that are chosen by patients who have multiple problems, and are carried out with more flexibility, allowing the therapists to exercise more clinical judgement [147]. Ultimately, results of such studies would help establish ‘practice-based evidence’, contrary to ‘evidence-based practice’ [148]. Additionally, some argue that long-term psychotherapy for severe and complex mental disorders such as personality disorders are especially difficult to study with a randomised clinical trial design, as these patients are prone to early treatment discontinuation, and because it is difficult to adhere to treatment

manuals of the interventions and co-interventions over a long course of time [149]. Additionally, the choice of control intervention is especially challenging in psychotherapy trials, and TAU control groups often suffer the problems that they become too similar to the experimental intervention [150].

From a patient perspective, however, it seems reasonable to let the treatment of severe psychiatric disorders be guided by the same strict methodological criteria as for any other medical specialty. As discussed in Papers 14-16, some of the methodological challenges faced in psychotherapy research for BPD can easily be mitigated by relatively simple changes in trial conduct. Furthermore, it is important that psychotherapy trials are implemented in routine clinical settings, ensuring the external validity of trial results. Hence, if clinicians have difficulties with adhering to treatment manuals over long courses of time – or on the contrary become increasingly better at implementing such manuals over the course of time – these fluctuations must be considered a naturalistic part of real-life clinical practice, which the randomised clinical trial should ideally capture. Only by implementing trials in real-world clinical settings without too much involvement from the researchers, it becomes possible to assess the true effects of psychotherapy.

Selecting the optimal psychotherapy duration for the individual patient

As mentioned previously, applying an evidence-based medicine approach to health care is a three-legged stool involving: 1) research evidence; 2) clinical experience; and 3) values, preferences, and characteristics of service users [70, 72]. Hence, applying an evidence-based medicine approach to selecting the optimal duration of psychotherapy for the individual BPD patient, all three legs of the stool must be considered.

Regarding research evidence, the ‘best available evidence’ has been presented in this doctoral dissertation as it has included evidence from the top of the evidence hierarchy (results from systematic reviews and randomised clinical trials), showing that there seem to be no evidence of a difference between short-term versus long-term psychotherapy (either mentalisation-based therapy or dialectical behavioural therapy) for outpatients with BPD when assessing self-harm, borderline symptoms, quality of life, level of functioning, and symptom distress, but long-term psychotherapy may increase the risk of psychiatric hospitalisations in the short run.

Regarding clinical experience, clinicians must consider their clinical judgement when selecting the optimal duration of psychotherapy for the individual patient. For example, patients presenting with severe levels of personality functioning or who exhibit great attachment insecurities, including severe abandonment anxiety and associated self-harming behaviours may seem difficult to treat in a short-term format [60]. Additionally, the termination phase of psychotherapy may be particularly challenging in such clinical cases. However, clinicians’ experiences sometimes vary substantially, and they must be mindful that given that short-term psychotherapy is relatively new, many clinicians (particularly the experienced ones) may have extensive experience with practicing long-term psychotherapy for BPD, while still being relatively inexperienced in deliberately shortening the treatment [151]. Furthermore, there is an inherent selection bias when relying on clinical experience in decision making, as clinicians naturally only recall patients who attended psychotherapy. Hence, it is likely that patients who prematurely dropped out of treatment are less represented in clinicians’ memories, and that such patients would have been sufficiently

helped within a short-term format, as this would potentially have reduced their risk of drop-out. Additionally, as the patient-clinician contact ends with the termination of therapy, clinicians may recall patients as responders, even if they later deteriorate after therapy has ended. As a result, there is a risk that relying on clinical experience may result in an overly reliance on long-term psychotherapy when recommending psychotherapy durations for BPD patients.

Regarding patient values, preferences and characteristics, such characteristics should also be considered when selecting the optimal psychotherapy duration for the individual patient. However, in public mental health services, it is not always possible to accommodate patient preferences to e.g. receive long-term psychotherapy, and until now, there are no distinct patient characteristics that necessarily warrant short-term or long-term psychotherapy. However, many clinicians and researchers believe that patients with more severe personality disorders should be offered longer-term psychotherapy interventions [60].

Psychotherapy for borderline personality disorder – towards a stepped care approach?

As noted, evidence-based psychotherapies for BPD last for one year or longer and are resource intensive [40]. As shown in the research presented in this doctoral dissertation, briefer alternatives to existing long-term evidence-based interventions could be considered equally effective. While we did not find evidence for any significant subgroup effects of patient characteristics with differential treatment outcomes of short-term versus long-term psychotherapy, most clinicians' clinical experiences is that there are still some patients who may benefit considerably from briefer psychotherapy interventions and may not need

additional treatment, while there are others, who still need additional treatment [60]. Hence, a step forward in routine clinical practice for outpatients with BPD could be to introduce briefer options to patients as part of a stepped care programme as an alternative to routine extended treatment [152, 153]. With stepped care, most BPD patients would receive a briefer treatment programme from the beginning, with longer psychotherapies reserved for those who do not respond. Implementing stepped care approaches to psychotherapy would make treatment cheaper and more accessible for patients [152, 154]. Not all patients with BPD can be treated effectively with brief psychotherapy, but a stepped-care model allows those with less severe symptomatology to be managed with fewer resources, freeing up more time and clinical personnel for the treatment of those patients with higher treatment demands. Of course the judgement of which patients' treatment to extend is not simple, particularly in light of often very difficult termination challenges facing clinicians at the final stages of psychotherapy [5]. Such clinical decisions warrant careful considerations from clinical staff and administrations, and one practical solution would be to free up access to extended treatment for a certain percentage of patients. Following the results from the MBT-RCT trial, such a stepped care approach has already been implemented at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, in which approximately 10% of patients from the start are offered extended psychotherapy in the mentalisation-based therapy programme. Whether this solution is viable and evidence-based is still unknown, as it would require a randomised clinical trial randomising participants to stepped-care versus traditional long-term psychotherapy, and such trial is indeed required in the future to move the field forward.

Outcome selection and measurement in borderline personality disorder trials

Excessive reliance on symptom scales

Randomised clinical trials including participants with BPD, like most other psychiatric disorders, often rely on symptom scales when assessing the effects of psychotherapy interventions [40]. Although symptom reduction is undeniably important to patients, there are some methodological problems with the excessive reliance on such symptom scales. For example, an emphasis on symptom reduction may inadvertently overlook other patient important outcomes, such as functional improvement or increased quality of life. Furthermore, when utilising symptom scales as outcome measures, it is essential to establish a quantifiable MCID. The MCID is the smallest value benefit to patients and is a patient-centred concept [155, 156]. However, it remains unclear whether any method currently suggested can produce an estimate that corresponds to real-world benefits.

The most frequently used methods to determine MCID - anchor-based and distribution-based approaches - have significant drawbacks. Anchor-based methods rely on comparisons to other scales but often involve extrapolation, making them vulnerable to bias. Distribution-based methods are statistical and do not necessarily reflect patient experiences or preferences. To this date, it can be argued that no symptom scale exists with a universally accepted, robust MCID, not even the Hamilton Depression Rating Scale (HDRS), which is the gold standard when assessing depression symptoms [75].

Beyond the methodological problems with defining MCIDs, symptom scales are also prone to biases stemming from a lack of blinding [133] and high rates of missing data, as previously documented in this doctoral dissertation. Trials including participants with

personality disorders, especially BPD, are particularly vulnerable, with average missing data rates of 33.1% [12]. These findings call for targeted methodological improvements in trial design, conduct, and reporting to strengthen the validity and clinical interpretability of psychotherapy research in BPD. Unless we address the pervasive issue of missing outcome data, the evidence base for psychotherapy for BPD risks being undermined by systematic error, causing reduced certainty of the accumulated body of evidence.

Heterogeneity in outcome selection across trials

There is large heterogeneity in outcome selection across BPD trials [40]. Even when borderline symptomatology is used as an outcome, several different scales exist and are used across trials. In the two trials included in the pooled analysis of individual patient data included in this doctoral dissertation assessing short-term versus long-term psychotherapy for BPD, we also observed heterogeneous outcome selection and measurement, particularly on the continuous outcomes. For example, borderline symptoms were assessed using the ZAN-BPD interview in the MBT-RCT trial and with the BSL-23 self-report scale in the FASTER trial. It is possible to perform aggregate data meta-analysis and individual patient data pooled analyses on outcomes assessed using different scales. Nevertheless, when heterogeneous scales are used across trials, it threatens the synthesis of data on a meta-analytical level.

In cases of heterogeneous outcome measurements, many meta-analyses summarise the outcomes using the ‘standardised mean difference’ (SMD). The SMD is a statistical concept that enables standardisation of the results to a uniform scale by dividing the mean difference between two groups by the pooled standard deviation of the measurement [76]. Nonetheless,

it is crucial to acknowledge the significant methodological limitations of the SMD [157]; the SMD is merely a statistical concept, making its results challenging to interpret clinically and convey to non-researchers [157]. Furthermore, it is not possible to assess the risk of random errors with Trial Sequential Analysis for a given meta-analytic outcome, if the outcome has been analysed using SMD. Hence, future research should focus on establishing and utilising core outcome sets (COS), which are minimum lists of patient-important outcomes for a specific patient group.

The Core Outcome Measures in Effectiveness Trials (COMET) initiative promotes creating and implementing core outcome sets [158] using advanced Delphi processes, including stakeholders such as patients, relatives, clinicians, researchers, administrators, and funders. One core outcome set already exists for ICD-11 personality disorder, but this core outcome set is only relevant for routine outcome monitoring in clinical practice, and thus, has not been developed with the intention to inform outcome selection in randomised clinical trials [159].

Hard outcomes

A possible future direction concerning outcome selection and measurement in BPD trials is to incorporate hard (often binary) patient-important outcomes, preferably as primary outcomes. Hard outcomes are conclusive regarding the disease progression, demonstrating a patient's feelings, functionality, or survival (e.g. suicide or suicide attempts, hospitalisations, employment status, convictions, on social welfare or not) [160]. Gathering data on hard outcome results is often simpler, and in certain countries, data acquisition (such as vital status) from registries can be done without any risk of missing information.

If hard outcomes are used as primary outcomes, more randomised participants will be needed before a psychotherapy type or duration can be deemed ‘evidence-based’, as sample size estimations using hard outcomes often result in larger required sample sizes to detect or reject realistic intervention effects [75]. Some may argue that such research is impractical due to unattainable sample sizes. However, such arguments are not rooted in scientific doubt about the validity of the methodology but rather in unsupported pragmatism. We have previously demonstrated that sample sizes calculated based on hard outcomes in psychiatric research are indeed possible, making results from psychiatric trials more comparable to somatic research, in which hard outcomes such as all-cause mortality are frequently used [75]. For example, the use of composite outcomes as used in the IPD paper of this doctoral dissertation can decrease the required sample size.

To ensure that patients receive psychotherapy in a duration that offers more benefits than harm (or the same benefits as long-term psychotherapy), we must employ valid research methods and continuously strive to enhance these methodologies, even if it means larger sample sizes and thus, more expensive research.

In the MBT-RCT trial, we assessed safety with the ICH-GCP definition of serious adverse events, which was assessed blinded by two independent investigators, who systematically went through the patient medical records of all randomised participants, hence, this result did not have any missing data. Later, in the individual patient data pooled analysis of two trials assessing short-term versus long-term psychotherapy for BPD, we decided to implement a hard outcome as the primary outcome due to the advantages presented here.

We selected some “disease-specific” serious adverse events and pruned them to a composite outcome assessing the proportion of participants with either a suicide, a suicide-attempts, or a psychiatric hospitalisation. By doing so, we could reduce the bias risk associated with the primary outcome (by reducing missing outcome data and increasing blinding), while still maintaining the primary focus of investigating the optimal duration of psychotherapy for BPD from a patient-important, evidence-based perspective.

The complexity of outcome selection in a heterogeneous group of patients

As previously noted, BPD is a heterogeneous group of psychiatric patients [26, 27, 161]. They can range from having mild dysfunction to severe symptomatic and functional impairment [26, 27, 161]. The heterogeneity of BPD was also present in our pooled analysis of individual patient data included in this doctoral dissertation, in which participants from the MBT-RCT trial had low proportions of suicidal behaviours at baseline compared with the FASTER trial, in which participants were only included if they had exhibited recent and chronic self-harming behaviours, including at least one episode in the past two months. Hence, it can be discussed whether different levels of disease severity call for different target outcomes. For example, suicide and suicide-attempts are unquestionably important outcomes for high-risk populations with high suicidal risk profiles, yet it could be argued that alternative outcomes, e.g., employment status, would be more fitting for less severe BPD populations such as the one included in the MBT-RCT trial. Perhaps the future implementation of the ICD-11 diagnostic system for personality disorder ranging from mild to severe personality functioning can be an important step towards selecting patient-important outcomes for each severity level of the personality disorder spectrum.

Most diseases (both psychiatric and somatic) have varying degrees of severity. In cases where a disease is on the border between being classified as severe or not, randomised clinical trials can provide clear evidence on whether an intervention is indicated, because both ‘severe’ and ‘non-severe’ patients will be equally distributed between the groups due to randomisation. Randomised clinical trials at low risk of bias, at low risk of random errors, and at low risk of design errors can via prospectively planned subgroup analyses suggest such indications [96, 109]. Due to their exploratory nature, subgroup analyses indicating an effect in one or more of the preplanned subgroups, can be used to inform new confirmatory randomised clinical trials on these subgroups [58].

CONCLUSION

The evidence for the optimal duration of psychotherapy for most adult psychiatric disorders is currently unclear due to sparse data from randomised clinical trials. For BPD, two trials currently exist assessing a short-term versus a long-term duration of the same psychotherapy type. We found no evidence of a difference between short-term (20 weeks) versus long-term (14 months) mentalisation-based therapy for outpatients with BPD either at 16 months or 24 months after randomisation. Long-term mentalisation-based therapy may lead to more serious adverse events at 16 months after randomisation, primarily driven by psychiatric hospitalisations. When pooling the data with the second trial, comparing short-term (six months) versus long-term (twelve months) dialectical behavioural therapy for BPD, we found the same results. Hence, there seems to be no evidence of a difference between short-term versus long-term psychotherapy (either mentalisation-based therapy or dialectical behavioural therapy) for outpatients with BPD when assessing borderline symptoms, level of functioning, self-harm, and quality of life, but long-term psychotherapy may increase the risk of psychiatric hospitalisations. More large trials powered to assess patient-important hard outcomes at low risk of systematic and random errors, assessing participants at different levels of ages and personality disorder severity, are needed to move the field forward. A stepped-care approach to psychotherapy may be a viable option in clinical practice to ensure that more patients get access to evidence-based care.

REFERENCES

1. Juul S, Poulsen S, Lunn S, Sorensen P, Jakobsen JC, Simonsen S. Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Syst Rev*. 2019;8(1):169.
2. Juul S, Jakobsen JC, Jørgensen CK, Poulsen S, Sørensen P, Simonsen S. The difference between shorter- versus longer-term psychotherapy for adult mental health disorders: a systematic review with meta-analysis. *BMC Psychiatry*. 2023;23(1):438.
3. Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196.
4. Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, 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;22(1):497.
5. Juul S, Simonsen S, Bateman A. The capacity to end: termination of mentalization-based therapy for borderline personality disorder. *J Contemp Psychother*. 2020;50:331-8.
6. Juul S, Jakobsen JC, Hestbaek E, Jørgensen CK, Olsen MH, Rishede M, et al. Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder: A Randomized Clinical Trial (MBT-RCT). *Psychotherapy and Psychosomatics*. 2023;92(5):329-39.
7. Juul S, Jakobsen JC, Hestbaek E, Kamp CB, Olsen MH, Rishede M, et al. Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder:

24 Months Follow-Up of a Randomized Clinical Trial. *Psychotherapy and Psychosomatics*. 2025; DOI: 10.1159/000544934

8. Juul S, Jakobsen JC, Hestbaek E, Kamp CB, Olsen MH, Rishede M, et al. Response to Philips et al.: Several Fundamental Misconceptions about Evidence-Based Practice. *Psychotherapy and Psychosomatics*. 2024;93(5):348-9.

9. Ribeiro JP, Juul S, Kongerslev MT, Jørgensen MS, Völlm BA, Edemann-Callesen H, et al. Pharmacological interventions for co-occurring psychopathology in people with borderline personality disorder: secondary analysis of the Cochrane systematic review with meta-analyses. *British Journal of Psychiatry*. 2025; 226(4):226-237

10. Juul S, Frandsen FW, Bo Hansen S, Sørensen P, Bateman A, Simonsen S. A clinical illustration of short-term mentalization-based therapy for borderline personality disorder. *J Clin Psychol*. 2022; 78(8):1567-1578

11. Simonsen S, Hestbæk E, Juul S. Adapting Short-Term Mentalization-Based Therapy to ICD-11 Personality Disorder in Adolescents. *Children*. 2023;10(1):93.

12. Juul S, Faltermeier P, Petersen JJ, Olsen MH, Andersen RK, Kamp CB, et al. Missing outcome data in randomised clinical trials of psychological interventions: a review of published trial reports in major psychiatry journals. *BMC Psychiatry*. 2024;24:798.

13. Juul S. Improving the methodological quality of randomized clinical trials assessing psychotherapy for borderline personality disorder: Recommendations for the future. *Front Psychiatry*. 2022;13:1053844.

14. Juul S. PhD thesis. Which length for whom: A systematic review, meta-analysis, and a randomised clinical trial of short-term versus long-term psychoterapy for adult psychiatric disorders. University of Copenhagen. [Defended April 20, 2021]

15. Stern A. Psychoanalytic investigation of and therapy in the borderline group of neuroses. *The Psychoanalytic Quarterly*. 1938;7(467-489).
16. Kernberg O. Borderline personality organization. *J Am Psychoanal Ass*. 1967;15(3):641-85.
17. Gunderson JG, Kolb JE. Discriminating features of borderline patients. *Am J Psychiatry*. 1978;135(7):792-6.
18. Gunderson JG, Weinberg I, Choi-Kain L. Borderline Personality Disorder. *Focus*. 2013;11(2):129-45.
19. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub; 2013. Report No.: 0890425574.
20. Zanarini MC, Frankenburg FR, Dubo ED, Sickel AE, Trikha A, Levin A, et al. Axis I comorbidity of borderline personality disorder. *Am J Psychiatry*. 1998;155:1733-9.
21. Leichsenring F, Leibing E, Kruse J, New AS, Leweke F. Borderline personality disorder. *Lancet*. 2011;377(9759):74-84.
22. Torgersen J. Prevalence, Sociodemographics, and Functional Impairment. In: Oldham J, Skodol AE, Bender DS, editors. *Textbook of personality disorders*. 2. edition ed. Washington DC: American Psychiatric Publishing; 2014. p. 109-31.
23. Gunderson JG, Herpertz SC, Skodol AE, Torgersen S, Zanarini MC. Borderline personality disorder. *Nature Reviews Disease Primers*. 2018;4(1):18029.
24. Michonski JD. The Underlying Factor Structure of DSM criteria in Youth BPD. In: Sharp C, Tackett JL, editors. *Handbook of Borderline Personality Disorder in Children and Adolescents*. New York, NY: Springer New York; 2014. p. 35-48.
25. Videler AC, Hutsebaut J, Schulkens JE, Sobczak S, Van Alphen SP. A life span perspective on borderline personality disorder. *Curr Psychiatry Reports*. 2019;21:1-8.

26. Cavelti M, Lerch S, Ghinea D, Fischer-Waldschmidt G, Resch F, Koenig J, et al. Heterogeneity of borderline personality disorder symptoms in help-seeking adolescents. *Borderline Personality Disorder and Emotion Dysregulation*. 2021;8:1-14.
27. Asnaani A, Chelminski I, Young D, Zimmerman M. Heterogeneity of borderline personality disorder: do the number of criteria met make a difference? *J Personality Dis*. 2007;21(6):615-25.
28. Jørgensen MS, Møller L, Bo S, Kongerslev M, Hastrup LH, Chanen A, et al. The course of borderline personality disorder from adolescence to early adulthood: a 5-year follow-up study. *Compr Psychiatry*. 2024;132:152478.
29. World Health Organization (WHO). *International Classification of Diseases - Eleventh Revision (ICD-11)*. 2019/2021.
30. Bach B, Kramer U, Doering S, di Giacomo E, Hutsebaut J, Kaera A, et al. The ICD-11 classification of personality disorders: a European perspective on challenges and opportunities. *Borderline Pers. Disord. Emot. Dysregulation*. 2022;9(1):12.
31. Ellison WD. Psychotherapy for Borderline Personality Disorder: Does the Type of Treatment Make a Difference? *Current Treatment Options in Psychiatry*. 2020;7(3):416-28.
32. Linehan M. *Cognitive-behavioral treatment of borderline personality disorder*: Guilford Press; 1993.
33. Lynch TR, Chapman AL, Rosenthal MZ, Kuo JR, Linehan MM. Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *J Clin Psychol*. 2006;62(4):459-80.
34. Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355-64.

35. Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563-9.
36. Bateman, Fonagy P. Mentalization-based treatment for personality disorders: a practical guide: Oxford University Press; 2016.
37. Bateman A, Bales D, MBT. HJQMf. A Quality Manual for MBT. 2020: Available at: <https://www.annafreud.org/media/14278/adherence-scale-august-2020.pdf> [Accessed January 10, 2025].
38. Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385(9969):735-43.
39. Kramer U. Case formulation for personality disorders: Tailoring psychotherapy to the individual client: Academic Press; 2019.
40. Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020(5).
41. National Institute for Health and Care Excellence. Borderline personality disorder: recognition and management [CG78]. 2009: Available at: <https://www.nice.org.uk/guidance/cg78/resources/2018-surveillance-of-personality-disorders-nice-guidelines-cg77-and-cg78-4906490080/chapter/Surveillance-decision?tab=evidence>. [Accessed June 10, 2025]
42. Bridler R, Häberle A, Müller ST, Cattapan K, Grohmann R, Toto S, et al. Psychopharmacological treatment of 2195 in-patients with borderline personality disorder: a comparison with other psychiatric disorders. *Eur Neuropsychopharmacol*. 2015;25(6):763-72.

43. Paris J. Why patients with severe personality disorders are overmedicated. *J Clin Psychiatry*. 2015;76(4):22479.
44. Stoffers-Winterling J, Storebø OJ, Lieb K. Pharmacotherapy for borderline personality disorder: an update of published, unpublished and ongoing studies. *Curr Psychiatry Rep*. 2020;22(8):1-10.
45. Lieb K, Vollm B, Rucker G, Timmer A, Stoffers JM. Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. *Br J Psychiatry*. 2010;196(1):4-12.
46. Stoffers-Winterling JM, Ribeiro JP, Kongerslev MT, Völlm BA, Mattivi JT, Faltinsen E, et al. Pharmacological interventions for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*. 2022(11).
47. Stoffers J, Völlm BA, Rucker G, Timmer A, Huband N, Lieb K. Pharmacological interventions for borderline personality disorder. *Cochrane Database of Syst Rev*. 2010(6):1465-858.
48. Stoffers-Winterling J, Storebo OJ, Lieb K. Pharmacotherapy for borderline personality disorder: an update of published, unpublished and ongoing studies. *Curr Psychiatry Rep*. 2020;22(8):37.
49. Leichsenring F, Rabung S. Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis. *Br J Psychiatry*. 2011;199(1):15-22.
50. Linehan MM, Comtois KA, Murray AM, Brown MZ, Gallop RJ, Heard HL, et al. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry*. 2006;63(7):757-66.

51. Giesen-Bloo J, van Dyck R, Spinhoven P, van Tilburg W, Dirksen C, van Asselt T, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Arch Gen Psychiatry*. 2006;63(6):649-58.
52. Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355-64.
53. Kvarstein EH, Pedersen G, Folmo E, Urnes Ø, Johansen MS, Hummelen B, et al. Mentalization-based treatment or psychodynamic treatment programmes for patients with borderline personality disorder – the impact of clinical severity. *Psychology Psychother*. 2019;92(1):91-111.
54. Kramer U, Kolly S, Charbon P, Ilagan GS, Choi-Kain LW. Brief psychiatric treatment for borderline personality disorder as a first step of care: Adapting general psychiatric management to a 10-session intervention. *Pers. Disord.: Theory Res. Treat*. 2022;13(5):516.
55. Spong A, Clare I, Galante J, Crawford M, Jones P. Brief psychological interventions for borderline personality disorder. A systematic review and meta-analysis of randomised controlled trials. *Clinical Psychology Review*. 2021;83:101937.
56. McMain SF, Guimond T, Barnhart R, Habinski L, Streiner DL. A randomized trial of brief dialectical behaviour therapy skills training in suicidal patients suffering from borderline disorder. *Acta Psychiatr Scand*. 2017;135(2):138-48.
57. Gratz K, Tull M, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099.

58. Blum N, St. John D, Pfohl B, Stuart S, McCormick B, Allen J, et al. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *Am J Psychiatry*. 2008;165(4):468-78.
59. Hestbæk E, Hasselby-Andersen M, Juul S, Beier N, Simonsen S. Mentalizing the patient–Patient experiences with short-term mentalization-based therapy for borderline personality disorder: A qualitative study. *Front Psychiatry*. 2022;13:1088872.
60. Søndergaard AA, Juul S, Poulsen S, Simonsen S. Mentalizing the therapist–Therapist experiences with short-term mentalization-based therapy for borderline personality disorder: A qualitative study. *Front Psychiatry*. 2023;14:1088865.
61. Bone C, Delgadillo J, Barkham M. A systematic review and meta-analysis of the good-enough level (GEL) literature. *J Counsel Psychol*. 2021;68(2):219.
62. Harnett P, O'Donovan A, Lambert MJ. The dose response relationship in psychotherapy: Implications for social policy. *Clin Psychologist*. 2010;14(2):39-44.
63. Kadera SW, Lambert MJ, Andrews AA. How much therapy is really enough?: A session-by-session analysis of the psychotherapy dose-effect relationship. *J Psychother Pract Res*. 1996;5(2):132.
64. Howard KI, Kopta SM, Krause MS, Orlinsky DE. The dose–effect relationship in psychotherapy. *Am Psychol*. 1986;41(2):159.
65. Barkham M, Rees A, Stiles WB, Shapiro DA, Hardy GE, Reynolds S. Dose–effect relations in time-limited psychotherapy for depression. *J Consult Clin Psychol*. 1996;64(5):927.

66. Cuijpers P, Huibers M, Ebert DD, Koole SL, Andersson G. How much psychotherapy is needed to treat depression? A metaregression analysis. *J Affect Disord.* 2013;149(1-3):1-13.
67. Castonguay LG, Barkham M, Lutz W, McAleavey A. Practice-oriented research: Approaches and applications. *Bergin and Garfield's handbook of psychotherapy and behavior change.* 2013;6:85-133.
68. Jakobsen JC, Gluud C. The necessity of randomized clinical trials. *Br J Med Med Res.* 2013;3(4):1453-68.
69. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, et al. Evaluating non-randomised intervention studies. *Health Technol Assess.* 2003;7(27).
70. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312(7023):71-2.
71. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet.* 2017;390(10092):415-23.
72. Druss BG, Jones N. Evidence-Based Practicing in Mental Health. *JAMA Psychiatry.* 2025.
73. Garattini S, Jakobsen JC, Wetterslev J, Bertelé V, Banzi R, Rath A, et al. Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med.* 2016;32:13-21.
74. Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. *N Eng J Med.* 2020;382(7):674-8.
75. Juul S, Siddiqui F, Faltermeier P, Petersen JJ, Kamp CB, Jakobsen RH, et al. Less focus on symptom scales in psychiatric trials: it is time to ensure research equality between psychiatry and other medical specialities. *Lancet Reg Health – Eur.* 2024;43.

76. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019). Cochrane, 2019. 2019.
77. IntHout J, Ioannidis JPA, Borm GF, Goeman JJ. Small studies are more heterogeneous than large ones: a meta-meta-analysis. *J Clin Epidemiol*. 2015;68(8):860-9.
78. Flather MD, Farkouh ME, Pogue JM, Yusuf S. Strengths and limitations of meta-analysis: Larger studies may be more reliable. *Control Clin Trials*. 1997;18(6):568-79.
79. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet*. 2005;365(9467):1348-53.
80. Edwards SJ, Lilford RJ, Braunholtz D, Jackson J. Why “underpowered” trials are not necessarily unethical. *Lancet*. 1997;350(9080):804-7.
81. Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *JAMA*. 2002;288(3):358-62.
82. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet*. 1993;341(8842):418-22.
83. Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Therapeutics*. 2017;102(6):924-33.
84. Bland JM, Altman DG. Regression towards the mean. *BMJ*. 1994;308(6942):1499.
85. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol*. 2004;34(1):215-20.
86. Garattini S, Jakobsen JC, Wetterslev J, Bertele V, Banzi R, Rath A, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med*. 2016;32:13-21.

87. Garattini S, Jakobsen JC, Wetterslev J, Bertelé V, Banzi R, Rath A, et al. Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med.* 2016;32:13-21.
88. Gluud LL. Bias in clinical intervention research. *Am J Epidemiol.* 2006;163(6):493-501.
89. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med.* 2001;135(11):982-9.
90. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ.* 2001;323(7303):42-6.
91. Moher D, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet.* 1998;352(9128):609-13.
92. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273(5):408-12.
93. Hrobjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol.* 2014;43(4):1272-83.
94. Hrobjartsson A, Skou Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ.* 2013;185(4).

95. Hrobjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ*. 2012;344:e1119.
96. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. 2011.
97. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343.
98. Sterne J, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:14898.
99. Crocker TF, Lam N, Jordão M, Brundle C, Prescott M, Forster A, et al. Risk-of-bias assessment using Cochrane's revised tool for randomized trials (RoB 2) was useful but challenging and resource-intensive: observations from a systematic review. *J Clin Epidemiol*. 2023;161:39-45.
100. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PloS One*. 2013;8(3):e59202.
101. Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. *BMC Med Res Methodol*. 2017;17:1-18.
102. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol*. 2008;61(8):763-9.
103. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol*. 2008;61(1):64-75.

104. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet*. 1998;351(9095):47-52.
105. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Control Clin Trials*. 1997;18(6):580-93.
106. Lan GK, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika*. 1983;70(3):659-63.
107. Devereaux P, Beattie WS, Choi PT, Badner NH, Guyatt GH, Villar JC, et al. How strong is the evidence for the use of perioperative β blockers in non-cardiac surgery? Systematic review and meta-analysis of randomised controlled trials. *BMJ*. 2005;331(7512):313-21.
108. Demets DL, Lan KG. Interim analysis: the alpha spending function approach. *Stat Med*. 1994;13(13-14):1341-52.
109. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error matrix approach for overviewing available evidence. *BMC Med Res Methodol*. 2010;10:1-14.
110. Villar J, Mackey ME, Carroli G, Donner A. Meta-analyses in systematic reviews of randomized controlled trials in perinatal medicine: comparison of fixed and random effects models. *Stat Med*. 2001;20(23):3635-47.
111. Schroll JB, Moustgaard R, Gøtzsche PC. Dealing with substantial heterogeneity in Cochrane reviews. Cross-sectional study. *BMC Med Res Methodol*. 2011;11:1-8.
112. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14(1):120.

113. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical research ed)*. 2008;336(7650):924-6.
114. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol*. 2011;64(4):380-2.
115. Schünemann HJ, Best D, Vist G, Oxman AD, Group GW. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *CMAJ*. 2003;169(7):677-80.
116. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 2021;10(1):89.
117. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Guideline: Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice (ICH-GCP). 2015.
118. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14(1):120.
119. Schünemann H, Brożek J, Guyatt G, Oxman A, The GRADE Working group. GRADE handbook for grading quality of evidence and strength of recommendations. 2013:Available from <https://gdt.gradepro.org/app/handbook/handbook.html> [Accessed June 19, 2025]
120. Philips B, Leichenring F, Jordet H, Karterud S. Flaws in the Design of a Randomized Controlled Trial Comparing Short-Term versus Long-Term Mentalization-

Based Therapy for Borderline Personality Disorder *Psychother Psychosom.* 2024;93(5), 346-347

121. Kvarstein EH, Pedersen G, Urnes Ø, Hummelen B, Wilberg T, Karterud S. Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—Does it make a difference? *Psychol Psychother: Theor Res Pract.* 2015;88(1):71-86.

122. Bateman A, Fonagy P. *Mentalization-based treatment for personality disorders: A practical guide.* Oxford, England: Oxford University Press; 2016.

123. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom.* 2022;91(6):382-97.

124. McMain SF, Chapman AL, Kuo JR, Guimond T, Streiner DL, Dixon-Gordon KL, et al. The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial protocol. *BMC Psychiatry.* 2018;18:1-16.

125. Fonagy P, Target M, Gergely G. Attachment and borderline personality disorder: A theory and some evidence. *Psychiatric Clinics.* 2000;23(1):103-22.

126. Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *JAACAP.* 2012;51(12):1304-13. e3.

127. Bo S, Sharp C, Kongerslev MT, Luyten P, Fonagy P. Improving treatment outcomes for adolescents with borderline personality disorder through a socioecological approach. *Borderline Pers. Disord. Emot. Dysregulation.* 2022;9(1):16.

128. Bevington D, Fuggle P, Fonagy P, Target M, Asen E. Innovations in practice: Adolescent mentalization-based integrative therapy (AMBIT)—A new integrated approach to working with the most hard to reach adolescents with severe complex mental health needs. *Child Adolesc Psychiatry Ment Health*. 2013;18(1):46-51.
129. Junqueira DR, Zorzela L, Golder S, Loke Y, Gagnier JJ, Julious SA, et al. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ*. 2023;381:e073725.
130. Zorzela L, Loke YK, Ioannidis JP, Golder S, Santaguida P, Altman DG, et al. PRISMA harms checklist: improving harms reporting in systematic reviews. *BMJ*. 2016;352:i157.
131. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Guideline for good clinical practice E6(R2)2016 26 July, 2024; Step 5 version.
132. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH E2A Clinical safety data management: definitions and standards for expedited reporting - Scientific guideline1995 26 July, 2024; Step 5 version. Available from: <https://www.ema.europa.eu/en/ich-e2a-clinical-safety-data-management-definitions-standards-expedited-reporting-scientific-guideline>.
133. Juul S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med*. 2020: DOI: bmjebm-2020-111407.
134. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18(3):280.

135. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol.* 2000;53(2):207-16.
136. Schulz KF, Altman DG, Moher D, The CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* 2010;152(11):726-32.
137. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019): Cochrane; 2019.
138. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol.* 2014;14(34).
139. Villar J, Carroli G, Belizán JM. Predictive ability of meta-analyses of randomised controlled trials. *Lancet.* 1995;345(8952):772-6.
140. Cappelleri JC, Ioannidis JPA, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC, et al. Large trials vs meta-analysis of smaller trials: how do their results compare? *JAMA.* 1996;276(16):1332-8.
141. LeLorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between Meta-Analyses and Subsequent Large Randomized, Controlled Trials. *N Eng J Med.* 1997;337(8):536-42.
142. Stoffers-Winterling J, Kongerslev M, Simonsen E, Lieb K, Storebø OJ. Clinical trials in personality disorders: Recommendations for study design and conduct. *Personal Dis: Theor Res Treat.* 2023;14(1):127.

143. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ*. 2000;321(7262):694-6.
144. Hawe P, Shiell A, Riley T. Complex interventions: how “out of control” can a randomised controlled trial be? *BMJ*. 2004;328(7455):1561-3.
145. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008;337:a1655.
146. Datta J, Petticrew M. Challenges to evaluating complex interventions: a content analysis of published papers. *BMC Pub Health*. 2013;13(1):568.
147. Starcevic V. Psychotherapy in the Era of Evidence-Based Medicine. *Australasian Psychiatry*. 2003;11(3):278-81.
148. Margison FR, Barkham M, Evans C, McGrath G, Clark JM, Audin K, et al. Measurement and psychotherapy: Evidence-based practice and practice-based evidence. *Br J Psychiatry*. 2000;177(2):123-30.
149. Philips B, Falkenström F. What research evidence is valid for psychotherapy research? *Front Psychiatry*. 2021;11:625380.
150. Faltinsen E, Todorovac A, Staxen Bruun L, Hróbjartsson A, Gluud C, Kongerslev MT, et al. Control interventions in randomised trials among people with mental health disorders. *Cochrane Database of Syst Rev*. 2022(4).
151. Hoyt MF. Therapist Resistances to Short–Term Dynamic Psychotherapy. *Journal of the American Academy of Psychoanalysis*. 1985;13(1):93-112.
152. Paris J. Stepped care for borderline personality disorder: making treatment brief, effective, and accessible: Academic Press; 2017.

153. Paris J, Black DW. Psychotherapy for Personality Disorders is Underfunded. *Ann Clin Psychiatry*. 2021;33(4):217-8.
154. Paris J. Access to psychotherapy for patients with personality disorders. *Personality and Mental Health*. 2020;14(3):246-53.
155. Wang Y, Devji T, Carrasco-Labra A, King MT, Terluin B, Terwee CB, et al. A step-by-step approach for selecting an optimal minimal important difference. *BMJ*. 2023;381:e073822.
156. Hengartner MP, Plöderl M. Estimates of the minimal important difference to evaluate the clinical significance of antidepressants in the acute treatment of moderate-to-severe depression. *BMJ Evid Based Med*. 2022;27(2):69-73.
157. Cuijpers P. Has the time come to stop using the “standardised mean difference”? *Clin Psychol Eur*. 2021;3(3):1-8.
158. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome Set-STAndards for development: the COS-STAD recommendations. *PLoS Med*. 2017;14(11):e1002447.
159. Prevolnik Rupel V, Jagger B, Fialho LS, Chadderton L-M, Gintner T, Arntz A, et al. Standard set of patient-reported outcomes for personality disorder. *Qual Life Res*. 2021;30(12):3485-500.
160. Grey P, Grey A, Bolland MJ. Outcomes, interventions and funding in randomised research published in high-impact journals. *Trials*. 2018;19(1):592.
161. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J Abnorm Psychol*. 2017;126(4):454-77.

Appendices

Paper 1

Paper 2

Paper 3

Paper 4

Paper 5

Paper 6

Paper 7

Paper 8

Paper 9

Paper 10

Paper 11

Paper 12

Paper 13

Paper 14

Paper 15

Paper 16

Appendices

Paper 1

PROTOCOL

Open Access



Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis

Sophie Juul^{1,2*} , Stig Poulsen², Susanne Lunn², Per Sørensen¹, Janus Christian Jakobsen^{3,4} and Sebastian Simonsen¹

Abstract

Background: Psychiatric disorders are highly prevalent and associated with great symptomatic, functional, and health economic burdens. Psychotherapy is among the recommended and used interventions for most psychiatric disorders and is becoming widely accessible in mental health systems. The effects of specific forms of psychotherapy (e.g., psychodynamic therapies, cognitive and behavioral therapies, humanistic therapies, and systemic therapies) have been assessed previously in systematic reviews, but the appropriate psychotherapy duration for psychiatric disorders has not been reviewed. The aim of this systematic review will be to synthesize the evidence of the effects of short-term compared with long-term psychotherapy for all adult psychiatric disorders.

Methods/design: A comprehensive search for relevant published literature will be undertaken in Cochrane Central Register of Controlled Trials (CENTRAL), Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica database (EMBASE), Latin American and Caribbean Health Sciences Literature (LILACS), PsycINFO, Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index—Science (CPCI-S), and Conference Proceedings Citation Index—Social Science & Humanities (CPCI-SSH) to identify relevant trials. We will search all databases from their inception to the present. We will include randomized clinical trials comparing a short-term and a long-term version of the same psychotherapy type for adult psychiatric disorders including attention deficit hyperactivity disorder, psychotic disorders, depressive disorders, bipolar disorders, anxiety disorders, obsessive-compulsive disorder, trauma- and stressor-related disorders, eating disorders, and personality disorders (as defined by standardized diagnostic criteria). We will rely on the trialists defining their compared interventions as short term and long term (or similar terminology). Primary outcomes will be quality of life, serious adverse events, and symptom severity. Secondary outcomes will be suicide or suicide attempts, self-harm, and level of functioning. Two review authors will independently extract data and perform risk of bias assessment using the Cochrane risk of bias tool. A meta-analysis will be performed as recommended by the Cochrane Handbook for Systematic Review of Interventions, bias will be assessed with domains, and Trial Sequential Analysis will be conducted to control random errors. Certainty of the evidence will be assessed by GRADE.

(Continued on next page)

* Correspondence: sophie.juul@regionh.dk

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark, Stolpegaardsvej 20, 2820 Gentofte, Denmark

²Department of Psychology, University of Copenhagen, Østre Farimagsgade 2A København K, 1353 Copenhagen, Denmark

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

(Continued from previous page)

Discussion: As psychotherapy is among the treatments of choice for most adult psychiatric disorders, a systematic review evaluating the benefits and harms of short-term compared with long-term psychotherapy is urgently needed. It is the hope that this review will be able to inform best practice in treatment and clinical research of these highly prevalent and burdensome disorders.

Systematic review registration: PROSPERO [CRD42019128535](https://www.crd.york.ac.uk/PROSPERO/record/CRD42019128535)

Keywords: Psychotherapy duration, Psychiatric disorders, Short-term psychotherapy, Long-term psychotherapy, Dose-effect

Background

It is estimated that each year, 38.2% of the European population suffer from a psychiatric disorder [1]. The economic burden from psychiatric disorders is excessive, not only because of high direct health care costs, but also because of indirect costs like sick days, disability, and early retirement [1]. Psychotherapy is among the recommended and widely used interventions for most disorders [2]. Specific types of psychotherapy have already been systematically reviewed, but the appropriate length of psychotherapy for all adult psychiatric disorders has not been reviewed previously. To present a complete overview of the evidence and to increase the statistical power, we will therefore in the present review include any adult psychiatric disorder. The major categories of adult psychiatric disorders listed in the *Diagnostic and Statistical Manual of Mental Disorders, 5th edition* (DSM-V) [3] are the following.

Attention Deficit Hyperactivity Disorder (ADHD) is characterized by a persistent pattern of inattention and/or hyperactivity and impulsivity that significantly interferes with functioning and development [3]. ADHD is one of the most common psychiatric disorders of childhood and adolescence, and it often persists into adulthood. The predominant characteristics of adult ADHD differ from typical ADHD characteristics in children. Symptoms of hyperactivity or impulsivity are typically less obvious in adults, whereas symptoms of inattention are more dominant [4]. Epidemiologic studies of adult ADHD have estimated the current prevalence to be 4.4% in the USA and 3.4% internationally [4, 5]. The total economic burden of ADHD in America has been estimated to be 31.6 billion US dollars in 2010 [6] including both direct costs, other health care costs, health care costs for family members, and work loss of patients and their relatives.

Psychotic disorders are characterized by abnormalities in one or more of the following five domains: delusions, hallucinations, disorganized thinking, disorganized or abnormal motor behavior, and negative symptoms [3]. The estimated annual prevalence of all psychotic disorders is 2.6% [7]. The most common psychotic disorder is schizophrenia with an estimated median

lifetime prevalence of 4.0 per 1000 and a lifetime morbid risk of 7.2 per 1000 [8]. Annual costs for the schizophrenia population have been systematically reviewed and estimated to range from 94 million to 102 billion US dollars. Indirect costs contributed to 50–85% of the total costs associated with schizophrenia [9].

Bipolar disorders are characterized by serious mood changes involving mood elevation (mania or hypomania) either alone or followed by major depressive episodes [3]. Bipolar disorder subtypes include bipolar I and bipolar II. Bipolar I disorder is associated with manic episodes nearly always followed by major depressive and hypomanic episodes. Bipolar II disorder is associated with at least one hypomanic episode, at least one major depressive episode, and the absence of manic episodes. The international annual prevalence is estimated to be 0.4% for bipolar I disorder and 0.3% for bipolar II disorder [10]. In 2009, the estimated annual direct and indirect costs of bipolar I and II disorders were 30.7 and 120.3 billion US dollars, respectively [11].

Depressive disorders are characterized by the presence of a sad, empty, and irritable mood often accompanied by somatic and cognitive changes resulting in significant functional impairment [3]. The most common depressive disorder is major depressive disorder (unipolar depression) with an annual prevalence of approximately 7% both in Europe [1] and in the USA [12]. The estimated annual economic burden of adults with major depressive disorder, including direct medical costs, workplace costs, and costs associated with comorbidities exceeded 200 billion US dollars [13] in the USA in 2010.

Anxiety disorders are characterized by excessive and counterproductive feelings of fear and anxiety often accompanied by behavioral disturbances such as pervasive avoidance behaviors [3]. Different anxiety disorders exist, which differ from one another in the types of objects or situations that induce hyperarousal or avoidance behavior [3]. The prevalence of anxiety disorders is estimated to be 18% in the USA [14] and 14% in European countries [1], placing them among the most prevalent psychiatric disorders worldwide. Costs associated with

anxiety disorders have previously been reported to be 46.6 billion US dollars in the USA [15] including both direct and indirect costs.

Obsessive-compulsive disorder (OCD) is characterized by recurrent intrusive thoughts, images, or urges (obsessions) with or without repetitive mental or behavioral acts (compulsions) [3]. OCD among adults has an annual prevalence of 1.2% and a lifetime prevalence of 2.3% [16, 17]. The annual economic burden of OCD is estimated to be 2272 euros per patient when including both direct and indirect costs [18].

Trauma- and stressor-related disorders are characterized by psychological distress following exposure to a traumatic or stressful event. The most common trauma disorder is post-traumatic stress disorder (PTSD) [3]. PTSD is a prevalent and disabling disorder associated with delayed help seeking [19]. The estimated annual prevalence of PTSD is 2% in Europe [1] and 4.7% in the USA [20], and the estimated lifetime prevalence is 3.9% across 26 countries ranging from low to high income [21]. The total costs of PTSD per patient have been estimated to 1082 million euros including both direct and indirect costs [18].

Eating disorders are characterized by a persistent disturbance in eating behavior resulting in altered consumption or absorption of food that significantly impairs health and psychosocial functioning [3]. The most common eating disorders are anorexia nervosa, bulimia nervosa, and binge-eating disorder. Lifetime prevalence of anorexia nervosa, bulimia nervosa, and binge-eating disorder are estimated to be 0.9, 1.5, and 3.5%, respectively, among women and 0.3, 0.5, and 2.0%, respectively, among men [22]. The estimated annual prevalence of eating disorders is 0.9% in the European population [1]. Annual costs per patient are estimated to range from 1288 to 8042 US dollars [23].

Personality disorders are characterized by enduring and inflexible patterns of emotional, behavioral, and interpersonal problems that deviate markedly from cultural expectations. According to DSM-V, the following nine personality disorders exist. Personality disorders onset in adolescence or early adulthood and are associated with great psychosocial distress and impairment [3]. In a systematic review of the economic burden of personality disorders, the estimated direct and indirect costs were 11,126 euros for patients 12 months prior to seeking treatment. Direct medical costs accounted for two thirds of these costs, while the remaining costs were related to productivity losses [24].

Description of the interventions

Different schools of psychotherapy exist. They are often divided into the following categories: *psychodynamic*

therapies, cognitive and behavioral therapies, humanistic therapies, and systemic therapies [2].

Psychodynamic (or psychoanalytical) therapies encompass the many approaches that are influenced by Freud's psychoanalysis but have developed into different independent schools [2]. Traditionally, psychodynamic therapies have been considered as long-term therapies, perhaps due to the notion that the uncovering of unconscious emotions and conflicts cannot be achieved with a fixed time limit [25]. Long-term psychoanalytical psychotherapy has previously been systematically reviewed yielding different results [26, 27]. Today, different lengths of psychodynamic therapies have been developed to treat different forms of psychopathology. In addition to traditional psychoanalysis, examples of long-term psychodynamic treatments are transference-focused psychotherapy (TFP), a psychodynamic treatment rooted in object relations theory lasting up to 3 years [28, 29], and mentalization-based therapy [30], an 18-month psychodynamic treatment rooted in attachment theory. Both are developed specifically to treat borderline personality disorder. Further, different variations of short-term psychodynamic therapy have been developed to treat a variation of common psychiatric disorders, most notably anxiety disorders, depressive disorders, certain behavior disorders, and personality disorders [31]. Short-term psychodynamic therapies vary in treatment duration but typically last between 12 and 24 sessions [31].

Cognitive and behavioral therapies (CBT) encompass many integrative approaches. Historically, behavior therapy (first wave CBT) developed from the learning theories of Pavlov [32] and Skinner [33]. An integration of a cognitive component to classical behavioral theories was first established by Beck [34, 35], who developed what is now often referred to as second-wave CBT. CBT is now often delivered as a short-term treatment, typically lasting between 12 and 20 sessions, for a variation of common psychiatric disorders like depressive disorders [34], anxiety disorders [36], obsessive-compulsive disorder [37], personality disorders [38], and eating disorders [39]. Different durations of CBT are also available for the treatment of schizophrenia [40, 41]. Today, so-called third-wave cognitive therapies have emerged, characterized by more integrative approaches to psychotherapy, incorporating techniques from Buddhist mindfulness, psychodynamic therapies, or Gestalt therapy [2]. These include dialectical behavior therapy [42] and schema-focused therapy (SFT) [29, 43], which are both long-term therapies for borderline personality disorder (up to 3 years), and acceptance and commitment therapy (ACT) [44] and compassion-focused therapy (CFT) [45], which are often delivered as short-term treatments for various psychiatric disorders [46, 47].

Humanistic therapies are characterized by psychotherapy approaches derived from humanistic and existentialist philosophy. Major approaches within this orientation are person-centered therapy [48], Gestalt therapy [49], existential psychotherapy [50], and process-experiential/emotion-focused therapy [51]. All humanistic therapies share the notion of empathic understanding, the promotion of in-therapy experiencing, and a belief in the uniquely human growth tendency by applying a consistent person-centered view involving concern for each patient's individual experience and differing needs [2]. Humanistic therapies have not been developed to treat specific types of disorders and are traditionally considered open-ended, which is also aligned with the person-centered way of thinking. However, different lengths of humanistic therapies have been studied, e.g., PE-EFT as a short-term treatment (down to 5 weeks) for depressive disorders [52] and as a 20-week treatment for trauma-related disorders [53].

Systemic therapies are characterized by a systemic approach to psychotherapy defining patients' problems as contextually rather than individually derived. Most often, the context of interest is the partner or the family, but it can also be a broader context, such as the extended family or a classroom [2]. Different systemic therapies exist for different types of psychopathology. Examples are family-based therapy for eating disorders [54], attachment-based family therapy for depressed adolescents (ABFT) [55], parent management training for childhood conduct disorders [56], psychoeducational family interventions for schizophrenia [57] and bipolar disorder [58], and systemic treatments for substance-use disorders [59, 60]. Different lengths of systemic therapies exist. However, the typical duration is between 10 and 25 sessions.

Other forms of psychotherapy exist, e.g., interpersonal therapy (IPT) [61] or cognitive-analytic therapy (CAT) [62]. However, it is beyond the scope of this review to mention all new approaches to psychotherapy since the field is constantly expanding. Further, despite the existence of well-established manualized and evidence-based approaches to psychotherapy, a large proportion of practicing psychotherapists define themselves as eclectic or integrative [63].

How the interventions might work

It is a common opinion among clinicians and researchers that patients suffering from complex psychiatric distress require longer and more intensive psychotherapy [27]. Complex psychiatric distress can be defined as disorders, which by definition are enduring and inflexible [27], such as personality disorders or schizophrenia, chronic psychiatric disorders (defined as lasting at least a year), or multiple psychiatric disorders. A related assumption is that complex and severe

problems typically take longer to improve than less complex or acute psychiatric distress [25, 64]. This is due to the inherent inflexibility of the psychopathology and the complexity of the required therapeutic techniques. Such potential therapeutic techniques could be provocation of affect or working with the therapeutic alliance [25]. These are techniques that are potentially hard to carry out when faced with time constraints. However, it is often argued that such techniques are essential to effective psychotherapy [65].

In contrast, one could argue that long-term therapies can become counterproductive, given that the same therapeutic techniques will be repeated for a long period of time without continuous assessment of their effects. It is possible that given the limited therapeutic time, planned short-term psychotherapy forces both patients and therapists to establish and maintain a focus throughout the treatment process [66]. Further, issues regarding termination of treatment are particularly important when conducting short-term psychotherapy, where concerns about termination are, almost by definition, always present [67, 68]. Thus, a possible advantage of short-term therapies is that both therapist and patient are forced to address difficult themes associated with separation and loss from the very beginning instead of postponing them for later.

Why is it important to do this review?

It is essential to investigate the optimal duration of psychotherapy for psychiatric disorders, because of the potential patient and health economic burden from long-term psychotherapy and because of the potential harmful effects of terminating treatment prematurely [69]. If short-term psychotherapy is the optimal treatment approach, then this could result in a reduction of waitlists and thus a greater access to evidence-based care. On the contrary, if long-term psychotherapy is the most optimal treatment, then it becomes sensible for mental health systems to invest in these treatments, as they would translate into greater health and occupational benefits [70].

The relationship between the number of sessions (dose) and patient improvement (effect) in psychotherapy has previously been studied with mixed results [70, 71]. There are studies indicating that increased number of sessions is associated with diminishing results [72]. There are also studies indicating that the speed of improvement is dependent on patients pretreatment functioning [73] and that some patients require different dosages to receive the same effect. However, most research on the association between dose and effect is based on uncontrolled studies [70–72, 74, 75] which can only show that patients improve during treatment. Whether this improvement can be attributed to the treatment, can only be established with randomized controlled trials, in which shorter and longer

therapies are directly compared. A systematic review of such randomized clinical trials might allow us to assess the safety profile of the different treatment options directly. We are already aware of two randomized clinical trials comparing a short-term and a long-term version of the same psychotherapy type for one or more adult psychiatric disorders [76, 77]. We have performed a preliminary literature search in the Cochrane Database of Systematic Reviews (search terms, short-term or brief and long-term or standard psychotherapy) for previous systematic reviews comparing a short-term and a long-term version of the same psychotherapy type for one or more adult psychiatric disorders. We identified 1114 hits. From this preliminary literature search, we have only identified one empty systematic review [78].

The present systematic review aims at forming the basis for evidence-based guideline recommendations for the optimal duration of psychotherapy for adult psychiatric disorders taking bias risk (systematic errors), play of chance (random errors), and certainty of the findings into consideration. The objective of this review will be to assess the beneficial and harmful effects of short-term psychotherapy compared with long-term psychotherapy for adult psychiatric disorders.

Methods

The present protocol has been registered in the PROSPERO database (registration number, CRD42019128535) and is being reported in accordance with the reporting guidance provided in the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P) statement [79, 80] (see checklist in Additional file 1).

Criteria for considering studies for this review

Types of studies

We will include randomized clinical trials irrespective of trial design, setting, publication status, publication year, and language. We will not include quasi-randomized trials and observational studies.

Types of participants

Adults (as defined by trialists) with a primary diagnosis of any of the following psychiatric disorders: attention deficit hyperactivity disorder, psychotic disorders, depressive disorders, bipolar disorders, anxiety disorders, obsessive-compulsive disorder, trauma- and stressor-related disorders, eating disorders, and personality disorders, as defined by standardized diagnostic criteria from either ICD-10 [81], DSM-5 [3], or earlier versions (ICD-10 codes: F20–29, F30–39, F40–49, F50–59, F60–69, and F90–90.9). Participants will be included irrespective of sex and comorbidities.

Types of interventions

Experimental group: we will accept any type of short-term psychotherapy (or similar terms used by the trialists).

Control group: we will accept any type of long-term psychotherapy (or similar terms used by the trialists).

We will rely on the trialists defining their compared interventions as short-term and long-term (or similar terminology). We will include trials comparing a short-term and a long-term version of the same psychotherapy type (e.g., short-term psychodynamic therapy compared to long-term psychodynamic therapy). We will not include trials comparing short-term psychotherapy (e.g., short-term cognitive behavioral therapy) with a different type of psychotherapy (e.g., long-term psychodynamic therapy) delivered as long-term therapy. Further, we will include trials with the same dose (sessions) but with different frequencies, e.g., 12 sessions delivered over 6 weeks compared to 12 sessions delivered over 12 weeks.

Outcome measures

Primary outcomes

1. Quality of life (continuous data)
2. Serious adverse events (dichotomous data). We will use the International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use—Good Clinical Practice (ICH-GCP) definition of a serious adverse event, which is any untoward medical occurrence that resulted in death, was life-threatening, required hospitalization or prolonging of existing hospitalization and resulted in persistent or significant disability or jeopardized the patient [82]. If the trialists do not use the ICH-GCP definition, we will include the data if the trialists use the term “serious adverse event.” If the trialists do not use the ICH-GCP definition nor use the term serious adverse event, then we will also include the data, if the event clearly fulfills the ICH-GCP definition for a serious adverse event.
3. Symptom severity assessed by any valid disease-specific symptom scale (continuous data). Symptoms will be analyzed separately for each disorder.

Secondary outcomes

1. Suicide or suicide attempts as defined by trialists (dichotomous data)
2. Self-harm as defined by trialists (dichotomous data)
3. Level of functioning as defined by trialists (continuous data)

Assessment time points

The primary assessment time point will be the time point closest to the end of treatment in the trials' long-term intervention group for all outcomes. For example, if a trial compares a 6-month and a 12-month version of the same psychotherapy type and outcomes are assessed every second month throughout the trial, we will select the assessment time point closest to the end of the 12-month intervention as the primary assessment time point for all outcomes. We will secondarily assess all outcomes at maximum follow-up if longer term follow-up is assessed.

Search methods for identification of studies**Electronic searches**

We will search Cochrane Central Register of Controlled Trials (CENTRAL), Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica database (EMBASE), Latin American and Caribbean Health Sciences Literature (LILACS), PsycINFO, Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index—Science (CPCI-S), and Conference Proceedings Citation Index—Social Science & Humanities (CPCI-SSH) to identify relevant trials. We will search all databases from their inception to the present. For a detailed search strategy for all electronic databases, see Additional file 2. The search strategy for PsycINFO will be given at the review stage.

Searching other resources

The reference lists of relevant publications will be checked for any unidentified randomized trials. We will contact the authors of included studies by email asking for unpublished randomized trials. Further, we will search for ongoing trials on the following:

- [ClinicalTrials.gov](http://www.clinicaltrials.gov) (www.clinicaltrials.gov)
- Google Scholar (<https://scholar.google.dk/>)
- The Turning Research into Practice (TRIP) Database (<https://www.tripdatabase.com/>)
- European Medicines Agency (EMA) (<http://www.ema.europa.eu/ema/>)
- US Food and Drug Administration (FDA) (www.fda.gov)
- China Food and Drug Administration (CFDA) (<http://eng.sfda.gov.cn/WS03/CL0755/>)
- Medicines and Healthcare Products Regulatory Agency (<https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency>)
- The World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) search portal (<http://apps.who.int/trialsearch/>)
- Cochrane Database of Systematic Reviews

- <http://www.evidencebasedpsychotherapies.org/index.php?id=25>

Additionally, we will hand search conference abstracts from psychiatry conferences for relevant trials. We will also consider relevant-for-the-review unpublished and gray literature trials if we identify these.

Data collection and analysis

We will perform the review following recommendations of the Cochrane Collaboration [83]. The analyses will be performed using Trial Sequential Analysis [84] and Stata version 16 (StataCorp LLC, College Station, TX, USA) [85].

Selection of studies

Two authors (SJ and SS) will independently screen titles and abstracts. We will retrieve all relevant full-text study reports/publications, and two review authors (SJ and SS) will independently screen the full text and identify and record reasons for exclusion of the ineligible studies. We will resolve any disagreement through discussion, or if required, we will consult a third person (JCJ). Trial selection will be displayed in an adapted flow diagram as per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [86].

Data extraction and management

Two authors (SJ and SS) will independently extract data from included trials. Disagreements will be resolved by discussion with a third author (JCJ). We will assess duplicate publications and companion papers of a trial together to evaluate all available data simultaneously (maximize data extraction, correct bias assessment). We will contact the trial authors by email to specify any additional data, which may not have been reported sufficiently or at all in the publication.

Trial characteristics

We will extract the following data: bias risk components (as defined below), trial design (parallel, factorial, or crossover), number of intervention arms, length of follow-up, estimation of sample size, and inclusion and exclusion criteria.

Participant characteristics and diagnosis

We will extract the following data: number of randomized participants, number of analyzed participants, number of participants lost to follow-up/withdrawals/crossover, compliance with interventions, age range (mean or median), sex ratio, and type of psychiatric disorder.

We will additionally report the proportion of participants in the compared groups who receive psychotropic medication.

Short-term psychotherapy characteristics

We will extract the following data: short-term psychotherapy type, treatment duration, number of sessions (dose), session lengths (minutes), number of sessions per week, and treatment format.

Long-term psychotherapy characteristics

We will extract the following data: long-term psychotherapy type, treatment duration, number of sessions (dose), session lengths (minutes), number of sessions per week, and treatment format.

Co-intervention characteristics

We will extract the following data: type of co-intervention, treatment duration of co-intervention, number of sessions (or dose), and treatment format.

Outcomes

All outcomes listed above will be extracted from each randomized clinical trial, and we will identify if outcomes are incomplete or selectively reported according to the criteria described later in “incomplete outcome data” bias domain and “selective outcome reporting” bias domain.

Notes

Funding of the trial and notable conflicts of interest of trial authors will be extracted, if available. We will note in the “Characteristics of included studies” table if outcome data were not reported in a usable way. Two review authors (SJ and SS) will independently transfer data into the Stata file [85]. Disagreements will be resolved through discussion, or if required, we will consult with a third author (JCJ).

Assessment of risk of bias in included studies

We will use the instructions given in the *Cochrane Handbook for Systematic Reviews of Interventions* [83] in our evaluation of the methodology and hence the risk of bias of the included trials. We will evaluate the methodology in respect of the following:

- Random sequence generation
- Allocation concealment
- Blinding of participants and treatment providers
- Blinding of outcome assessment
- Incomplete outcome data
- Selective outcome reporting
- Other risk of bias
- Overall risk of bias

Random sequence generation

- *Low risk*: If sequence generation was achieved using computer random number generator or a random number table. Drawing lots, tossing a coin, shuffling cards, and throwing dice will also be considered adequate if performed by an independent adjudicator.
- *Unclear risk*: If the method of randomization was not specified, but the trial was still presented as being randomized
- *High risk*: If the allocation sequence was not randomized or only quasi-randomized. These trials will be excluded.

Allocation concealment

- *Low risk*: If the allocation of patients was performed by a central independent unit, on-site locked computer, identical-looking numbered sealed envelopes, or containers prepared by an independent investigator
- *Uncertain risk*: If the trial was classified as randomized but the allocation concealment process was not described
- *High risk*: If the allocation sequence was familiar to the investigators who assigned participants

Blinding of participants and treatment providers

- *Low risk*: If the participants and the treatment providers were blinded to intervention allocation and this was described
- *Uncertain risk*: If the procedure of blinding was insufficiently described
- *High risk*: If blinding of participants and the treatment providers was not performed

Blinding of outcome assessment

- *Low risk of bias*: If it was mentioned that outcome assessors were blinded and this was sufficiently described
- *Uncertain risk of bias*: If it was not mentioned if the outcome assessors in the trial were blinded or the extent of blinding was insufficiently described
- *High risk of bias*: If no blinding or incomplete blinding of outcome assessors was performed

Incomplete outcome data

- *Low risk of bias*: If missing data were unlikely to make treatment effects depart from plausible values. This could be either (1) there were no drop-outs or

withdrawals for all outcomes or (2) the numbers and reasons for the withdrawals and drop-outs for all outcomes were clearly stated and could be described as being similar to both groups. Generally, the trial is judged as at a low risk of bias due to incomplete outcome data if drop-outs are less than 5%. However, the 5% cut-off is not definitive.

- *Uncertain risk of bias*: If there was insufficient information to assess whether missing data were likely to induce bias on the results.
- *High risk of bias*: If the results were likely to be biased due to missing data either because the pattern of drop-outs could be described as being different in the two intervention groups or the trial used improper methods in dealing with the missing data (e.g., last observation carried forward).

Selective outcome reporting

- *Low risk of bias*: If a protocol was published before or at the time the trial begun and the outcomes specified in the protocol were reported on
- *Uncertain risk of bias*: If no protocol was published
- *High risk of bias*: If the outcomes in the protocol were not reported on

Other risk of bias

- *Low risk of bias*: If the trial appears to be free of other components (for example, academic bias or for-profit bias) that could put it at risk of bias
- *Unclear risk of bias*: If the trial may or may not be free of other components that could put it at risk of bias
- *High risk of bias*: If there are other factors in the trial that could put it at risk of bias (for example, authors conducted trials on the same topic, for-profit bias)

Overall risk of bias

- *Low risk of bias*: The trial will be classified as overall “low risk of bias” only if all of the bias domains described in the above paragraphs are classified as low risk of bias.
- *High risk of bias*: The trial will be classified as “high risk of bias” if any of the bias risk domains described above are classified as “unclear” or high risk of bias.

We will assess the domains “blinding of outcome assessment,” “incomplete outcome data,” and “selective outcome reporting” for each outcome result. Thus, we can assess the bias risk for each outcome assessed in addition to each trial. Our primary conclusions will be

based on the results of our primary outcome results with overall low risk of bias. Both our primary and secondary conclusions will be presented in the summary of findings tables.

Differences between protocol and the review

We will conduct the review according to this published protocol and report any deviations from it in the “Differences between the protocol and the review” section of the systematic review.

Measures of treatment effect

Dichotomous outcomes We will calculate risk ratios (RRs) with 95% confidence interval (CI) for dichotomous outcomes, as well as the Trial Sequential Analysis-adjusted CIs (see below).

Continuous outcomes We will calculate the mean differences (MDs) and consider calculating the standardized mean difference (SMD) with 95% CI for continuous outcomes. We will also calculate trial sequential analysis-adjusted CIs (see below).

Dealing with missing data

We will, as the first option, contact all trial authors to obtain any relevant missing data (i.e., for data extraction and for assessment of risk of bias, as specified above).

Dichotomous outcomes We will not impute missing values for any outcomes in our primary analysis. In our sensitivity analyses (see paragraph below), we will impute data.

Continuous outcomes We will primarily analyze scores assessed at single time points. If only changes from baseline scores are reported, we will analyze the results together with follow-up scores [83]. If standard deviations (SDs) are not reported, we will calculate the SDs using trial data, if possible. We will not use intention-to-treat data if the original report did not contain such data. We will not impute missing values for any outcomes in our primary analysis. In our sensitivity analysis (see paragraph below) for continuous outcomes, we will impute data.

Assessment of heterogeneity

We will primarily investigate forest plots to visually assess any sign of heterogeneity. We will secondly assess the presence of statistical heterogeneity by chi² test (threshold $P < 0.10$) and measure the quantities of heterogeneity by the I^2 statistic [87, 88]. We will investigate possible heterogeneity through subgroup analyses. We

may ultimately decide that a meta-analysis should be avoided [83].

Assessment or reporting biases

We will use a funnel plot to assess reporting bias if ten or more trials are included. We will visually inspect funnel plots to assess the risk of bias. We are aware of the limitations of a funnel plot (i.e., a funnel plot assesses bias due to small sample size). From this information, we assess possible reporting bias. For dichotomous outcomes, we will test asymmetry with the Harbord test [89] if τ^2 is less than 0.1 and with the Rücker test if τ^2 is more than 0.1. For continuous outcomes, we will use the regression asymmetry test [90] and the adjusted rank correlation [91].

Unit of analysis issues We will only include randomized clinical trials. For trials using crossover design, only data from the first period will be included [83, 92]. There will therefore not be any unit of analysis issues. We will not include cluster randomized trials.

Data synthesis

Meta-analysis We will undertake the meta-analysis according to the recommendations stated in the *Cochrane Handbook for Systematic Reviews of Interventions* [83], Keus et al. [93], and the eight-step assessment suggested by Jakobsen et al. [94]. We will use the statistical software Stata version 16 [85] to analyze data. We will assess our intervention effects with both random-effects meta-analyses [95] and fixed-effects meta-analyses [96]. We will use the more conservative point estimate of the two [94]. The more conservative point estimate is the estimate closest to zero effect. If the two estimates are similar, we will use the estimate with the widest CI. We assess a total of six primary and secondary outcomes, and we will therefore consider a *P* value of 0.014 or less as the threshold for statistical significance [94]. We will investigate possible heterogeneity through subgroup analyses. Ultimately, we may decide that a meta-analysis should be avoided [83]. We will use the eight-step procedure to assess if the thresholds for significance are crossed [94]. Our primary conclusion will be based on results with low risk of bias [94]. Where multiple trial arms are reported in a single trial, we will include only the relevant arms. If two comparisons are combined in the same meta-analysis, we will halve the control group to avoid double-counting [83]. Trials with a factorial design will be included. In case of, e.g., a 2×2 factorial designed trial, the two groups receiving short-term interventions will be considered short-term control groups, while the two groups receiving long-term control interventions will be considered long-term control groups. If quantitative

synthesis is not appropriate due to considerable heterogeneity or a small number of included trials, we will report the results in a narrative way.

Trial sequential analysis Traditional meta-analysis runs the risk of random errors due to sparse data and repetitive testing of accumulating data when updating reviews. We wish to control the risks of type I errors and type II errors. We will therefore perform Trial Sequential Analysis on the outcomes, in order to calculate the required information size (that is, the number of participants needed in a meta-analysis to detect or reject a certain intervention effect) and the cumulative *Z*-curve's breach of relevant trial sequential monitoring boundaries [84, 97–104]. A more detailed description of trial sequential analysis can be found in the trial sequential analysis manual [103] and at <http://www.ctu.dk/tsa/>.

For dichotomous outcomes, we will estimate the required information size based on the observed proportion of patients with an outcome in the control group (the cumulative proportion of patients with an event in the control groups relative to all patients in the control groups), a relative risk reduction of 20%, an alpha of 1.4% for all our outcomes, a beta of 20%, and the observed diversity as suggested by the trials in the meta-analysis. For continuous outcomes, we will in the trial sequential analysis use the observed SD, a mean difference of the observed SD/2, an alpha of 1.4% for all outcomes, a beta of 20%, and the observed diversity as suggested by the trials in the meta-analysis.

Subgroup analysis and integration of heterogeneity

Subgroup analysis We will perform the following subgroup analyses when analyzing the primary outcomes (quality of life, serious adverse events, and symptom severity).

1. High risk of bias trials compared to low risk of bias trials
2. Types of psychiatric disorders
3. Types of psychotherapy comparisons
4. Trials above and below the mean difference in intervention lengths

We will use the formal test for subgroup interactions in Stata [85].

Sensitivity analysis To assess the potential impact of the missing data for dichotomous outcomes, we will perform the two following sensitivity analyses on both the primary and secondary outcomes.

- *“Best-worst-case” scenario*: We will assume that all participants lost to follow-up in the short-term experimental group had no serious adverse event, had no suicides, had no suicide attempts, and had no self-harm and that all those participants lost to follow-up in the long-term control group did not survive, had a serious adverse event, had a suicide attempt, and had at least one episode of self-harm.
- *“Worst-best-case” scenario*: We will assume that all participants lost to follow-up in the short-term control group did not survive, had serious adverse event, had a suicide attempt, and had at least one episode of self-harm and that all those participants lost to follow-up in the long-term control group have survived, had no serious adverse event, had no suicide attempts, and had no self-harm.

We will present results of both scenarios in our review. When analyzing quality of life, symptom severity, and level of functioning, a “beneficial outcome” will be the group mean plus two standard deviations (SDs) (we will secondly use one SD in another sensitivity analysis) of the group mean and a “harmful outcome” will be the group mean minus two SDs (we will secondly use one SD in another sensitivity analysis) of the group mean [94]. To assess the potential impact of missing SDs for continuous outcomes, we will perform the following sensitivity analysis.

- Where SDs are missing and it is not possible to calculate them, we will impute SDs from trials with similar populations and low risk of bias. If we find no such trials, we will impute SDs from trials with a similar population. As the final option, we will impute SDs from all trials.

We will present results of this scenario in our review. Other post hoc sensitivity analyses might be warranted if unexpected clinical or statistical heterogeneity is identified during the analysis of the review results [94].

“Summary of findings” table We will create a summary of findings table using each of the prespecified outcomes (quality of life, serious adverse events, symptom severity, suicide and suicide attempts, self-harm, and level of functioning) We will use the five GRADE considerations (bias risk of the trials, consistency of effect, imprecision, indirectness, and publication bias) to assess the quality of a body of evidence as it relates to the studies which contribute data to the meta-analyses for the prespecified outcomes [94, 105–107]. We will assess

imprecision using Trial Sequential Analysis. Otherwise, we will use methods and recommendations described in the *Cochrane Handbook for Systematic Reviews of Interventions* [83] using GRADEpro software. We will justify all decisions to downgrade the quality of studies using footnotes, and we will make comments to aid the reader’s understanding of the review where necessary. Firstly, we will present our results in the Summary of Findings table based on the results from the trials with low risk of bias, and secondly, we will present the results based on all trials.

Discussion

This protocol aims at comparing the effects of short-term psychotherapy with the effects of long-term psychotherapy for common adult psychiatric disorders to determine the best length of treatment. The outcomes will be quality of life, serious adverse events, symptom severity, suicide or suicide attempts, self-harm, and level of functioning.

This protocol has a number of strengths. The predefined methodology is based on the *Cochrane Handbook for Systematic Reviews of Interventions* [83], the eight-step assessment suggested by Jakobsen et al. [94], Trial Sequential Analysis [84], and GRADE assessment [105–107]. Hence, this protocol considers both risks of random errors and risks of systematic errors. Another strength of this protocol is that we pragmatically compare two overall treatment strategies with each other, i.e., the results of this review will potentially reflect the effects of the two strategies in clinical everyday practice.

Our protocol also has some limitations. The primary limitation is the potential for large heterogeneity as a result of including all psychiatric disorders and all types of psychotherapy. Therefore, we may ultimately decide that a meta-analysis is not warranted. Further, psychotherapy always consists of multiple treatment elements and it is likely that different interventions have different effects. Hence, if we show a difference between the compared strategies, it will be difficult to conclude what exactly caused the difference in effect. To minimize this limitation, a number of subgroups are planned, but results of subgroup analyses should always be interpreted with great caution. Another limitation is the large number of comparisons which increase the risk of type 1 error. We have adjusted our thresholds for significance according to the number of primary outcomes, but as mentioned, we have also included multiple subgroup analyses. This large risk of type 1 error will be considered when interpreting the review results. Further, we expect that no trials will have blinded treatment providers and patients. Even though blinding of patients should be relatively easy, blinding of treatment providers is theoretically possible but much more difficult to carry out. Finally, we rely on the trialists defining their

compared interventions as short-term and long-term (or similar terminology). Hence, we will not include trials comparing a short-term and a long-term version of the same psychotherapy type, if the trialists did not explicitly define their interventions with such terminology. Using trialists' definitions of short-term and long-term psychotherapy potentially introduces problems with heterogeneity. However, we believe that our choice of methodology from a pragmatic point of view is the best solution there is. First, trialists often report poorly and often do not themselves use thresholds and important data might be excluded from our review if we demand exact definitions of lengths. Further, we do not expect to include many trials in this systematic review. Hence, relying on trialists' definitions of short-term versus long-term psychotherapy may increase the number of trials being eligible for inclusion. Finally, we believe this pragmatic methodology will lead to the inclusion of the most relevant trials.

Additional files

Additional file 1: PRISMA-P 2015 Checklist. (DOCX 30 kb)

Additional file 2: Search strategies. (DOC 46 kb)

Abbreviations

ABFT: Attachment-based family therapy; ACT: Acceptance and commitment therapy; ADHD: Attention deficit hyperactivity disorder; CAT: Cognitive-analytic therapy; CBT: Cognitive behavioral therapy; CENTRAL: Cochrane Central Register of Controlled Trials; CFT: Compassion-focused therapy; CI: Confidence interval; CPCI-S: Conference Proceedings Citation Index—Science; CPCI-SSH: Conference Proceedings Citation Index—Social Science & Humanities; DSM-V: Diagnostic and statistical manual of mental disorders, 5th edition; EMA: European Medicines Agency; EMBASE: Excerpta Medica database; GRADE: The Grading of Recommendations Assessment, Development and Evaluation; ICD-10: International Classification of Diseases and Related Health Problems – 10th edition; ICH-GCP: International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use – Good Clinical Practice; ICTRP: International Clinical Trials Registry Platform; MD: Mean differences; MEDLINE: Medical Literature Analysis and Retrieval System Online; PRISMA: Preferred reporting items for systematic review and meta-analysis; PRISMA-P: Preferred reporting items for systematic review and meta-analysis – protocols; PROSPERO: International Prospective Register of Systematic Reviews; PTSD: Post-traumatic stress disorder; RR: Risk ratio; SCI-EXPANDED: Science index citation expanded; SD: Standard deviation; SFT: Schema-Focused Therapy; SMD: Standardized mean difference; SSCI: Social Science Citation Index; TFP: Transference-focused psychotherapy; TRIP: Turning research into practice; WHO: World Health Organization

Acknowledgements

The expert help from Sarah Louise Klingenberg (Information Specialist, The Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Copenhagen, Denmark) in making the search strategy is hugely appreciated.

Authors' contributions

SJ wrote up the protocol with regular supervision from SP, SL, JCJ, and SS. JCJ and SJ wrote the methods section. PS read and commented on the final manuscript before it was submitted for publication. All authors read and approved the final manuscript.

Funding

This study is funded by the Mental Health Services, Capital Region of Denmark Research Fund and TrygFoundation, grant no. 123488. The funding

bodies will not be involved in design, collection, analysis, interpretation of data, and in writing up the manuscripts.

Availability of data and materials

Data sharing is not applicable to this protocol article. We will publish all data including code in the supplementary material of the systematic review.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark, Stolpegaardsvej 20, 2820 Gentofte, Denmark. ²Department of Psychology, University of Copenhagen, Østre Farimagsgade 2A København K, 1353 Copenhagen, Denmark. ³Copenhagen Trial Unit, Tagensvej 22, København N, 2200 Copenhagen, Denmark. ⁴Department of Cardiology, Holbæk Hospital, Smedelundsgade 60, 4300 Holbæk, Denmark.

Received: 15 March 2019 Accepted: 5 July 2019

Published online: 13 July 2019

References

- Wittchen HU, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011; 21(9):655–79.
- Barlow DH. The Oxford handbook of clinical psychology: updated edition. New York: Oxford University Press; 2014.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Arlington: American Psychiatric Pub; 2013.
- Kessler RC, et al. The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. *Am J Psychiatry*. 2006;163(4):716–23.
- Fayyad J, et al. Cross-national prevalence and correlates of adult attention-deficit hyperactivity disorder. *Br J Psychiatry*. 2007;190:402–9.
- Birnbaum HG, et al. Costs of attention deficit-hyperactivity disorder (ADHD) in the US: excess costs of persons with ADHD and their family members in 2000. *Curr Med Res Opin*. 2005;21(2):195–205.
- Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe—a critical review and appraisal of 27 studies. *Eur Neuropsychopharmacol*. 2005; 15(4):357–76.
- McGrath J, et al. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev*. 2008;30(1):67–76.
- Chong HY, et al. Global economic burden of schizophrenia: a systematic review. *Neuropsychiatr Dis Treat*. 2016;12:357.
- Merikangas KR, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry*. 2011;68(3): 241–51.
- Dilsaver SC. An estimate of the minimum economic burden of bipolar I and II disorders in the United States: 2009. *J Affect Disord*. 2011;129(1–3):79–83.
- Compton WM, et al. Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991–1992 and 2001–2002. *Am J Psychiatry*. 2006;163(12):2141–7.
- Greenberg PE, et al. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry*. 2015;76(2):155–62.
- Kessler RC, et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593–602.
- DuPont RL, et al. Economic costs of anxiety disorders. *Anxiety*. 1996;2(4): 167–72.
- Kessler RC, et al. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617–27.
- Ruscio AM, et al. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Mol Psychiatry*. 2010;15(1):53–63.

18. Olesen J, et al. The economic cost of brain disorders in Europe. *Eur J Neurol*. 2012;19(1):155–62.
19. Kessler RC, et al. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1995;52(12):1048–60.
20. Goldstein RB, et al. The epidemiology of DSM-5 posttraumatic stress disorder in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions-III. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51(8):1137–48.
21. Koenen K, et al. Posttraumatic stress disorder in the world mental health surveys. *Psychol Med*. 2017;47(13):2260–74.
22. Hudson JI, et al. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol Psychiatry*. 2007;61(3):348–58.
23. Stuhldreher N, et al. Cost-of-illness studies and cost-effectiveness analyses in eating disorders: a systematic review. *Int J Eat Disord*. 2012;45(4):476–91.
24. Soeteman DI, et al. The economic burden of personality disorders in mental health care. *J Clin Psychiatry*. 2008;69:259–65.
25. Crits-Christoph P, Barber JP. Long-term psychotherapy. In: Snyder CR, Ingram RE, editors. *Handbook of psychological change: Psychotherapy processes & practices for the 21st century*. New York: Wiley, Inc.; 2000. p. 455–73.
26. Smit Y, et al. The effectiveness of long-term psychoanalytic psychotherapy - a meta-analysis of randomized controlled trials. *Clin Psychol Rev*. 2012;32(2):81–92.
27. Leichsenring F, Rabung S. Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. *JAMA*. 2008;300(13):1551–65.
28. Kernberg OF, et al. Transference focused psychotherapy: overview and update. *Int J Psychoanal*. 2008;89(3):601–20.
29. Giesen-Bloo J, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Arch Gen Psychiatry*. 2006;63(6):649–58.
30. Bateman A, Fonagy P. *Mentalization-based treatment for personality disorders: a practical guide*. New York: Oxford University Press; 2016.
31. Abbass AA, et al. Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database Syst Rev*. 2014;(7):CD004687.
32. Pavlov I. *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. New York: Oxford University Press; 1927.
33. Skinner BF. *Science and human behavior*. New York: Macmillan; 1953.
34. Beck AT. *Cognitive therapy of depression*. New York: Guilford press; 1979.
35. Beck AT. *Depression: clinical, experimental, and theoretical aspects*. Philadelphia: University of Pennsylvania Press; 1967.
36. Beck A, Rush A. A cognitive model of anxiety formation and anxiety resolution. *Stress Anxiety*. 1975;2:69–80.
37. Salkovskis PM. Obsessional-compulsive problems: a cognitive-behavioural analysis. *Behav Res Ther*. 1985;23(5):571–83.
38. Beck AT, Davis DD, Freeman A. *Cognitive therapy of personality disorders*. New York: Guilford Publications; 2015.
39. Fairburn CG. *Cognitive behavior therapy and eating disorders*. New York: Guilford Press; 2008.
40. Grant PM, et al. Randomized trial to evaluate the efficacy of cognitive therapy for low-functioning patients with schizophrenia. *Arch Gen Psychiatry*. 2012;69(2):121–7.
41. Sensky T, et al. A randomized controlled trial of cognitive-behavioral therapy for persistent symptoms in schizophrenia resistant to medication. *Arch Gen Psychiatry*. 2000;57(2):165–72.
42. Linehan MM, et al. Dialectical behavior therapy for high suicide risk in individuals with borderline personality disorder: a randomized clinical trial and component analysis. *JAMA Psychiatry*. 2015;72(5):475–82.
43. Farrell JM, Shaw IA, Webber MA. A schema-focused approach to group psychotherapy for outpatients with borderline personality disorder: a randomized controlled trial. *J Behav Ther Exp Psychiatry*. 2009;40(2):317–28.
44. Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behav Ther*. 2004; 35(4):639–65.
45. Gilbert P. An introduction to compassion focused therapy in cognitive behavior therapy. *Int J Cogn Ther*. 2010;3(2):97–112.
46. Ruiz FJ. A review of acceptance and commitment therapy (ACT) empirical evidence: correlational, experimental psychopathology, component and outcome studies. *Int J Psychol Psychol Ther*. 2010;10(1): 125.
47. Leaviss J, Uttley L. Psychotherapeutic benefits of compassion-focused therapy: an early systematic review. *Psychol Med*. 2015;45(5):927–45.
48. Rogers CR. A theory of therapy, personality and interpersonal relationships in the client-centred framework. In: Koch S, editor. *Psychology: a study of science*. Vol III, Formulations of the person in social context. New York: McGraw Hill; 1959. p. 184–256.
49. Perls F, Hefferline G, Goodman P. *Gestalt therapy. Excitement and growth in the human personality*. New York: Julian; 1951.
50. Yalom ID. *Existential psychotherapy*. Vol. 1. New York: Basic Books; 1980.
51. Greenberg LS, Van Balen R. The theory of experience-centered therapies. In: Greenberg LS, Watson JC, Lietta G, editors. *Handbook of experiential psychotherapy*. New York: Guilford Press; 1998. p. 28–57.
52. Shinohara K, et al. Behavioural therapies versus other psychological therapies for depression. *Cochrane Database Syst Rev*. 2013;(10).
53. Paivio SC, Nieuwenhuis JA. Efficacy of emotion focused therapy for adult survivors of child abuse: a preliminary study. *J Trauma Stress*. 2001;14(1):115–33.
54. Lock J, le Grange D. Family-based treatment of eating disorders. *Int J Eat Disord*. 2005;37(S1):S64–7.
55. Diamond G, Siqueland L, Diamond GM. Attachment-based family therapy for depressed adolescents: programmatic treatment development. *Clin Child Fam Psychol Rev*. 2003;6(2):107–27.
56. Brestan EV, Eyberg SM. Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *J Clin Child Psychol*. 1998;27(2):180–9.
57. Lam DH. Psychosocial family intervention in schizophrenia: a review of empirical studies. *Psychol Med*. 1991;21(2):423–41.
58. Miklowitz DJ, et al. A randomized study of family-focused psychoeducation and pharmacotherapy in the outpatient management of bipolar disorder. *Arch Gen Psychiatry*. 2003;60(9):904–12.
59. O'Farrell TJ, Clements K. Review of outcome research on marital and family therapy in treatment for alcoholism. *J Marital Fam Ther*. 2012;38(1):122–44.
60. Stanton MD, Shadish WR. Outcome, attrition, and family-couples treatment for drug abuse: a meta-analysis and review of the controlled, comparative studies. *Psychol Bull*. 1997;122(2):170.
61. Cuijpers P, et al. Interpersonal psychotherapy for depression: a meta-analysis. *Am J Psychiatr*. 2011;168(6):581–92.
62. Ryle A, Poynton AM, Brockman BJ. Cognitive-analytic therapy: active participation in change: A new integration in brief psychotherapy. Oxford: Wiley; 1990.
63. Norcross JC. A primer on psychotherapy integration. In: Norcross JC, Goldfried MR, editors. *Handbook of psychotherapy integration*. New York: Basic Books; 2005. p. 3–23.
64. Critchfield KL, Benjamin LS. Principles for psychosocial treatment of personality disorder: summary of the APA Division 12 Task Force/NASPR review. *J Clin Psychol*. 2006;62(6):661–74.
65. Norcross J. Empirically supported therapy relationships. In: Norcross JC, editor. *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. New York: Oxford University Press; 2002. p. 3–16.
66. Piper WE, et al. A comparative study of four forms of psychotherapy. *J Consult Clin Psychol*. 1984;52(2):268.
67. Hoyt MF. Therapist resistances to short-term dynamic psychotherapy. *J Am Acad Psychoanal*. 1985;13(1):93–112.
68. Bloom B. Planned short-term psychotherapies. In: Snyder CR, Ingram RE, editors. *Handbook of psychological change: Psychotherapy processes & practices for the 21st century*. New York: Wiley; 2000. p. 429–54.
69. Lilienfeld S. Psychological treatments that cause harm. *Perspect Psychol Sci*. 2007;2(1):53–70.
70. Cuijpers P, et al. How much psychotherapy is needed to treat depression? A metaregression analysis. *J Affect Disord*. 2013;149(1–3):1–13.
71. Castonguay LG, et al. Practice-oriented research: approaches and applications. *Bergin Garfield's Handb Psychother Behav Chang*. 2013;6:85–133.
72. Howard KI, et al. The dose-effect relationship in psychotherapy. *Am Psychol*. 1986;41(2):159–64.
73. Lambert MJ, Hansen NB, Finch AE. Patient-focused research: using patient outcome data to enhance treatment effects. *J Consult Clin Psychol*. 2001;69(2):159.
74. Stulz N, et al. Dose-effect relationship in routine outpatient psychotherapy: does treatment duration matter? *J Couns Psychol*. 2013;60(4):593–600.
75. Kopta SM, et al. Patterns of symptomatic recovery in psychotherapy. *J Consult Clin Psychol*. 1994;62(5):1009–16.
76. Knekt P, et al. Randomized trial on the effectiveness of long-and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*. 2008;38(5):689–703.

77. Lorentzen S, et al. Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial. *Br J Psychiatry*. 2013;203(3):280–7.
78. Naeem F, Farooq S, Kingdon D. Cognitive behavioural therapy (brief versus standard duration) for schizophrenia. *Cochrane Database Syst Rev*. 2015;(10).
79. Moher D, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):1.
80. Shamseer L, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. 2015;349:7647.
81. WorldHealthOrganization, International classification of diseases, 10th revision (ICD-10). 2009.
82. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH harmonised guideline: integrated addendum to ICH E6(R1): guideline for good clinical practice (ICH-GCP). 2015.
83. Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions* 5.1.0: The Cochrane Collaboration; 2011.
84. Copenhagen Trial Unit. TSA - trial sequential analysis. <http://www.ctu.dk/tsa/>. [Web page]. Accessed 10 July 2019.
85. StataCorp. Stata statistical software: release 16. College Station, TX: StataCorp LLC; 2019. <http://www.stata.com>
86. Moher D, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9.
87. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–58.
88. Higgins JP, et al. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557.
89. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*. 2006;25(20):3443–57.
90. Egger M, et al. Bias in meta-analysis detected by a simple, graphical test. *Bmj*. 1997;315(7109):629–34.
91. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50:1088–101.
92. Elbourne DR, et al. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol*. 2002;31(1):140–9.
93. Keus F, et al. Evidence at a glance: error matrix approach for over-viewing available evidence. *BMC Med Res Methodol*. 2010;10(1):90.
94. Jakobsen JC, et al. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14(1):120.
95. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–88.
96. DeMets DL. Methods for combining randomized clinical trials: strengths and limitations. *Stat Med*. 1987;6(3):341–8.
97. Wetterslev J, et al. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol*. 2008;61(1):64–75.
98. Brok J, et al. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol*. 2008;61(8):763–9.
99. Brok J, et al. Apparently conclusive meta-analyses may be inconclusive—trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *Int J Epidemiol*. 2008;38(1):287–98.
100. Thorlund K, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *Int J Epidemiol*. 2008;38(1):276–86.
101. Wetterslev J, et al. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol*. 2009;9(1):86.
102. Thorlund K, Anema A, Mills E. Interpreting meta-analysis according to the adequacy of sample size. An example using isoniazid chemoprophylaxis for tuberculosis in purified protein derivative negative HIV-infected individuals. *Clin Epidemiol*. 2010;2:57.
103. Thorlund K, et al. User manual for trial sequential analysis (TSA). http://www.ctu.dk/tsa/files/tsa_manual.pdf, 2011.
104. Imberger G, et al. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. *BMJ Open*. 2016;6(8):e011890.
105. Guyatt GH, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical research ed)*. 2008;336(7650):924–6.
106. Schünemann HJ, et al. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *Can Med Assoc J*. 2003;169(7):677–80.
107. Guyatt GH, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol*. 2011;64(4):380–2.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Additional file 1: PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 4:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P: Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 5:15

Section/topic	#	Checklist item	Information reported		Line number(s)			
			Yes	No				
ADMINISTRATIVE INFORMATION								
Title								
Identification	1a	Identify the report as a protocol of a systematic review	☒	☐	2			
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☒	-			
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	☒	☐	73			
Authors								
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	☒	☐	4-29			
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	☒	☐	687-690			
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☒	-			
Support								
Sources	5a	Indicate sources of financial or other support for the review	☒	☐	682-685			
Sponsor	5b	Provide name for the review funder and/or sponsor	☒	☐	682-685			
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	☒	☐	684-685			
INTRODUCTION								
Rationale	6	Describe the rationale for the review in the context of what is already known	☒	☐	79-252			

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	☒	☐	254-258
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	☒	☐	265-315
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	☒	☐	317-347
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	☒	☐	324-325
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	☒	☐	348-352
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	☒	☐	354-367
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	☒	☐	354-367
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	☒	☐	370-403
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	☒	☐	290-307
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	☒	☐	406-486
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	☒	☐	532-601
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2 , Kendall's tau)	☒	☐	532-601

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	☒	☐	567-601
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	☒	☐	548-550
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☒	☐	520-526
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☒	☐	603-613

Additional file 2

Search strategies for Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and Trial Sequential Analysis (Juul et al.)

Cochrane Central Register of Controlled Trials (CENTRAL; 2019, Issue 6) in the Cochrane Library (1 hits)

- #1 MeSH descriptor: [Attention Deficit Disorder with Hyperactivity] explode all trees
- #2 (attention deficit hyperactivity disorder or adhd).ti,ab
- #3 #1 or #2
- #4 MeSH descriptor: [Psychotic Disorders] explode all trees
- #5 MeSH descriptor: [Schizophrenia] explode all trees
- #6 (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) near (think* or motor*))) or schizophreni*).ti,ab
- #7 #4 or #5 or #6
- #8 MeSH descriptor: [Bipolar Disorder] explode all trees
- #9 (bipolar or mood elevation or mania or hypomania or depress*).ti,ab
- #10 #8 or #9
- #11 MeSH descriptor: [Depressive Disorder] explode all trees
- #12 (depressi* or mood or unipolar).ti,ab
- #13 #11 or #12
- #14 MeSH descriptor: [Anxiety Disorders] explode all trees
- #15 (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia).ti,ab
- #16 #14 or #15
- #17 MeSH descriptor: [Obsessive-Compulsive Disorder] explode all trees
- #18 (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive near (mental or behavior*)))ti,ab
- #19 #17 or #18
- #20 MeSH descriptor: [Stress Disorders, Post-Traumatic] explode all trees
- #21 (post-trauma* or trauma*).ti,ab
- #22 #20 or #21
- #23 MeSH descriptor: [Feeding and Eating Disorders] explode all trees
- #24 (eating behavior or anorexia* or bulimia* or binge-eating*).ti,ab
- #25 #23 or #24
- #26 MeSH descriptor: [Personality Disorders] explode all trees
- #27 (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*).ti,ab
- #28 #26 or #27
- #29 3 or 7 or 10 or 13 or 16 or 19 or 22 or 25 or 28
- #30 MeSH descriptor: [Psychotherapy] explode all trees
- #31 (((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care).ti,ab
- #32 #30 or #31
- #33 (brief or extended or standard or intensiv* or ((short* or long*) and term)).ti,ab
- #34 #32 and #33
- #35 #29 and #34

MEDLINE Ovid (1946 to June 2019) (5249 hits)

1. exp Attention Deficit Disorder with Hyperactivity/
2. (attention deficit hyperactivity disorder or adhd).ti,ab.
3. 1 or 2
4. exp Psychotic Disorders/
5. exp Schizophrenia/
6. (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) adj (think* or motor*))) or schizophreni*).ti,ab.
7. 4 or 5 or 6
8. exp Bipolar Disorder/
9. (bipolar or mood elevation or mania or hypomania or depress*).ti,ab.

10. 8 or 9
11. exp Depressive Disorder/
12. (depressi* or mood or unipolar).ti,ab.
13. 11 or 12
14. exp Anxiety Disorders/
15. (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia).ti,ab.
16. 14 or 15
17. exp Obsessive-Compulsive Disorder/
18. (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive adj (mental or behavior*))).ti,ab.
19. 17 or 18
20. exp Stress Disorders, Post-Traumatic/
21. (post-trauma* or trauma*).ti,ab.
22. 20 or 21
23. exp "Feeding and Eating Disorders"/
24. (eating behavior or anorexia* or bulimia* or binge-eating*).ti,ab.
25. 23 or 24
26. exp Personality Disorders/
27. (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*).ti,ab.
28. 26 or 27
29. 3 or 7 or 10 or 13 or 16 or 19 or 22 or 25 or 28
30. exp psychotherapy/
31. (((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care).ti,ab.
32. 30 or 31
33. (brief or extended or standard or intensiv* or ((short* or long*) and term)).ti,ab.
34. 32 and 33
35. 29 and 34
36. limit 35 to ("all adult (19 plus years)" or "adolescent (13 to 18 years)" or "young adult (19 to 24 years)" or "adult (19 to 44 years)" or "young adult and adult (19-24 and 19-44)" or "middle age (45 to 64 years)" or "middle aged (45 plus years)" or "all aged (65 and over)" or "aged (80 and over)")
37. (random* or blind* or placebo* or meta-analys*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
38. 36 and 37

Embase Ovid (1974 to June 2019) (5891 hits)

1. exp attention deficit disorder/
2. (attention deficit hyperactivity disorder or adhd).ti,ab.
3. 1 or 2
4. exp psychosis/
5. exp schizophrenia/
6. (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) adj (think* or motor*)) or schizophreni*).ti,ab.
7. 4 or 5 or 6
8. exp bipolar disorder/
9. (bipolar or mood elevation or mania or hypomania or depress*).ti,ab.
10. 8 or 9
11. exp depression/
12. (depressi* or mood or unipolar).ti,ab.
13. 11 or 12
14. exp anxiety disorder/
15. (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia).ti,ab.
16. 14 or 15
17. exp obsessive compulsive disorder/
18. (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive adj (mental or behavior*))).ti,ab.
19. 17 or 18
20. exp posttraumatic stress disorder/
21. (post-trauma* or trauma*).ti,ab.
22. 20 or 21

23. exp eating disorder/
24. (eating behavior or anorexia* or bulimia* or binge-eating*).ti,ab.
25. 23 or 24
26. exp personality disorder/
27. (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*).ti,ab.
28. 26 or 27
29. 3 or 7 or 10 or 13 or 16 or 19 or 22 or 25 or 28
30. exp psychotherapy/
31. (((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care).ti,ab.
32. 30 or 31
33. (brief or extended or standard or intensiv* or ((short* or long*) and term)).ti,ab.
34. 32 and 33
35. 29 and 34
36. limit 35 to (adult <18 to 64 years> or aged <65+ years>)
37. (random* or blind* or placebo* or meta-analys*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]
38. 36 and 37

LILACS (Bireme; 1982 to June 2019) (1163 hits)

(attention deficit hyperactivity disorder or adhd) or (psychotic or delusion\$ or hallucination\$ or ((disorgani\$ed or abnormal) and (think\$ or motor\$)) or schizophreni\$) or (bipolar or mood elevation or mania or hypomania or depress\$) or (depressi\$ or mood or unipolar) or (anxiet\$ or fear or avoidance behavior\$ or phobia\$ or panic\$ or agoraphobia) or (obsessive compulsive disorder or OCD or urge\$ or obsessi\$ or (repetitive and (mental or behavior\$))) or (post-trauma\$ or trauma\$) or (eating behavior or anorexia\$ or bulimia\$ or binge-eating\$) or (schizotypal\$ or paranoid\$ or schizoid\$ or histrionic\$ or narcissistic\$ or antisocial\$ or borderline\$ or avoidant\$ or dependent or obsessive-compulsi\$) [Words] and (((psycho\$ or cognitive or behavior\$ or humanistic or systemic) and therap\$) or psychotherap\$ or self care or self-care) [Words] and (brief or extended or standard or intensiv\$ or (short\$ or long\$) and term)) [Words]

Science Citation Index Expanded (SCI-EXPANDED) (1900 to June 2019); Social Sciences Citation Index (SSCI) (1956 to June 2019); Conference Proceedings Citation Index- Science (CPCI-S) (1990 to June 2019); and Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) (1990 to June 2019) (Web of Science) (5401 hits)

#7 #6 AND #5

#6 TS=(random* or blind* or placebo* or meta-analys*)

#5 #4 AND #1

#4 #3 AND #2

#3 TS=(brief or extended or standard or intensiv* or ((short* or long*) and term))

#2 TS=(((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care)

#1 TI=((attention deficit hyperactivity disorder or adhd) or (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) and (think* or motor*))) or schizophreni*) or (bipolar or mood elevation or mania or hypomania or depress*) or (depressi* or mood or unipolar) or (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia) or (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive near (mental or behavior*))) or (post-trauma* or trauma*) or (eating behavior or anorexia* or bulimia* or binge-eating*) or (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*))

Paper 2

RESEARCH

Open Access



The difference between shorter- versus longer-term psychotherapy for adult mental health disorders: a systematic review with meta-analysis

Sophie Juul^{1,2*}, Janus Christian Jakobsen^{2,3}, Caroline Kamp Jørgensen^{2,3}, Stig Poulsen⁴, Per Sørensen¹ and Sebastian Simonsen¹

Abstract

Background The optimal psychotherapy duration for mental health disorders is unclear. Our aim was to assess the beneficial and harmful effects of shorter- versus longer-term psychotherapy for adult mental health disorders.

Method We searched relevant databases and websites for published and unpublished randomised clinical trials assessing different durations of the same psychotherapy type before June 27, 2022. Our methodology was based on Cochrane and an eight-step procedure. Primary outcomes were quality of life, serious adverse events, and symptom severity. Secondary outcomes were suicide or suicide-attempts, self-harm, and level of functioning.

Results We included 19 trials randomising 3,447 participants. All trials were at high risk of bias. Three single trials met the required information size needed to confirm or reject realistic intervention effects. One single trial showed no evidence of a difference between 6 versus 12 months dialectical behavioral therapy for borderline personality when assessing quality of life, symptom severity, and level of functioning. One single trial showed evidence of a beneficial effect of adding booster sessions to 8 and 12 weeks of internet-based cognitive behavioral therapy for depression and anxiety when assessing symptom severity and level of functioning. One single trial showed no evidence of a difference between 20 weeks versus 3 years of psychodynamic psychotherapy for mood- or anxiety disorders when assessing symptom severity and level of functioning. It was only possible to conduct two pre-planned meta-analyses. Meta-analysis showed no evidence of a difference between shorter- and longer-term cognitive behavioural therapy for anxiety disorders on anxiety symptoms at end of treatment (SMD: 0.08; 95% CI: -0.47 to 0.63; $p=0.77$; $I^2=73\%$; four trials; very low certainty). Meta-analysis showed no evidence of a difference between shorter and longer-term psychodynamic psychotherapy for mood- and anxiety disorders on level of functioning (SMD 0.16; 95% CI -0.08 to 0.40; $p=0.20$; $I^2=21\%$; two trials; very low certainty).

Conclusions The evidence for shorter versus longer-term psychotherapy for adult mental health disorders is currently unclear. We only identified 19 randomised clinical trials. More trials at low risk of bias and at low risk of random errors assessing participants at different levels of psychopathological severity are urgently needed.

Systematic review registration PROSPERO CRD42019128535.

*Correspondence:

Sophie Juul

sophie.juul@regionh.dk

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Mental health disorders, Duration of psychotherapy, Systematic review, Meta-analysis

Background

The annual prevalence of mental health disorders is estimated to be 38.2% of the European population [1]. The economic burden from mental health disorders is high, both because of direct health care costs, but also because of indirect costs like sick days, disability, and early retirement [1–3]. Psychotherapy is among the recommended and widely used interventions for most disorders [4]. Accordingly, it would be highly relevant to identify the optimal duration of psychotherapy for various mental health disorders and conditions. If short-term psychotherapy is the optimal treatment approach for a given disorder, this could result in a reduction of waitlists and thus a greater access to evidence-based care. On the contrary, if long-term psychotherapy is the most optimal treatment, it would be sensible for mental health systems to invest in these treatments, as they would translate into greater long-term health and occupational benefits [5, 6].

The relationship between dose and effect in psychotherapy has been studied with mixed results in non-controlled studies [5, 7]. While several non-controlled studies indicate that there is a linear or negatively accelerating relationship between number of psychotherapy sessions and outcome for most mental health disorders [8, 9], these findings have been criticized on methodological grounds [10].

The inconclusiveness of the existing research and the general lack of internal validity of non-controlled studies [11, 12] indicate the need for a systematic review of well-designed randomised clinical trials directly comparing psychotherapies of different durations for clearly specified populations, including patients treated for mental health disorders in secondary mental health care settings [11, 12]. However, such systematic review has not previously been performed [6].

The present systematic review aims at forming the basis for evidence-based guideline recommendations for the optimal duration of psychotherapy for adult mental health disorders taking both benefits and harms, bias risk (systematic errors), play of chance (random errors), and certainty of the findings into consideration.

Methods

We report this systematic review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [13]. A PRISMA 2020 checklist can be found in Supplementary material 1. The Cochrane methodology used in this

systematic review is described in detail in our protocol [6], which was also registered in the PROSPERO database (CRD42019128535) prior to the systematic literature search.

Search strategy and selection criteria

Electronic searches

An experienced information specialist searched for eligible trials comparing a shorter with a longer-term version of the same psychotherapy type for one or more adult mental health published before June 27, 2022 in the following databases: Cochrane Central Register of Controlled Trials (CENTRAL), Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica database (EMBASE), Latin American and Caribbean Health Sciences Literature (LILACS), PsycINFO, Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index—Science (CPCI-S), and Conference Proceedings Citation Index—Social Science & Humanities (CPCI-SSH). The electronic search strategies can be found in Supplementary material 2. Additionally, we checked the reference lists of relevant publications for any unidentified trials, and we hand searched conference abstracts from psychiatry conferences for relevant trials. We also considered unpublished and gray literature trials if these were identified.

Inclusion and exclusion criteria

We only included randomized clinical trials. Trials were included irrespective of setting, publication status, publication year, language, and the reporting of our outcomes. We relied on the trialists defining their compared interventions as shorter and longer-term (or similar terminology). We did not include cluster randomized trials, quasi randomized trials, or observational studies.

Data extraction and risk of bias assessment

Two review authors (SJ, CKJ) independently screened relevant trials, extracted data using a standardised data extraction sheet, and assessed the risk of bias according to the Risk of Bias (ROB) assessment tool provided in Cochrane Handbook of Systematic Reviews of Interventions [14]. Any discrepancies were resolved through discussion or, if required, through discussion with a third author (JCJ, SS). We contacted trial authors by e-mail if relevant data were unclear or missing. For more information on our risk of bias assessments, see our protocol [6].

Outcomes and subgroup analyses

Our primary outcomes were quality of life, serious adverse events (as defined by the ICH-GCP guidelines) [15], and symptom severity. Our secondary outcomes were suicide or suicide attempts (dichotomous data), self-harm (dichotomous data), and level of functioning. For all outcomes, we used the trial results reported at the time point closest to the end of treatment in the long-term treatment group.

We planned the following subgroup analyses on our primary outcomes:

- High risk of bias trials compared to low risk of bias trials
- Types of mental health disorders
- Types of psychotherapy comparisons
- Trials above and below the mean difference in intervention lengths

Assessment of statistical and clinical significance

We performed our meta-analyses according to the recommendations stated in the Cochrane Handbook for Systematic Reviews of Interventions [14], Keus et al. [16], and the eight-step procedure suggested by Jakobsen et al. [17] for better validation of meta-analytic results in systematic reviews. Review Manager 5.4 and Stata 16 were used for all meta-analyses [18, 19]. We planned to use risk ratios (RR) for dichotomous outcomes, mean differences (MD) for continuous outcomes assessed with homogeneous measures, and standardised mean difference (SMD) for continuous outcomes with heterogeneous measures. We reported both the random-effects and the fixed-effect meta-analysis results, but primarily emphasized the most conservative result (highest *P* value) of the two results, and considered the less conservative results a sensitivity analysis [17]. We used the best–worst/worst–best case scenarios to assess the potential impact of missing outcome data [6, 17]. We planned to use Trial Sequential Analysis to control for random errors and to report Trial Sequential Analysis-adjusted CIs if the cumulative *Z*-curves did not reach the futility area or passed the diversity-adjusted required information size (DARIS) [6, 17, 20–28]. Trial Sequential Analysis estimates the DARIS (that is the number of participants needed in a meta-analysis to detect or reject a certain intervention effect). When analysing continuous outcomes, we pragmatically anticipated an intervention effect equal to the MD of the observed SD/2 [29]. Heterogeneity was assessed by calculating inconsistency (I^2) for traditional meta-analyses and diversity (D^2) for Trial Sequential Analysis. If it was not possible to perform Trial Sequential Analysis to estimate if there was enough

information, we calculated the required information size for each single trial result and assessed if there was adequate power to confirm or reject realistic intervention effects of single trial results. For dichotomous outcomes, we used the proportion of participants with an event in the control group, a relative risk reduction of 20%, an alpha of 1.4%, and a beta of 20% as predefined in our protocol [6]. For continuous outcomes, we used the observed mean and standard deviation for the control group, the observed mean in the control group plus or minus the observed standard deviation in the control group/2 for the experimental group, an alpha of 1.4%, and a beta of 20% as predefined in our protocol [6]. We assessed a total of six primary and secondary outcome and, hence, considered a *p*-value of 0.014 as the threshold for statistical significance [17, 30]. We performed independent samples *t*-tests to calculate *p*-values for single trial results for continuous outcomes, and Fisher's exact test for single trial results for dichotomous outcomes. We used The Grading of Recommendations Assessment, Development and Evaluation (GRADE) to assess the certainty of evidence [17, 31–33].

Results

Study characteristics

On June 27, 2022 our literature search identified a total of 31,689 records after duplicates were removed (Fig. 1). We included 19 randomised clinical trials enrolling a total of 3,447 participants [34–52] (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished) (Supplementary material 3). A list of excluded studies with reasons can be found in Supplementary material 4.

Characteristics of included trials can be found in Table 1. All trials were assessed as at high risk of bias (Supplementary material 5). Five trials assessed the difference between shorter- and longer-term cognitive behavioural therapy for anxiety disorders [36–38, 42, 48]. Four trials assessed the difference between shorter- and longer-term cognitive behavioural therapy for major depressive disorder [39–41, 43]. Three trials assessed the difference between shorter- and longer-term psychodynamic psychotherapy for major depressive disorder [40, 41, 44]. Two trials assessed the difference between shorter- and longer-term psychodynamic psychotherapy for mood- and anxiety disorders [34, 35]. Three trials assessed the difference between shorter- and longer-term prolonged exposure for post-traumatic stress disorder [46, 47, 49]. One trial assessed the difference between shorter- and longer-term interpersonal therapy for major depressive disorder [39]. One trial assessed the difference

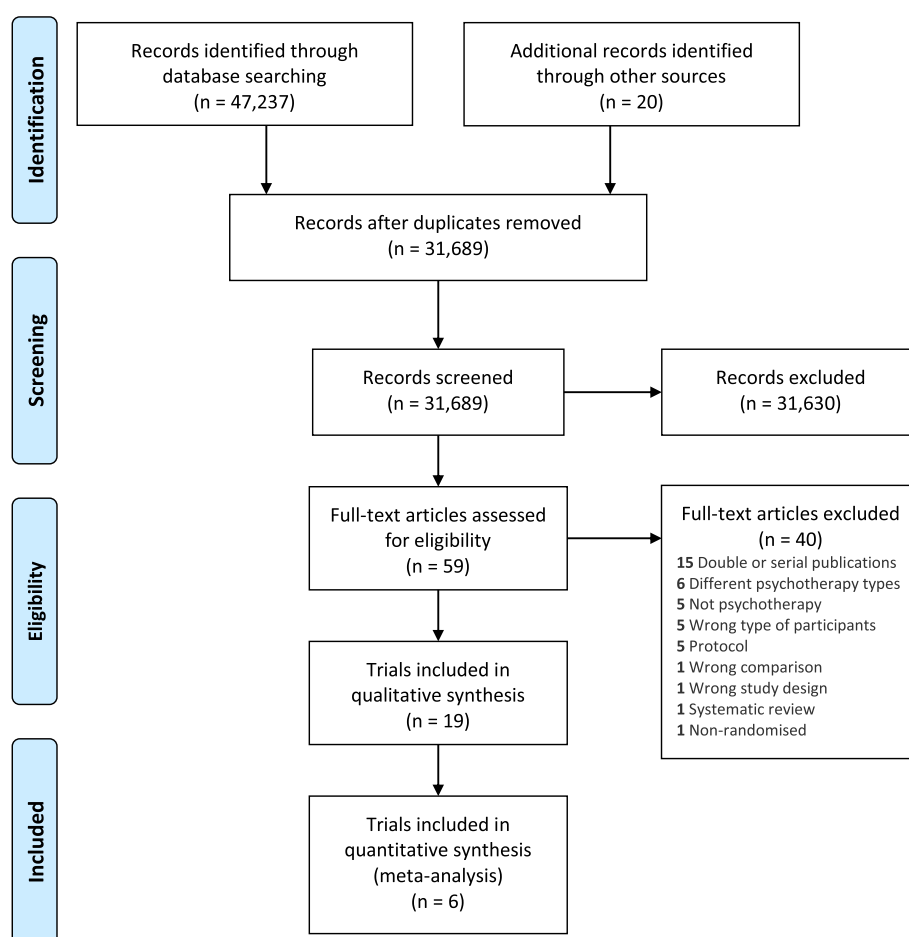


Fig. 1 PRISMA flow diagram

between shorter- and longer-term cognitive behavioural therapy for post-traumatic stress disorder [45]. One trial assessed the difference between shorter- and longer-term internet-based cognitive behavioural therapy for post-traumatic stress disorder [51]. One factorial trial compared internet-based cognitive behavioral therapy for 8 versus 12 weeks with or without booster sessions for depression and anxiety [50]. One trial assessed the difference between shorter- and longer-term dialectical behavioural therapy for borderline personality disorder [52, 53] (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished).

All trials compared different durations (weeks of treatment), dosages (number of sessions), and session lengths (minutes) (Table 1). Furthermore, trialists' definitions of short-term and long-term psychotherapy were not consistent across studies. Most trials compared different numbers of sessions delivered over different durations (e.g. 8 sessions delivered over 8 weeks compared with

16 sessions delivered over 16 weeks) [34, 35, 40–42, 44, 45, 48, 50, 51] (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished). Some trials compared different numbers of sessions delivered over the same duration (e.g. six sessions delivered over 12 weeks compared with 12 sessions delivered over 12 weeks) [37, 38]. Some trials compared the same number of sessions over different durations (e.g. 10 sessions delivered over two weeks compared with 10 sessions delivered over 10 weeks) [36, 39, 46, 49]. Two trials compared the same number of sessions, but with different sessions lengths in minutes (e.g. 10–15 sessions of 60 min compared with 10–15 sessions of 90 min) [43, 47]. We planned to assess serious adverse events. However, only one of the trials reported on this outcome (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished). For several of

Table 1 Characteristics of included trials

Trial	Country	Number of randomised participants	Mental health disorder	Shorter-term intervention	Longer-term intervention	Overall risk of bias	Primary outcome
Barkham et al. 1996^a [40]	United Kingdom	54	Major depressive disorder	8 sessions CBT (8 weeks) 8 sessions psychodynamic-interpersonal therapy (8 weeks)	16 sessions CBT (18 weeks) 16 sessions psychodynamic-interpersonal therapy (18 weeks)	High	No primary outcome was reported
Bohni et al. 2009^a [42]	Denmark	48	Panic disorder	8 sessions CBT (3 weeks)	13 sessions CBT (13 weeks)	High	No primary outcome was reported
Böttche et al. 2021 [51]	Egypt, Saudi Arabia, Algeria, Syria, Morocco, Palestine	224	PTSD	6 sessions internet-based CBT (3 weeks)	10 sessions internet-based CBT (5 weeks)	High	Primary outcome was PTSD symptoms as measured with the Posttraumatic Stress Diagnostic Scale (PDS)
Brujiniks et al. 2020 [39]	The Netherlands	200	Major depressive disorder	20 sessions CBT (24 weeks) 20 sessions interpersonal therapy (24 weeks)	20 sessions CBT (16 weeks) 20 sessions interpersonal therapy (16 weeks)	High	Primary outcome was depression severity as measured with the BDHI
Christensen et al. 2006^a [43]	Australia	931	Major depressive disorder	Brief online CBT and problem solving (unclear duration)	Extended online CBT and problem solving (unclear duration)	High	Primary outcome was depression severity as measured with the Goldberg Depression Scale
Clark et al. 1999 [37]	United Kingdom	29	Panic disorder	5 sessions CBT (12 weeks)	12 sessions CBT (12 weeks)	High	No primary outcome was reported
Dekker et al. 2005 [44]	The Netherlands	103	Major depressive disorder	8 sessions short psychodynamic supportive therapy (8 weeks)	16 sessions short psychodynamic supportive psychotherapy (24 weeks)	High	No primary outcome was reported
Dell et al. 2022 [49]	Australia	138	PTSD	10 sessions prolonged exposure (2 weeks)	10 sessions prolonged exposure (10 weeks)	High	Primary outcome was severity of PTSD symptoms assessed with the Clinician-Administered PTSD Scale (CAPS) ₃
Ehlers et al. 2014 [45]	United Kingdom	61	PTSD	14 sessions cognitive therapy (5 weeks)	12 sessions cognitive therapy (12 weeks)	High	Primary outcome was severity of PTSD symptoms assessed with the Clinician-Administered PTSD Scale (CAPS)
Foa et al. 2018 [46]	USA	219	PTSD	10 sessions prolonged exposure therapy (2 weeks)	10 sessions prolonged exposure therapy (8 weeks)	High	Primary outcome was severity of PTSD symptoms assessed with the PTSD Symptom Scale-Interview (PSS-I)

Table 1 (continued)

Trial	Country	Number of randomised participants	Mental health disorder	Shorter-term intervention	Longer-term intervention	Overall risk of bias	Primary outcome
Hadjistavropoulos et al. 2022 [50]	Canada	434	Depression and/or anxiety	8 sessions internet-based CBT (8 weeks)	11 sessions internet-based CBT (8 weeks + 3 booster sessions)	High	Primary outcomes were severity of depression and anxiety assessed with the Patient Health Questionnaire (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD-7)
Herbert et al. 2004 [36]	USA	34	Social anxiety disorder	12 sessions CBT (12 weeks)	12 sessions CBT (18 weeks)	High	No primary outcome was reported
Kenardy et al. 2003 [48]	Australia and Scotland	81 ^b	Panic disorder	6 sessions CBT (6 weeks)	12 sessions CBT (12 weeks)	High	Primary measures included a comprehensive battery of panic and anxiety measures
Knekt et al. 2008 [34]	Finland	229	Mood- and anxiety disorders	20 sessions psychodynamic therapy (20 weeks)	468 sessions psychodynamic therapy (156 weeks)	High	Primary outcomes were depressive and anxiety symptoms
Lorentzen et al. 2013 [35]	Norway	167	Mood- anxiety- and personality disorders	20 sessions psychodynamic group therapy (20 weeks)	80 sessions psychodynamic group therapy (80 weeks)	High	No primary outcome was reported
McMain et al	Canada	240	BPD	26 sessions DBT (26 weeks)	52 sessions DBT (52 weeks)	High	Primary outcome was frequency of suicidal or non-suicidal self-injurious episodes
Nacasch et al. 2015 [47]	Israel	40	PTSD	10–15 sessions prolonged exposure therapy (60 min)	10–15 sessions prolonged exposure therapy (90 min)	High	Primary outcome was severity of PTSD symptoms assessed with the Post-traumatic Symptom Scale-Interview (PSS-I)
Roberge et al. 2008 [38]	Canada	65	Panic disorder with agoraphobia	7 sessions CBT (16 weeks)	14 sessions CBT (15 weeks)	High	No primary outcome was reported
Shapiro et al. 1994 ^a [41]	United Kingdom	150	Major depressive disorder	8 sessions CBT (8 weeks) 8 sessions psychodynamic-relationship-oriented therapy (8 weeks)	16 sessions CBT (18 weeks) 16 sessions of psychodynamic-relationship-oriented therapy (18 weeks)	High	No primary outcome was reported

^aThe results of these trials were not reported in a usable way; i.e. the results were reported in a graph, and standard deviations were not provided for the point estimates

^bThis trial randomised a total of 186 participants to four groups. The number of randomised participants for the two relevant groups were not sufficiently reported, as only the number of participants who commenced treatment was reported

BPD Borderline personality disorder, CBT Cognitive behavior therapy, DBT Dialectical behavior therapy, PTSD Post-traumatic stress disorder

our review outcomes it was not possible to conduct meta-analysis due to insufficient data. Four trials did not report the results in a usable way [40–43], i.e. they reported the results on a graph and/or did not include standard deviations for each point estimate on a group level. We contacted trial authors to receive relevant data, but we have not received any responses. It was not possible to perform Trial Sequential Analyses to assess the risk of random errors on any of our review outcomes because of lack of relevant data. Only a few trials reported on our dichotomous outcomes, and the continuous outcomes were assessed with heterogeneous measures. We therefore performed sample size calculations for all single trial results to estimate the required information size needed to confirm or reject realistic intervention effects for all outcomes. Results of these sample size calculations can be found in Supplementary material 6.

Due to the large heterogeneity in participants, interventions, and lengths of trials included in this review, we will present the single trial results first. Second, we will present the meta-analysis results.

Single trial results

Trials including participants with borderline personality disorder

We identified one trial randomising 240 participants with borderline personality disorder to six months versus 12 months dialectical behavioral therapy [52, 53] (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished). We retrieved the data through the published trial report and personal communication with the trialists. This trial reported data on all our pre-defined review outcomes. It was not possible to include the trial in a pre-defined meta-analysis, as it was the only trial including participants with borderline personality disorder. The trial reached their pre-calculated sample size of 240 participants [52, 53]. The trial showed no evidence of a difference between short-term and long-term dialectical behavioral therapy when assessing quality of life ($p=0.831$, required information size reached), serious adverse events ($p=1$, required information size not reached), symptom severity ($p=0.833$, required information size reached), suicide or suicide attempts ($p=1$, required information size not reached), self-harm ($p=0.28$, required information size not reached), and level of functioning ($p=0.731$, required information size reached) (Table 2, Supplementary material 6). This trial was assessed as at overall high risk of bias due to lack of blinding of participants and personnel, and due

to incomplete outcome data (Supplementary material 5), and the certainty of evidence was assessed as “very low” for all outcomes (Supplementary material 7).

Trials including participants with mood- and anxiety disorders

We identified three trials assessing the effects of shorter- versus longer-term psychotherapy for mood- and anxiety disorders [34, 35, 50].

One trial randomising 229 participants with mood- and anxiety disorders to 20 weeks versus 156 weeks of psychodynamic psychotherapy [34] showed no evidence of a difference when assessing symptom severity ($p=0.037$, required information size reached), considering our adjusted threshold for significance was pre-defined at 0.014 in our protocol [6], or level of functioning ($p=0.066$, required information size reached). The trial almost reached their sample size (230 participants) [34], but it was unclear whether this sample size was pre-defined. One trial randomising 167 participants with mood- and anxiety disorders to 20 weeks versus 80 weeks of psychodynamic psychotherapy [35] showed no evidence of a difference when assessing the proportion of participants with a suicide or a suicide attempts (zero events in both groups) or level of functioning ($p=0.889$, required information size not reached) (Table 2, Supplementary material 6). Both trials were assessed at high risk of bias (Supplementary material 5) and the certainty of evidence was assessed as “very low” for all outcomes (Supplementary material 8). These two trials are included in a meta-analysis (see below).

We also identified one factorial trial randomising 496 participants with major depressive disorder and anxiety disorders to internet-based cognitive behavioral therapy for 8 versus 12 weeks with or without 3 booster sessions [50]. This trial showed no evidence of a difference when assessing quality of life for either of the two pairwise comparisons (8 weeks versus 8 weeks plus boosters $p=0.858$; 12 weeks versus 12 weeks plus boosters $p=0.089$; required information size reached). The trial showed evidence of a beneficial effect of adding booster sessions in both pairwise comparisons when assessing symptom severity (8 weeks versus 8 weeks plus boosters $p=0.01$; 12 weeks versus 12 weeks plus boosters $p=0.01$; required information size reached) and level of functioning (8 weeks versus 8 weeks plus boosters $p=0.01$; 12 weeks versus 12 weeks plus boosters $p=0.01$; required information size reached) (Table 2, Supplementary material 6). Both trials were assessed at high risk of bias (Supplementary material 5), and the certainty of evidence was assessed as “very low” for all outcomes (Supplementary materials 9 and 10).

Table 2 Single trial results

Trials including participants with major depressive disorder ^a									
Trial characteristics		Primary review outcomes			Secondary review outcomes			Trialists' own conclusions	
Trial ID	Shorter-term intervention	Longer-term intervention	Quality of Life	Serious Adverse Events	Symptom severity	Suicide/Suicide attempts	Self-harm	Level of functioning	
Barkham et al. 1996 [40]	8 sessions CBT (8 weeks) 8 sessions psychodynamic-interpersonal therapy (8 weeks)	16 sessions CBT (18 weeks) 16 sessions psychodynamic-interpersonal therapy (18 weeks)	-	-	-	-	-	-	Clients given 16 sessions showed a statistically significant advantage over clients given 8 sessions on some measures at some assessments
Bruijniks et al. 2020 [39]	20 sessions CBT (24 weeks) 20 sessions interpersonal therapy (24 weeks)	20 sessions CBT (16 weeks) 20 sessions interpersonal therapy (16 weeks)	CBT: The mean RAND-36 scores at EoT were 50.13 (22.20) for the short-term group (n = 49) and 51.53 (22.36) for the long-term group (n = 39) (p = 0.77) IPT: The mean RAND-36 scores at EoT were 46.8 (20.46) in the short-term group (n = 36) and 53.46 (20.67) in the long-term group (n = 47) (p = 0.14)	-	CBT: The mean (SD) BDI scores at EoT were 24.16 (15.09) for the short-term group (n = 37) and 21.25 (12.90) for the long-term group (n = 35) (p = 0.38) IPT: The mean (SD) BDI scores at EoT were 22.91 (14.75) for the short-term group (n = 34) and 20.02 (16.05) for the long-term group (n = 39) (p = 0.42)	-	-	-	In clinical practice settings, delivery of twice weekly sessions of CBT and IPT for depression was superior to once weekly sessions when assessing depression outcomes
Christensen 2006 [43]	Brief online CBT and problem solving (unclear duration)	Extended online CBT and problem solving (unclear duration)	-	-	-	-	-	-	Brief CBT-based interventions are not as effective as extended interventions

Table 2 (continued)

Dekker et al. 2005 [44]	8 sessions short psychodynamic supportive psychotherapy (8 weeks)	16 sessions short psychodynamic supportive psychotherapy (24 weeks)	The mean (SD) QIDS scores at EoT were 22.6 (8.6) for the short-term group (n=45) and 22.8 (8.3) for the long-term group (n=45) ($p=0.911$)	-	The mean (SD) HDRS scores at EoT were 11.1 (6.8) for the short-term group (n=45) and 12.1 (7.6) for the long-term group (n=45) ($p=0.512$)	-	Eight or 16 psychotherapy sessions in addition to 8 sessions of pharmacotherapy over a period of 6 months would appear to be equally effective in terms of dealing with symptoms
Shapiro et al. 1994 [41]	8 sessions CBT (8 weeks)	16 sessions CBT (18 weeks)	-	-	-	-	There is no added benefit from 16 treatment sessions compared with 8
Trials including participants with anxiety disorders^a							
Trial characteristics							
Trial ID	Shorter-term intervention	Longer-term intervention	Quality of Life	Serious Adverse Events	Symptom severity	Suicide/Suicide attempts	Self-harm
Bohni et al. 2009 [42]	8 sessions CBT (3 weeks)	13 sessions CBT (13 weeks)	-	-	-	-	-
Secondary review outcomes							Level of functioning
Trialsists' own conclusions							Patients in massed CBT achieved their results at a faster rate than patients in spaced CBT, with outcomes after 3 weeks in massed CBT comparable with those achieved after approximately 3 months in spaced CBT

Table 2 (continued)

Clark et al. 1999 [37]	5 sessions CBT (12 weeks)	12 sessions CBT (12 weeks)	-	-	The mean (SD) BAI scores at EoT were 9.8 (6.7) for the short-term group (<i>n</i> = 14) and 8.4 (8.0) for the long-term group (<i>n</i> = 15). This result is included in a meta-analysis	-	-	Brief CT did not differ from full CT at posttreatment or at follow-up, and effect sizes were essentially the same
Herbert et al. 2004 [36]	12 sessions CBT (12 weeks)	12 sessions CBT (18 weeks)	-	-	The mean (SD) SPAI-SP scores at EoT were 76.71 (47.18) for the short-term group (<i>n</i> = 15) and 113.77 (39.56) for the long-term group (<i>n</i> = 19). This result is included in a meta-analysis	-	-	The results revealed that the standard treatment program in which therapy was provided over 12 successive weeks resulted in more rapid symptom reduction and lower dropout relative to the extended treatment delivered over 18 weeks
Kenardy et al. 2003 [48]	6 sessions CBT (6 weeks)	12 sessions CBT (12 weeks)	-	-	The mean (SD) STAI-T scores at EoT were 47.86 (12.31) for the short-term group (<i>n</i> = 39) and 41.10 (13.14) for the long-term group (<i>n</i> = 42). This result is included in a meta-analysis	-	-	A brief version performs significantly worse than the standard duration treatment at post-treatment

Table 2 (continued)

Roberge et al. 2008 [38]	7 sessions CBT (16 weeks)	14 sessions CBT (15 weeks)	-	-	The mean (SD) PAS scores at EoT were 10.2 (8.8) for the short-term group ($n = 32$) and 9.5 (10.3) for the long-term group ($n = 33$) ($p = 0.77$). This result is included in a meta-analysis	-	-	Brief CBT effectiveness appears comparable to standard CBT in the short term
Trials including participants with mood- and anxiety disorders^a								
Trial characteristics			Primary review outcomes			Secondary review outcomes		
Trial ID	Shorter-term intervention	Longer-term intervention	Quality of Life	Serious Adverse Events	Symptom severity	Suicide/Suicide attempts	Self-harm	Trialists' own conclusions
Knekt et al. 2008 [34]	Short-term psychodynamic therapy	Long-term psychodynamic therapy	-	-	The mean (SD) HDRS scores at EoT were 10.8 (5.65) for the short-term group ($n = 83$) and 9.0 (6.0) for the long-term group ($n = 107$) ($p = 0.037$)	-	-	Patients receiving short-term psychodynamic psychotherapy recovered faster from both depressive and anxiety symptoms during the first year of follow-up. During the following 2 years, the symptoms persisted at the level reached in the brief therapy group, whereas in the long-term psychodynamic psychotherapy group the improvement continued during the entire 3-year period. In the long run, long-term psychodynamic psychotherapy thus gave greater benefits than those achieved by the brief therapies

Table 2 (continued)

Lorentzen et al. 2013 [35]	Short-term psychodynamic group therapy	Long-term psychodynamic group therapy	-	-	There were 0/77 suicides or suicide attempts in the short-term group compared to 0/90 in the long-term group ($p = \text{not applicable}$)	The mean (SD) GAF scores at EoT were 67.8 (11.7) for the short-term group ($n = 71$) and 68.1 (14.2) for the long-term group ($n = 79$) ($p = 0.889$). This result is included in a meta-analysis	We observed that short- and long-term therapy were equally effective across 3 years, using IIP, GAF-S and GAF-F as the outcome variables. However, there was a trend in favour of long-term therapy ($P = 0.10$) using GAF-S as the outcome variable
Hadjistavropoulos et al. 2022 [50]	Internet-based CBT (8 weeks)	Internet-based CBT (8 weeks + 3 booster sessions)	-	-			
			The mean (SD) EQ-5D-5L scores at EoT were 71.92 (18.93) for the short-term group ($n = 79$) and 71.36 (21.34) for the long-term group ($n = 87$) ($p = 0.858$)	The mean (SD) PHQ-9 scores at EoT were 7.93 (5.36) for the short-term group ($n = 79$) and 5.84 (5.07) for the long-term group ($n = 87$) ($p = 0.01$) The mean (SD) GAD-7 scores at EoT were 7.56 (5.37) for the short-term group ($n = 79$) and 5.56 (4.60) for the long-term group ($n = 87$) ($p = 0.01$)		The mean (SD) SDS scores at EoT were 13.86 (7.86) for the short-term group ($n = 79$) and 10.66 (8.7) for the long-term group ($n = 87$) ($p = 0.01$)	No significant group differences were found in this study

Table 2 (continued)

Trials including participants with post-traumatic stress disorder ^a									
Trial characteristics			Primary review outcomes				Secondary review outcomes		Trialists' own conclusions
Trial ID	Shorter-term intervention	Longer-term intervention	Quality of Life	Serious Adverse Events	Symptom severity	Suicide/Suicide attempts	Self-harm	Level of functioning	
Böttche et al. 2021 [51]	6 sessions internet-based CBT (3 weeks)	10 sessions internet-based CBT (5 weeks)	The mean (SD) EUROHIS-QOL-8 scores at EoT were 5.53 (0.83) for the short-term group (n = unclear) and 5.11 (1.02) for the long-term group (n = unclear) (p = 0.75)	-	The mean (SD) PDS change scores at EoT were -14.73 (1.45) for the short-term group (n = unclear) and -15.03 (1.64) for the long-term group (n = unclear) (n = unclear) (p = 0.89)	-	-	-	The shorter condition results in the same symptom change and dropout rate as the longer condition
	Internet-based CBT (12 weeks)	Internet-based CBT (12 weeks + 3 booster sessions)	The mean (SD) EQ-5D-5L scores at EoT were 74.06 (15.94) for the short-term group (n = 87) and 69.13 (21.93) for the long-term group (n = 91) (p = 0.089)	-	The mean (SD) PHQ-9 scores at EoT were 6.52 (5.23) for the short-term group (n = 87) and 7.55 (6.24) for the long-term group (n = 91) (p = 0.235)	-			The mean (SD) SDS scores at EoT were 10.29 (8.01) for the short-term group (n = 87) and 11.39 (8.46) for the long-term group (n = 91) (p = 0.374)

Table 2 (continued)

Dell et al. 2022 [49]	Massed prolonged exposure	Standard prolonged exposure	-	0 events	The mean (SD) CAPS scores at EoT were 27.69 (18.42) for the short-term group ($n=63$) and 25.68 (16.59) for the long-term group ($n=71$) ($p=0.664$)	0 events	-	Massed prolonged exposure was non-inferior to standard prolonged exposure in reducing symptoms of PTSD
Ehlers et al. 2014 [45]	Intensive cognitive therapy	Standard cognitive therapy	The mean (SD) Q-LES-Q scores at EoT were 52.67 (20.21) for the short-term group ($n=30$) and 62.93 (21.70) for the long-term group ($n=31$) ($p=0.061$)	-	The mean (SD) CAPS scores at EoT were 32.22 (27.20) for the short-term group ($n=30$) and 26.97 (28.68) for the long-term group ($n=31$) ($p=0.466$)	-	The mean (SD) SDS scores at EoT were 9.30 (8.20) for the short-term group ($n=30$) and 10.02 (9.76) for the long-term group ($n=31$) ($p=0.757$)	A novel 7-day intensive version of cognitive therapy for PTSD was well tolerated, achieved faster symptom reduction, and led to comparable overall outcomes as the standard once-weekly cognitive therapy delivered over 3 months
Foa et al. 2018 [46]	Massed prolonged exposure	Extended prolonged exposure	-	-	The mean (SD) PSS-I scores at EoT were 18.88 (no SD reported) for the short-term group ($n=110$) and 18.34 (no SD reported) for the long-term group ($n=110$) ($p=$ not applicable)	-	-	Among active duty military personnel with PTSD, massed prolonged exposure therapy (10 sessions delivered over 2 weeks) was noninferior to spaced prolonged exposure therapy (10 sessions delivered over 8 weeks)

Table 2 (continued)

Nacasch et al. 2015 [47]	60 min sessions of prolonged exposure	90 min sessions of prolonged exposure	-	-	The mean (SD) PSS-I scores at EoT were 13.3 (9.52) for the short-term group ($n = 20$) and 12.24 (8.02) for the long-term group ($n = 17$) ($p = 0.719$)	-	-	In sum, 20-min imaginal exposure within 60-min sessions yielded non-inferior outcomes in PTSD symptoms and posttraumatic negative cognitions at posttreatment and follow-up to the 40-min imaginal exposures and 90-min sessions
Trials including participants with borderline personality disorder^a								
Trial characteristics								
Trial	Shorter-term intervention	Longer-term intervention	Quality of Life	Serious Adverse Events	Symptom severity	Suicide/Suicide attempts	Self-harm	Level of functioning
McMain et al. 2022 [52]	6 months of DBT	12 months of DBT	The mean (SD) overall EQ5DL scores at EoT were 60.7 (21.43) for the short-term group ($n = 91$) compared with 61.41 (23.17) in the long-term group ($n = 90$) ($p = 0.831$)	2 / 90 participants had one or more serious adverse events in the short-term group at EoT compared with 2 / 93 in the long-term group ($p = 1$) (based on suicide/ suicide attempt data only)	The mean (SD) BSL scores at EoT were 38.6 (22.4) for the short-term group ($n = 90$) compared with 39.3 (22.2) in the long-term group ($n = 91$) ($p = 0.833$)	2 / 90 participants had one suicide or suicide-attempts in the short-term group at EoT compared with 2 / 93 in the long-term group ($p = 1$)	28 / 90 participants had one or more deliberate self-harm incidents in the short-term group at EoT compared with 37 / 93 in the long-term group ($p = 0.28$)	The mean (SD) SAS scores at EoT were 2.51 (0.58) for the short-term group ($n = 90$) compared with 2.54 (0.59) in the long-term group ($n = 91$) ($p = 0.731$)
effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished)								Half the dose of the standard DBT yielded noninferior improvements across time points for the primary outcome, total self-harm frequency, as well as several clinical outcomes

^a Data is presented for the primary time-point of assessment (end of treatment)

BAI Beck Anxiety Inventory, BSL Borderline Symptom List-23, DBT Dialectical Behavioural Therapy, CAPS Clinician Administered PTSD Scale, CBT Cognitive Behavioural Therapy, EoT End of treatment, EQ5DL Euroqol-5D-5, HDPS Hamilton Depression Rating Scale, IPT Interpersonal Therapy, PAS Panic and Agoraphobia Scale, PSS-I PTSD Symptom Scale Interview, SAS Social Adjustment Scale, SD Standard deviation, SDS Sheehan Disability Scale, SPAI-SP Social Phobia Anxiety Inventory – Social Phobia, STAI-T State Trait Anxiety Inventory-Trait, QLDS Quality of Life Depression Scale, Q-LES-Q Quality of Life Enjoyment and Satisfaction Questionnaire

Trials including participants with major depressive disorder

We identified five trials including eight comparisons assessing the effects of shorter- versus longer-term psychotherapy for participants with major depressive disorder [39–41, 43, 44]. Four trials compared shorter- versus longer-term cognitive behavioural therapy for major depressive disorder [39–41, 43]. Three trials compared shorter- versus longer-term psychodynamic psychotherapy for major depressive disorder [40, 41, 44]. One trial compared shorter- versus longer-term interpersonal therapy for major depressive disorder [39]. It was not possible to perform meta-analyses, as the trials differed in the assessed psychotherapy traditions, and only two trials reported on our pre-defined review outcomes [39, 44].

One trial randomising 200 participants with major depressive disorder to once- versus twice weekly cognitive behavioral therapy or interpersonal therapy [39] showed no evidence of a difference when assessing quality of life and symptom severity for either cognitive behavioral therapy ($p=0.77$ and $p=0.38$, required information size not reached) or interpersonal therapy ($p=0.14$ and $p=0.42$, required information size not reached). One trial randomising 103 participants with major depressive disorder to eight versus 16 sessions of short-term psychodynamic supportive psychotherapy [44] showed no evidence of a difference when assessing quality of life ($p=0.911$, required information size not reached) or symptom severity ($p=0.512$, required information size not reached) (Table 2, Supplementary material 6). Both trials were assessed at high risk of bias (Supplementary material 5) and the certainty of evidence was assessed as “very low” for all outcomes (Supplementary materials 11, 12, and 13).

Trials including participants with post-traumatic stress disorder

We identified five trials assessing the effects of shorter- versus longer-term psychotherapy for participants with post-traumatic stress disorder [45–47, 49, 51]. Three trials compared shorter- versus longer-term prolonged exposure for post-traumatic stress disorder [46, 47, 49]. One trial compared shorter- versus longer-term cognitive behavioral therapy for post-traumatic stress disorder [45]. One trial compared shorter- versus longer-term internet-based cognitive behavioral therapy for post-traumatic stress disorder [51]. It was not possible to perform meta-analyses, as the trials differed in the assessed psychotherapy traditions, and one of them did not report standard deviations [46]. The two remaining trials reported on some of our pre-defined review outcomes.

One trial randomising 224 participants with post-traumatic stress disorder to 6 versus 10 assignments

of internet-based cognitive behavioral therapy showed no evidence of a difference when assessing quality of life ($p=0.75$, required information size not reached) and symptom severity ($p=0.89$, required information size not reached) [51]. One trial randomising 138 participants with post-traumatic stress disorder to massed prolonged exposure (10 sessions delivered over 2 weeks) versus standard prolonged exposure (10 sessions delivered over 10 weeks) showed no evidence of a difference when assessing symptom severity ($p=0.664$; required information size not reached) [49]. One similar trial did not report standard deviations, but the trialists concluded that massed prolonged exposure therapy (10 sessions delivered over 2 weeks) was noninferior to spaced prolonged exposure therapy (10 sessions delivered over 8 weeks) [46]. One trial randomising 61 participants with post-traumatic stress disorder to intensive (5 weeks) versus standard (12 weeks) cognitive therapy [45] showed no evidence of a difference when assessing quality of life ($p=0.061$, required information size not reached), symptom severity ($p=0.466$, required information size not reached), or level of functioning ($p=0.757$, required information size not reached). One trial randomising 40 participants with post-traumatic stress disorder to 60 min versus 90 min sessions of prolonged exposure therapy [47] showed no evidence of a difference when assessing symptom severity ($p=0.719$, required information size not reached) (Table 2, Supplementary material 6). All trials were assessed at high risk of bias (Supplementary material 5) and the certainty of evidence was assessed as “very low” for all outcomes (Supplementary materials 14, 15, and 16).

Trials including participants with anxiety disorders

We identified five trials assessing the effects of shorter- versus longer-term cognitive behavioral therapy for anxiety disorders [36–38, 42, 48]. One trial did not report the results in a usable way; i.e. the results were reported on a graph and standard deviations were not reported [42].

One trial randomising 29 participants with panic disorder to five versus 12 sessions cognitive behavioral therapy [37] showed no evidence of a difference when assessing symptom severity ($p=0.615$, required information size not reached). One trial randomising 34 participants with social anxiety disorder to 12 versus 18 weeks of cognitive behavioral therapy [36] showed no evidence of a difference when assessing symptom severity ($p=0.018$, required information size not reached), considering our adjusted threshold for significance was pre-defined at 0.014 in our protocol [6]. One trial randomising 81

participants to six versus 12 weeks of cognitive behavioral therapy for participants with panic disorder [48] showed no evidence of a difference when assessing symptom severity ($p=0.0195$, required information size not reached), considering our adjusted threshold for significance was pre-defined at 0.014 in our protocol [6]. One trial randomising 65 participants with panic disorder and agoraphobia to 7 sessions versus 14 sessions cognitive behavioral therapy [38] showed no evidence of a difference when assessing symptom severity ($p=0.77$, required information size not reached). All trials were assessed at high risk of bias (Supplementary material 5) and the certainty of evidence was assessed as “very low” for all outcomes (Supplementary material 17).

It was only possible to perform two pre-planned meta-analyses: one assessing the effects of shorter- versus longer-term cognitive behavioral therapy for anxiety disorders at end of treatment and at maximum follow-up, and another one assessing the effects of shorter- versus longer-term psychodynamic psychotherapy for mood and anxiety disorders at end of treatment.

Shorter- versus longer-term cognitive behavioural therapy for anxiety disorders

We identified five trials assessing the effects of shorter-versus longer-term cognitive behavioural therapy for anxiety disorders [36–38, 42, 48]. All trials were assessed as at high risk of bias (Supplementary material 5). One trial was not eligible for meta-analysis, as the results were not reported in a usable way; i.e. the results were reported on a graph and standard deviations were not reported [42].

Four trials randomising a total of 209 participants reported on anxiety symptoms [36–38, 48]. Four different symptom scales were used: Beck Anxiety Inventory (BAI) [37], Social Phobia Anxiety Inventory – Social Phobia [36], State Trait Anxiety Inventory-Trait (STAI-T) [48], and Panic and Agoraphobia Scale (PAS) [38]. One trial included participants with social anxiety disorder [36]. Two trials included participants with panic disorder [37, 48]. One trial included participants with panic disorder

and agoraphobia [38]. We chose to analyse anxiety symptoms using SMD.

Meta-analysis of anxiety symptoms at end of treatment

Random-effects meta-analysis showed no evidence of a difference between shorter (5, 6, 7, 12 weeks) and longer-term (12, 12, 14, 18 weeks) cognitive behavioural therapy for anxiety disorders (including social anxiety disorder, panic disorder, and panic disorder with agoraphobia) on anxiety symptoms at end of treatment (SMD: 0.08; 95% CI: -0.47 to 0.63; $p=0.77$; $I^2=73\%$; four trials; very low certainty) (Fig. 2). Visual inspection of the forest plot and measures to quantify heterogeneity indicated substantial heterogeneity ($I^2=73\%$). The end of treatment assessment time point was 12 weeks [37, 48], 15 weeks [38], and 18 weeks [36]. It was not possible to assess the possible impact of missing outcome data, due to unclear or lack of reporting of number of analysed participants in some of the included trials. It was not possible to perform Trial Sequential Analysis for this outcome, because the outcome was assessed using SMD [24]. This outcome result was assessed as at high risk of bias. Certainty of the evidence was assessed as ‘very low’. See Supplementary material 17. The fixed-effect meta-analysis showed similar results (SMD 0.16; 95% CI: -0.11, 0.44; $p=0.25$; $I^2=73\%$; four trials; very low certainty) Supplementary material 18.

Shorter- versus longer-term psychodynamic therapy for mood and anxiety disorders

We identified two trials assessing the effects of shorter-versus longer-term psychodynamic therapy for mood-and anxiety disorder [34, 35, 54]. Both trials were assessed as at high risk of bias (Supplementary material 4).

Two trials randomising a total of 393 participants reported on level of functioning [34, 35]. Two different assessment scales were used, including Global Assessment of Functioning – Function (GAF-F) [35] and the work subscale (SAS-Work) of the Social Adjustment Scale [34]. We chose to analyze level of functioning using standardised mean difference. In order to assure the

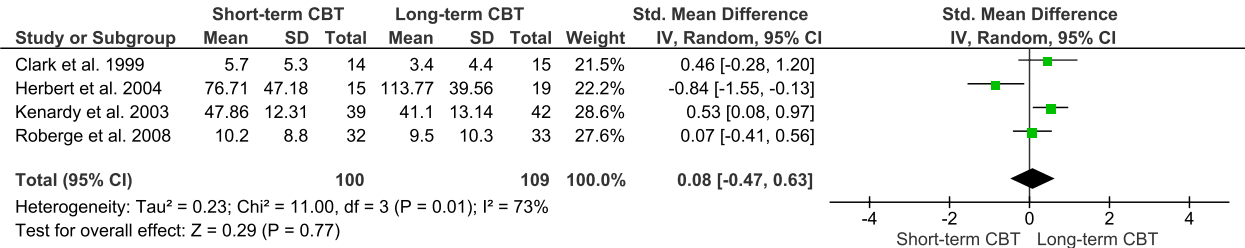


Fig. 2 Forest plot of shorter- versus longer-term cognitive behavioural therapy for anxiety disorders on severity of anxiety symptoms at end of treatment

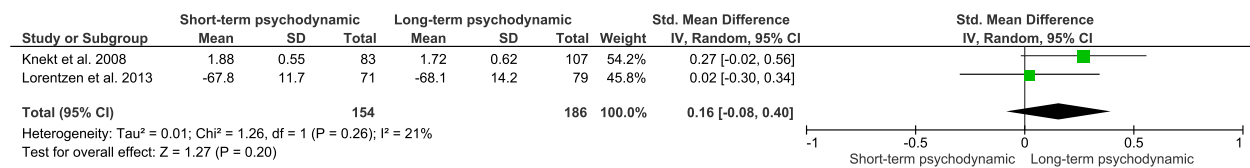


Fig. 3 Forest plot of shorter- versus longer-term psychodynamic therapy for mood- and anxiety disorders on level of functioning at end of treatment

scales pointed in the right direction, we multiplied the mean in one of the trials with ‘-1’.

Meta-analysis of level of functioning at end of treatment

Random effects meta-analysis showed no evidence of a difference between shorter- (20 and 20 weeks) and longer-term (80 and 156 weeks) psychodynamic psychotherapy for mood and anxiety disorders on level of functioning at end of treatment (SMD 0.16; 95% CI -0.08 to 0.40; $p=0.20$; $I^2=21\%$; two trials; very low certainty) (Fig. 3). Visual inspection of forest plot and measures to quantify heterogeneity ($I^2=21\%$) showed some heterogeneity. The end of treatment time point of assessment was 36 months after randomisation for both trials. It was not possible to perform Trial Sequential Analysis for this outcome, because the outcome was assessed using SMD [24]. This outcome result was assessed as at high risk of bias. Certainty of the evidence was assessed as ‘very low’. See Supplementary material 8. The fixed-effect meta-analysis showed similar results (SMD 0.16; 95% CI: -0.05, 0.37; $p=0.14$; $I^2=21\%$; two trials; very low certainty) Supplementary material 19.

Incomplete outcome data

Random effects meta-analysis of the best–worst case scenario adding 2 SD (SMD -0.16; 95% CI -8.13 to 7.81; $p < 0.00001$; $I^2=95\%$) and adding 1 SD (SMD -0.15; 95% CI -4.26 to 3.95; $p < 0.94$; $I^2=100\%$) for missing data showed no evidence of a difference between shorter- and longer-term psychodynamic psychotherapy. Random effects meta-analysis of the worst-best case scenario adding 2 SD (SMD -0.14; 95% CI -7.62 to 7.35; $p < 0.97$; $I^2=100\%$) and adding 1 SD (SMD -0.14; 95% CI -3.76 to 3.48; $p < 0.94$; $I^2=100\%$) for missing values showed no evidence of a difference between shorter- and longer-term psychodynamic psychotherapy.

Because of lack of relevant data, it was not possible to conduct other pre-defined meta-analyses. It was only possible to perform one sensitivity analysis (best–worst worst-best scenarios) to assess the potential impact of incomplete outcome data. We also planned several subgroup analyses to test for heterogeneity [6], but it was not possible to conduct them because of lack of relevant data. Further, it was not possible to assess the risk of

publication bias by testing for funnel plot asymmetry due to lack of trials. Last, it was not possible to perform Trial Sequential Analyses because all included outcomes were assessed using SMD.

The possible contribution of ongoing trials

We identified two ongoing trials [55, 56] that might contribute to the current evidence on shorter- versus longer-term psychotherapy for adult mental health disorders. These ongoing trials will contribute to the evidence on quality of life, serious adverse events, symptom severity, suicide and suicide attempts, self-harm, and level of functioning.

Discussion

We conducted the first systematic review assessing the difference between shorter- and longer-term psychotherapy for adult mental health disorders. We included 19 trials randomising a total of 3,447 participants to a shorter or a longer-term version of the same psychotherapy type. All trials and outcome results were at high risk of bias, and the certainty of the evidence according to GRADE was ‘very low’ for all outcomes.

One single trial showed no evidence of a difference between shorter- versus longer-term dialectical behavioral therapy for borderline personality disorder and reached the required information size needed to confirm or reject realistic intervention effects when assessing quality of life, symptom severity, and level of functioning [53] (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished). One single trial showed evidence of a beneficial effect of adding booster sessions to 8 and 12 weeks of internet-based cognitive-behavioral therapy when assessing symptom severity and level of functioning and reached the required information size needed to confirm or reject realistic intervention effects [50]. One single trial showed no evidence of a difference between shorter- versus longer-term psychodynamic psychotherapy for mood- or anxiety disorders and reached the required information size needed to confirm or reject realistic intervention effects when assessing symptom severity and level of

functioning [34]. The remaining single trials did not meet the required information size needed to confirm or reject realistic intervention effects. It was only possible to perform two pre-planned meta-analyses. Meta-analysis showed no evidence of a difference between short-term and long-term cognitive behavioural therapy for anxiety symptoms at end of treatment or at maximum follow-up. Meta-analysis showed no evidence of a difference between short-term and long-term psychodynamic psychotherapy on level of functioning at end of treatment. All trials and outcomes were assessed as at high risk of bias, and the certainty of evidence was assessed as ‘very low’ for all outcomes. It was not possible to perform Trial Sequential Analysis or tests for publication bias. Further, due to poor reporting in the included trials, we only performed one planned sensitivity analysis to assess the potential impact of missing data. Only one trial reported on serious adverse events (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished). Two trials reported on suicide and suicide attempts [35] (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished), and one trial reported on self-harm (McMain S: The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial, Unpublished).

Our review has several strengths. We followed our protocol which was registered prior to the systematic literature search (PROSPERO ID: CRD42019128535). Data were double-extracted by independent authors minimizing the risk of inaccurate data extraction, and we assessed the risk of bias in all trials according to Cochrane methodology [14]. We used GRADE to assess the certainty of the evidence [31–33], and the eight-step assessment suggested by Jakobsen et al. to assess if the thresholds for significance were crossed [17]. Hence, this systematic review considered both risks of random errors and risks of systematic errors which adds further robustness to our results and conclusions. Another strength of our review is that we pragmatically accepted any short-term psychotherapy type and any long-term psychotherapy type, thus results may therefore guide a clinician when choosing between different treatment durations.

Our review also has several limitations. First, due to large heterogeneity in participants, interventions, comparisons, and outcomes, we decided to primarily report the results narratively and only perform two small pre-planned meta-analyses. The observed heterogeneity is

due to our pre-defined broad inclusion criteria, i.e. we used the trialists’ own definitions of short-term and long-term psychotherapy. However, we believe that this choice of methodology from a pragmatic point of view is the best solution there is, as introducing specific thresholds may have excluded important data from our review [6]. If we had used a specific threshold distinguishing short-term from long-term psychotherapy, e.g. by applying a definition of short-term psychotherapy as including up to 24 sessions and long-term psychotherapy as including at least 50 sessions or having a duration of at least one year as suggested by others [57, 58], we would have only been able to include three trials in the review, and the aim of presenting a complete overview would not be possible. Second, all trials were at high risk of bias. Therefore, there is a risk that our results overestimated the beneficial effects and underestimated the harmful effects of the experimental interventions being studied [59–66]. Third, we only identified 19 trials, and it was not possible to assess the risk of random errors in the meta-analyses with Trial Sequential Analysis due to the inclusion of continuous outcomes assessed with heterogeneous measures (i.e. we assessed the effects with standardised mean difference). This is a major limitation, as we cannot assess if the shown lack of difference is an indication of a “true” lack of difference, or if it is an indication that more trials are needed. We calculated the required information sizes for single trial results post-hoc, but these should primarily be considered exploratory, as they rely on the observed means and standard deviations instead of pre-defined minimal clinically important differences on the assessed scales. Fourth, only few trials reported on serious adverse events, suicide, suicide attempts, and self-harm. It is of utmost importance to always assess beneficial *and* harmful intervention effects on patient-important outcomes [14, 67].

We have identified one previous systematic review comparing short-term and long-term psychotherapy for schizophrenia [68]. However, the review did not identify any trials. We have also identified a meta-regression study investigating the effects of psychotherapy for major depressive disorder [5]. This study found no significant association between the duration of psychotherapy and effect-size, which is similar to the conclusion of the present review. However, in the meta-regression study, there was a strong association between number of sessions per week and effect size. An increase from one to two sessions per week increased the effect size with $g=0.45$, while keeping the total number of treatment sessions constant [5]. The results of the present review could neither confirm nor reject that two sessions per week were more efficacious than one session per week.

The included trials in this review typically assessed the effects of different durations of psychotherapy for anxiety disorders, major depressive disorder, and post-traumatic stress disorder. Our findings indicate that there may be no evidence of a difference between short-term and long-term psychotherapy when assessing symptom severity and level of functioning. There are, however, indications from non-controlled studies that patients with complex and severe psychopathology, defined by the presence of, e.g., co-occurring mental health disorders, longer duration and early onset of the disorder, and unemployment, may have better outcomes in high-intensity than in low-intensity treatments [69, 70]. We included one trial including participants with borderline personality disorder. This trial did not find evidence of a difference between six versus 12 months dialectical behavioral therapy, and the trial reached the required information size needed to confirm or reject realistic intervention effects for quality of life, symptom severity, and level of functioning. However, the trial was assessed as at high risk of bias and the certainty of evidence was “very low” for all outcomes. Accordingly, future randomised clinical trials comparing the outcomes of short- and long-term psychotherapy for patients with low and high problem complexity should be conducted. We are currently performing a similar randomised clinical trial assessing the effects of five months versus 14 months of mentalization-based therapy for borderline personality disorder [55, 71]. We are planning a protocol for an individual patient data meta-analysis of shorter- versus longer-term psychotherapy for borderline personality disorder, which will be conducted once data from the two trials become available. Results of the individual patient data meta-analysis will increase the possibility of identifying subgroups of participants with specific effects of the assessed interventions. We identified no trials including participants with other severe personality pathology, schizophrenia, or other psychotic disorders. Hence, it is still unclear whether patients with severe psychopathology requires short-term or long-term psychotherapy.

Evidence-based practice and decision-making should be based on the best available evidence, patient preferences, and the clinician's expertise [72]. For severe and complex cases there is evidence of beneficial effects of psychotherapy of specific treatment lengths (e.g. long-term specialized treatment for borderline personality disorder [73]) but very low certainty evidence to guide clinicians in choosing the optimal treatment duration. Evidently, clinicians should by default offer psychotherapy in a duration supported by the best available evidence. But when there is a question of treatment duration, e.g. a patient asking for a shorter treatment because of life circumstances, the clinician is advised to balance this

preference with clinical experience which may include knowledge of specific prognostic factors such as early onset or co-occurring disorders, while also considering the poor evidence regarding the optimal treatment duration currently available.

Conclusions

The evidence for shorter- versus longer-term psychotherapy for adult mental health disorders is currently unclear. We only identified 19 randomised clinical trials. More trials at low risk of bias and at low risk of random errors assessing participants at different levels of psychopathological severity are urgently needed.

Differences between the protocol and the review

In addition to assessing all outcomes at end of treatment, we planned to assess all outcomes at maximum follow-up as a secondary analysis. However, only few trials reported data at maximum follow-up. Because of lack of relevant data, we chose to only report data at end of treatment.

Abbreviations

CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence Interval
CPCI-S	Conference Proceedings Citation Index—Science
CPCI-SSH	Conference Proceedings Citation Index—Social Science & Humanities
DARIS	Diversity Adjusted Required Information Size
EMBASE	Excerpta Medica database
GRADE	Grades of Recommendation, Assessment, Development, and Evaluation
LILACS	Latin American and Caribbean Health Sciences Literature
MEDLINE	Medical Literature Analysis and Retrieval System Online
MD	Mean Difference
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
ROB	Risk of Bias
RR	Risk Ratio
SCI-EXPANDED	Science Citation Index Expanded
SD	Standard Deviation
SMD	Standardised Mean Difference
SSCI	Social Sciences Citation Index
TSA	Trial Sequential Analysis

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12888-023-04895-6>.

Additional file 1. PRISMA 2020 Checklist for “Short-term versus long-term psychotherapy for adult psychiatric disorders: A systematic review with meta-analysis”

Additional file 2.

Additional file 3.

Additional file 4.

Additional file 5:Supplementary material 5. Risk of bias table.

Additional file 6.

Additional file 7.

Additional file 8.**Additional file 9.****Additional file 10.****Additional file 11.****Additional file 12.****Additional file 13.****Additional file 14.****Additional file 15.****Additional file 16.****Additional file 17.**

Additional file 18:Supplementary material 18. Fixed-effect meta-analysis of short-term versus long-term cognitivebehavioral therapy for anxiety disorders on anxiety symptoms (sensitivity analysis).

Additional file 19:Supplementary material 19. Fixed-effect meta-analysis of short-term versus long-term psychodynamic psychotherapy for mood- and anxiety disorders on level of functioning (sensitivity analysis).

Acknowledgements

The expert help from Sarah Louise Klingenberg (Information Specialist, The Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Copenhagen, Denmark) in making the search strategy and conducting electronic searches is hugely appreciated.

Authors' contributions

SJ and CKJ independently assessed eligibility, extracted data, and performed risk of bias assessments with ongoing supervision from JCJ and SS. SJ conducted data analyses and wrote up the manuscript draft. All authors read, commented on, and approved the final manuscript.

Funding

Open access funding provided by Royal Danish Library. This work was funded by Trygfonden (SJ, grant number: 123488), and by the Mental Health Services in the Capital Region of Denmark Research Fund (SJ and SS, grant number: N/A).

Availability of data and materials

All data generated or analysed during this study are included in this published article (and its supplementary information files).

Declarations**Ethical approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Gentofte, Denmark. ²Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital – Rigshospitalet, The Capital Region of Denmark, Copenhagen, Denmark. ³Department of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. ⁴Department of Psychology, University of Copenhagen, Copenhagen, Denmark.

Received: 6 September 2022 Accepted: 22 May 2023

Published online: 16 June 2023

References

- Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jonsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011;21(9):655–79.
- Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jonsson B, group Cs, et al. The economic cost of brain disorders in Europe. *Eur J Neurol*. 2012;19(1):155–62.
- Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe—a critical review and appraisal of 27 studies. *Eur Neuropsychopharmacol*. 2005;15(4):357–76.
- Barlow DH. *The Oxford Handbook of Clinical Psychology: Updated Edition*. Oxford Library of Psychology. Oxford: Oxford University Press; 2014.
- Cuijpers P, Huibers M, Ebert DD, Koole SL, Andersson G. How much psychotherapy is needed to treat depression? A meta-regression analysis. *J Affect Disord*. 2013;149(1–3):1–13.
- Juul S, Poulsen S, Lunn S, Sorensen P, Jakobsen JC, Simonsen S. Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Syst Rev*. 2019;8(1):169.
- Castonguay LG, Barkham M, Lutz W, McAleavey A. Practice-oriented research: Approaches and applications. In: Lambert MJ, editors. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 6th edition. Hoboken: Wiley; 2013. p. 85–133.
- Howard KI, Kopta SM, Krause MS, Orlinsky DE. The Dose-Effect Relationship in Psychotherapy. *Am Psychol*. 1986;41(2):159–64.
- Robinson L, Delgadillo J, Kellett S. The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychother Res*. 2020;30(1):79–96.
- Bone C, Delgadillo J, Barkham M. A systematic review and meta-analysis of the good-enough level (GEL) literature. *Journal of Counseling Psychology*. 2021;68(2):219.
- Jakobsen JC, Gluud C. The necessity of randomized clinical trials. *Br J Med Res*. 2013;3(4):1453–68.
- Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, et al. Evaluating non-randomised intervention studies. *Health Technol Assess*. 2003;7(27):iii–x, 1–173.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 2021;10(1):89.
- Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V. *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.1). The Cochrane Collaboration. Wiley. 2011. Available at: <https://handbook-5-1.cochrane.org/>. Accessed 2 Mar 2023.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Guideline: Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice (ICH-GCP). 2015. Available at: <https://ichgcp.net/da>. Accessed 2 Mar 2023.
- Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error matrix approach for overviewing available evidence. *BMC Med Res Methodol*. 2010;10(1):90.
- Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14(1):120.
- Stata Statistical Software: Release 16. College Station: StataCorp LLC; 2019. Available at: <http://www.stata.com>. Accessed May 26 2023.
- The Cochrane Collaboration. Review manager (Version 5.4) [Computer Software]. The Nordic Cochrane Centre; 2020. Available at: <https://rev-man.cochrane.org>. Accessed 2 Mar 2023.
- Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol*. 2008;61(1):64–75.
- Copenhagen Trial Unit. TSA - Trial Sequential Analysis. [Computer Software] Available at: <http://www.ctu.dk/tsa>. Accessed 2 Mar 2023.
- Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol*. 2008;61(8):763–9.
- Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive—trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *Int J Epidemiol*. 2008;38(1):287–98.

24. Thorlund K, Engström J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA). 2011. Available at: https://ctu.dk/wp-content/uploads/2021/03/2017-10-10-TSA-Manual-ENG_ER.pdf. Accessed 2 Mar 2023.
25. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol*. 2009;9(1):86.
26. Thorlund K, Anema A, Mills E. Interpreting meta-analysis according to the adequacy of sample size. An example using isoniazid chemoprophylaxis for tuberculosis in purified protein derivative negative HIV-infected individuals. *Clin Epidemiol*. 2010;2:57.
27. Thorlund K, Devereaux P, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *Int J Epidemiol*. 2008;38(1):276–86.
28. Imberger G, Thorlund K, Gluud C, Wetterslev J. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. *BMJ Open*. 2016;6(8):e011890.
29. Norman GR, Sloan JA, Wywich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582–92.
30. Jakobsen JC, Wetterslev J, Lange T, Gluud C. Viewpoint: taking into account risks of random errors when analysing multiple outcomes in systematic reviews. *Cochrane Database of Syst Rev*. 2016(3)ED000111. <https://doi.org/10.1002/14651858.ED000111>.
31. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol*. 2011;64(4):380–2.
32. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6.
33. Schünemann HJ, Best D, Vist G, Oxman AD, Group GW. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *CMAJ*. 2003;169(7):677–80.
34. Knekt P, Lindfors O, Hakanen T, Valikoski M, Virtala E, Laaksonen MA, et al. Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*. 2008;38(5):689–703.
35. Lorentzen S, Ruud T, Fjeldstad A, Hoglend P. Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial. *Br J Psychiatry*. 2013;203(3):280–7.
36. Herbert JD, Rheingold AA, Gaudiano BA, Myers VH. Standard Versus Extended Cognitive Behavior Therapy for Social Anxiety Disorder: A Randomized-Controlled Trial. *Behav Cognitive Psychother*. 2004;32(2):131–47.
37. Clark DM, Salkovskis PM, Hackmann A, Wells A, Ludgate J, Gelder M. Brief Cognitive Therapy for Panic Disorder: A Randomized Controlled Trial. *J Consult Clin Psychol*. 1999;67(4):583–9.
38. Roberge P, Marchand A, Reinharz D, Savard P. Cognitive-Behavioral Treatment for Panic Disorder With Agoraphobia A Randomized, Controlled Trial and Cost-Effectiveness Analysis. *J Behavior Mod*. 2008;32(3):333–51.
39. Bruijninks SJE, Lemmens L, Hollon SD, Peeters F, Cuijpers P, Arntz A, et al. The effects of once- versus twice-weekly sessions on psychotherapy outcomes in depressed patients. *Br J Psychiatry*. 2020;216(4):222–30.
40. Barkham M, Rees A, Shapiro DA, Stiles WB, Agnew RM, Halstead J, et al. Outcomes of time-limited psychotherapy in applied settings: replicating the Second Sheffield Psychotherapy Project. *J Consult Clin Psychol*. 1996;64(5):1079.
41. Shapiro DA, Barkham M, Rees A, Hardy GE, Reynolds S, Startup M. Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *J Consult Clin Psychol*. 1994;62(3):522.
42. Bohni M, Spindler H, Arendt M, Hougaard E, Rosenberg N. A randomized study of massed three-week cognitive behavioural therapy schedule for panic disorder. *Acta Psychiatr Scand*. 2009;120(3):187–95.
43. Christensen H, Griffiths K, Mackinnon A, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. *Psychol Med*. 2006;36(12):1737.
44. Dekker J, Molenaar PJ, Kool S, Van Aalst G, Peen J, de Jonghe F. Dose-effect relations in time-limited combined psycho-pharmacological treatment for depression. *Psychol Med*. 2005;35(1):47–58.
45. Ehlers A, Hackmann A, Grey N, Wild J, Liness S, Albert I, et al. A randomized controlled trial of 7-day intensive and standard weekly cognitive therapy for PTSD and emotion-focused supportive therapy. *Am J Psychiatry*. 2014;171(3):294–304.
46. Foa EB, McLean CP, Zang Y, Rosenfield D, Yadin E, Yarvis JS, et al. Effect of prolonged exposure therapy delivered over 2 weeks vs 8 weeks vs present-centered therapy on PTSD symptom severity in military personnel: A randomized clinical trial. *JAMA*. 2018;319(4):354–64.
47. Nacasch N, Huppert JD, Su Y-J, Kivity Y, Dinshtein Y, Yeh R, et al. Are 60-Minute Prolonged Exposure Sessions With 20-Minute Imaginal Exposure to Traumatic Memories Sufficient to Successfully Treat PTSD? A Randomized Noninferiority Clinical Trial. *Behav Ther*. 2015;46(3):328–41.
48. Kenardy JA, Dow MG, Johnston DW, Newman MG, Thomson A, Taylor CB. A comparison of delivery methods of cognitive-behavioral therapy for panic disorder: an international multicenter trial. *J Consult Clin Psychol*. 2003;71(6):1068.
49. Dell L, Sbisá AM, Forbes A, O'Donnell M, Bryant R, Hodson S, et al. Effect of massed v. standard prolonged exposure therapy on PTSD in military personnel and veterans: a non-inferiority randomised controlled trial. *Psychol Med*. 2022;1–8. <https://doi.org/10.1017/S0033291722000927>.
50. Hadjistavropoulos H, Peynenburg V, Thiessen D, Nugent M, Karin E, Dear B, et al. A randomized factorial trial of internet-delivered cognitive behavioural therapy: An 8-week program with or without extended support and booster lesson. *Internet Int*. 2022;27:100499.
51. Böttche M, Wagner B, Vöhringer M, Heinrich M, Stein J, Selmo P, et al. Is only one cognitive technique also effective? Results from a randomized controlled trial of two different versions of an internet-based cognitive behavioural intervention for post-traumatic stress disorder in Arabic-speaking countries. *Eur J Psychotraumatol*. 2021;12(1):1943870.
52. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022;91(6):382–97.
53. McMain SF, Chapman AL, Kuo JR, Guimond T, Streiner DL, Dixon-Gordon KL, et al. The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial protocol. *BMC Psychiatry*. 2018;18(1):230.
54. Knekt P, Lindfors O, Laaksonen MA, Raitasalo R, Haaramo P, Jarvikoski A, et al. Effectiveness of short-term and long-term psychotherapy on work ability and functional capacity—a randomized clinical trial on depressive and anxiety disorders. *J Affect Disord*. 2008;107(1–3):95–106.
55. Juul S, Lunn S, Poulsen S, Sorensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196.
56. Kool M, Van HL, Bartak A, de Maat SCM, Arntz A, van den Eshof JW, et al. Optimizing psychotherapy dosage for comorbid depression and personality disorders (PsyDos): a pragmatic randomized factorial trial using schema therapy and short-term psychodynamic psychotherapy. *BMC Psychiatry*. 2018;18(1):252.
57. Gabbard, G. Long-term Psychodynamic Psychotherapy: A Basic Text. Arlington: American Psychiatric Publishing Inc; 2004.
58. Crits-Christoph P, Barber JP. Long-term psychotherapy. In: Ingram RE, Snyder CR, editors. *Handbook of Psychological Change: Psychotherapy Processes & Practices for the 21st Century*. Hoboken: John Wiley & Sons; 2000. p. 455–73.
59. Gluud LL. Bias in clinical intervention research. *Am J Epidemiol*. 2006;163(6):493–501.
60. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med*. 2001;135(11):982–9.
61. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ*. 2001;323(7303):42–6.
62. Moher D, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998;352(9128):609–13.
63. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273(5):408–12.
64. Hrobjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol*. 2014;43(4):1272–83.

65. Hrobjartsson A, Skou Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ*. 2013;185(4):E201-11.
66. Hrobjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ*. 2012;344:e1119.
67. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V. *Cochrane Handbook for Systematic Reviews of Interventions* (Version 6.3). The Cochrane Collaboration. Wiley; 2019. Available at: <https://training.cochrane.org/handbook>. Accessed 2 Mar 2023.
68. Naeem F, Farooq S, Kingdon D. Cognitive behavioural therapy (brief versus standard duration) for schizophrenia. *Cochrane Database Syst Rev*. 2015;10:CD010646.
69. Delgadillo J, Moreea O, Lutz W. Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behav Res Ther*. 2016;79:15–22.
70. Delgadillo J, Huey D, Bennett H, McMillan D. Case complexity as a guide for psychological treatment selection. *J Consult Clin Psychol*. 2017;85(9):835.
71. Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, et al. Detailed statistical analysis plan for the short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder randomized clinical trial (MBT-RCT). *Trials*. 2020;22:1–8.
72. Sackett DL, Rosenberg WM, Gray JM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;321:71–2.
73. Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020(5). <https://doi.org/10.1002/14651858.CD012955.pub2>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



PRISMA 2020 Checklist for “Short-term versus long-term psychotherapy for adult psychiatric disorders: A systematic review with meta-analysis”

Item	Place/location in manuscript
1	Title
2	Abstract
3	Introduction
4	Introduction
5	Inclusion and exclusion criteria
6	Search strategy and selection criteria
7	Search strategy and selection criteria and Supplementary material 2
8	Search strategy and selection criteria
9	Data extraction and risk of bias assessment
10a	Outcomes and subgroup analyses
10b	Protocol
11	Data extraction and risk of bias assessment
12	Outcomes and subgroup analyses
13a	Assessment of statistical and clinical significance and protocol
13b	Assessment of statistical and clinical significance and protocol
13c	Assessment of statistical and clinical significance and protocol
13d	Assessment of statistical and clinical significance and protocol
13e	Assessment of statistical and clinical significance and protocol
13f	Assessment of statistical and clinical significance and protocol
14	Assessment of statistical and clinical significance and protocol
15	Assessment of statistical and clinical significance and protocol
16a	Study characteristics
16b	Supplementary material 4 – list of excluded studies
17	Study characteristics and Table 1
18	Supplementary material 5
19	Single trial results
20a	Single trial results and meta-analysis sections
20b	Meta-analysis sections
20c	Meta-analysis sections
20d	Incomplete outcome data
21	N/A
22	Supplementary materials

23a	Discussion
23b	Discussion
23c	Discussion
23d	Discussion
24a	Abstract and Methods
24b	Methods
24c	N/A
25	Funding
26	Competing interests
27	Availability of data and materials

PRISMA 2020 expanded checklist

Note: This expanded checklist details elements recommended for reporting for each PRISMA 2020 item. Non-italicized elements are considered 'essential' and should be reported in the main report or as supplementary material for all systematic reviews (except for those preceded by "If...", which should only be reported where applicable). Elements written in italics are 'additional', and while not essential, provide supplementary information that may enhance the completeness and usability of systematic review reports. Note that elements presented here are an abridged version of those presented in the explanation and elaboration paper, with references and some examples removed. Consulting the explanation and elaboration paper is recommended if further clarity or information is required (doi:10.1136/bmj.n160).

Section and Topic	Item #	Elements recommended for reporting
TITLE		
TITLE	1	- Identify the report as a systematic review in the title. - Report an informative title that provides key information about the main objective or question the review addresses (e.g. the population(s) and intervention(s) the review addresses). <i>Consider providing additional information in the title, such as the method of analysis used, the designs of included studies, or an indication that the review is an update of an existing review, or a continually updated ("living") systematic review.</i>
ABSTRACT		
ABSTRACT	2	Report an abstract addressing each item in the PRISMA 2020 for Abstracts checklist.
INTRODUCTION		
RATIONALE	3	- Describe the current state of knowledge and its uncertainties. - Articulate why it is important to do the review. - If other systematic reviews addressing the same (or a largely similar) question are available, explain why the current review was considered necessary. If the review is an update or replication of a particular systematic review, indicate this and cite the previous review. - If the review examines the effects of interventions, also briefly describe how the intervention(s) examined might work. <i>If there is complexity in the intervention or context of its delivery (or both) (e.g. multi-component interventions, equity considerations), consider presenting a logic model to visually display the hypothesised relationship between intervention components and outcomes.</i>
OBJECTIVES	4	- Provide an explicit statement of all objective(s) or question(s) the review addresses, expressed in terms of a relevant question formulation framework. - If the purpose is to evaluate the effects of interventions, use the Population, Intervention, Comparator, Outcome (PICO) framework or one of its variants, to state the comparisons that will be made.
METHODS		
ELIGIBILITY CRITERIA	5	- Specify all study characteristics used to decide whether a study was eligible for inclusion in the review, that is, components described in the PICO framework or one of its variants, and other characteristics, such as eligible study design(s) and setting(s), and minimum duration of follow-up. - Specify eligibility criteria with regard to report characteristics, such as year of dissemination, language, and report status (e.g. whether reports, such as unpublished manuscripts and conference abstracts, were eligible for inclusion). - Clearly indicate if studies were ineligible because the outcomes of interest were not measured, or ineligible because the results for the outcome of interest were not reported. - Specify any groups used in the synthesis (e.g. intervention, outcome and population groups) and link these to the comparisons specified in the objectives (item #4). <i>Consider providing rationales for any notable restrictions to study eligibility.</i>
INFORMATION SOURCES	6	- Specify the date when each source (e.g. database, register, website, organisation) was last searched or consulted. - If bibliographic databases were searched, specify for each database its name (e.g. MEDLINE, CINAHL), the interface or platform through which the database was searched (e.g. Ovid, EBSCOhost), and the dates of coverage (where this information is provided). - If study registers, regulatory databases and other online repositories were searched, specify the name of each source and any date restrictions that were applied.

Section and Topic	Item #	Elements recommended for reporting
		• If websites, search engines or other online sources were browsed or searched, specify the name and URL of each source. • If organisations or manufacturers were contacted to identify studies, specify the name of each source. • If individuals were contacted to identify studies, specify the types of individuals contacted (e.g. authors of studies included in the review or researchers with expertise in the area). • If reference lists were examined, specify the types of references examined (e.g. references cited in study reports included in the systematic review, or references cited in systematic review reports on the same or similar topic). • If cited or citing reference searches (also called backward and forward citation searching) were conducted, specify the bibliographic details of the reports to which citation searching was applied, the citation index or platform used (e.g. Web of Science), and the date the citation searching was done. • If journals or conference proceedings were consulted, specify of the names of each source, the dates covered and how they were searched (e.g. handsearching or browsing online).
SEARCH STRATEGY	7	• Provide the full line by line search strategy as run in each database with a sophisticated interface (such as Ovid), or the sequence of terms that were used to search simpler interfaces, such as search engines or websites. • Describe any limits applied to the search strategy (e.g. date or language) and justify these by linking back to the review's eligibility criteria. • If published approaches, including search filters designed to retrieve specific types of records or search strategies from other systematic reviews, were used, cite them. If published approaches were adapted, for example if search filters are amended, note the changes made. • If natural language processing or text frequency analysis tools were used to identify or refine keywords, synonyms or subject indexing terms to use in the search strategy, specify the tool(s) used. • If a tool was used to automatically translate search strings for one database to another, specify the tool used. • If the search strategy was validated, for example by evaluating whether it could identify a set of clearly eligible studies, report the validation process used and specify which studies were included in the validation set. • If the search strategy was peer reviewed, report the peer review process used and specify any tool used such as the Peer Review of Electronic Search Strategies (PRESS) checklist. • If the search strategy structure adopted was not based on a PICO-style approach, describe the final conceptual structure and any explorations that were undertaken to achieve it.
SELECTION PROCESS	8	<i>Recommendations for reporting regardless of the selection processes used:</i> • Report how many reviewers screened each record (title/abstract) and each report retrieved, whether multiple reviewers worked independently at each stage of screening or not, and any processes used to resolve disagreements between screeners. • Report any processes used to obtain or confirm relevant information from study investigators. • If abstracts or articles required translation into another language to determine their eligibility, report how these were translated. <i>Recommendations for reporting in systematic reviews using automation tools in the selection process:</i> • Report how automation tools were integrated within the overall study selection process. • If an externally derived machine learning classifier was applied (e.g. Cochrane RCT Classifier), either to eliminate records or to replace a single screener, include a reference or URL to the version used. If the classifier was used to eliminate records <i>before screening</i>, report the number eliminated in the PRISMA flow diagram as 'Records marked as ineligible by automation tools'. • If an internally derived machine learning classifier was used to assist with the screening process, identify the software/classifier and version, describe how it was used (e.g. to remove records or replace a single screener) and trained (if relevant), and what internal or external validation was done to understand the risk of missed studies or incorrect classifications. • If machine learning algorithms were used to prioritise screening (whereby unscreened records are continually re-ordered based on screening decisions), state the software used and provide details of any screening rules applied. <i>Recommendations for reporting in systematic reviews using crowdsourcing or previous 'known' assessments in the selection process:</i>

Section and Topic	Item #	Elements recommended for reporting
		• If crowdsourcing was used to screen records, provide details of the platform used and specify how it was integrated within the overall study selection process. • If datasets of already-screened records were used to eliminate records retrieved by the search from further consideration, briefly describe the derivation of these datasets.
DATA COLLECTION PROCESS	9	• Report how many reviewers collected data from each report, whether multiple reviewers worked independently or not, and any processes used to resolve disagreements between data collectors. • Report any processes used to obtain or confirm relevant data from study investigators. • If any automation tools were used to collect data, report how the tool was used, how the tool was trained, and what internal or external validation was done to understand the risk of incorrect extractions. • If articles required translation into another language to enable data collection, report how these articles were translated. • If any software was used to extract data from figures, specify the software used. • If any decision rules were used to select data from multiple reports corresponding to a study, and any steps were taken to resolve inconsistencies across reports, report the rules and steps used.
DATA ITEMS (outcomes)	10a	• List and define the outcome domains and time frame of measurement for which data were sought. • Specify whether all results that were compatible with each outcome domain in each study were sought, and if not, what process was used to select results within eligible domains. • If any changes were made to the inclusion or definition of the outcome domains, or to the importance given to them in the review, specify the changes, along with a rationale. • If any changes were made to the processes used to select results within eligible outcome domains, specify the changes, along with a rationale. • <i>Consider specifying which outcome domains were considered the most important for interpreting the review's conclusions and provide rationale for the labelling (e.g. "a recent core outcome set identified the outcomes labelled 'critical' as being the most important to patients").</i>
DATA ITEMS (other variables)	10b	• List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). • Describe any assumptions made about any missing or unclear information from the studies. • If a tool was used to inform which data items to collect, cite the tool used.
STUDY RISK OF BIAS ASSESSMENT	11	• Specify the tool(s) (and version) used to assess risk of bias in the included studies. • Specify the methodological domains/components/items of the risk of bias tool(s) used. • Report whether an overall risk of bias judgement that summarised across domains/components/items was made, and if so, what rules were used to reach an overall judgement. • If any adaptations to an existing tool to assess risk of bias in studies were made, specify the adaptations. • If a new risk of bias tool was developed for use in the review, describe the content of the tool and make it publicly accessible. • Report how many reviewers assessed risk of bias in each study, whether multiple reviewers worked independently, and any processes used to resolve disagreements between assessors. • Report any processes used to obtain or confirm relevant information from study investigators. • If an automation tool was used to assess risk of bias, report how the automation tool was used, how the tool was trained, and details on the tool's performance and internal validation.
EFFECT MEASURES	12	• Specify for each outcome (or type of outcome [e.g. binary, continuous]), the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results. • State any thresholds (or ranges) used to interpret the size of effect (e.g. minimally important difference; ranges for no/trivial, small, moderate and large effects) and the rationale for these thresholds.

Section and Topic	Item #	Elements recommended for reporting
		- If synthesized results were re-expressed to a different effect measure, report the method used to re-express results (e.g. meta-analysing risk ratios and computing an absolute risk reduction based on an assumed comparator risk). <i>Consider providing justification for the choice of effect measure.</i>
SYNTHESIS METHODS (eligibility for synthesis)	13a	Describe the processes used to decide which studies were eligible for each synthesis.
SYNTHESIS METHODS (preparing for synthesis)	13b	Report any methods required to prepare the data collected from studies for presentation or synthesis, such as handling of missing summary statistics, or data conversions.
SYNTHESIS METHODS (tabulation and graphical methods)	13c	- Report chosen tabular structure(s) used to display results of individual studies and syntheses, along with details of the data presented. - Report chosen graphical methods used to visually display results of individual studies and syntheses. <i>If studies are ordered or grouped within tables or graphs based on study characteristics (e.g. by size of the study effect, year of publication), consider reporting the basis for the chosen ordering/grouping.</i> <i>If non-standard graphs were used, consider reporting the rationale for selecting the chosen graph.</i>
SYNTHESIS METHODS (statistical synthesis methods)	13d	- If statistical synthesis methods were used, reference the software, packages and version numbers used to implement synthesis methods. - If it was not possible to conduct a meta-analysis, describe and justify the synthesis methods or summary approach used. - If meta-analysis was done, specify: - the meta-analysis model (fixed-effect, fixed-effects or random-effects) and provide rationale for the selected model. - the method used (e.g. Mantel-Haenszel, inverse-variance). - any methods used to identify or quantify statistical heterogeneity (e.g. visual inspection of results, a formal statistical test for heterogeneity, heterogeneity variance (τ^2), inconsistency (e.g. I^2), and prediction intervals). - If a random-effects meta-analysis model was used: - specify the between-study (heterogeneity) variance estimator used (e.g. DerSimonian and Laird, restricted maximum likelihood (REML)). - specify the method used to calculate the confidence interval for the summary effect (e.g. Wald-type confidence interval, Hartung-Knapp-Sidik-Jonkman). <i>consider specifying other details about the methods used, such as the method for calculating confidence limits for the heterogeneity variance.</i> - If a Bayesian approach to meta-analysis was used, describe the prior distributions about quantities of interest (e.g. intervention effect being analysed, amount of heterogeneity in results across studies). - If multiple effect estimates from a study were included in a meta-analysis, describe the method(s) used to model or account for the statistical dependency (e.g. multivariate meta-analysis, multilevel models or robust variance estimation). - If a planned synthesis was not considered possible or appropriate, report this and the reason for that decision.
SYNTHESIS METHODS (methods to explore heterogeneity)	13e	- If methods were used to explore possible causes of statistical heterogeneity, specify the method used (e.g. subgroup analysis, meta-regression). - If subgroup analysis or meta-regression was performed, specify for each: - which factors were explored, levels of those factors, and which direction of effect modification was expected and why (where possible). - whether analyses were conducted using study-level variables (i.e. where each study is included in one subgroup only), within-study contrasts (i.e. where data on subsets of participants within a study are available, allowing the study to be included in more than one subgroup), or some combination of the above. - how subgroup effects were compared (e.g. statistical test for interaction for subgroup analyses). - If other methods were used to explore heterogeneity because data were not amenable to meta-analysis of effect estimates (e.g. structuring tables to examine variation in results across studies based on subpopulation), describe the methods used, along with the factors and levels. - If any analyses used to explore heterogeneity were not pre-specified, identify them as such.

Section and Topic	Item #	Elements recommended for reporting
SYNTHESIS METHODS (sensitivity analyses)	13f	- If sensitivity analyses were performed, provide details of each analysis (e.g. removal of studies at high risk of bias, use of an alternative meta-analysis model). - If any sensitivity analyses were not pre-specified, identify them as such.
REPORTING BIAS ASSESSMENT	14	- Specify the methods (tool, graphical, statistical or other) used to assess the risk of bias due to missing results in a synthesis (arising from reporting biases). - If risk of bias due to missing results was assessed using an existing tool, specify the methodological components/domains/items of the tool, and the process used to reach a judgement of overall risk of bias. - If any adaptations to an existing tool to assess risk of bias due to missing results were made, specify the adaptations. - If a new tool to assess risk of bias due to missing results was developed for use in the review, describe the content of the tool and make it publicly accessible. - Report how many reviewers assessed risk of bias due to missing results in a synthesis, whether multiple reviewers worked independently, and any processes used to resolve disagreements between assessors. - Report any processes used to obtain or confirm relevant information from study investigators. - If an automation tool was used to assess risk of bias due to missing results, report how the automation tool was used, how the tool was trained, and details on the tool's performance and internal validation.
CERTAINTY ASSESSMENT	15	- Specify the tool or system (and version) used to assess certainty (or confidence) in the body of evidence. - Report the factors considered (e.g. precision of the effect estimate, consistency of findings across studies) and the criteria used to assess each factor when assessing certainty in the body of evidence. - Describe the decision rules used to arrive at an overall judgement of the level of certainty, together with the intended interpretation (or definition) of each level of certainty. - If applicable, report any review-specific considerations for assessing certainty, such as thresholds used to assess imprecision and ranges of magnitude of effect that might be considered trivial, moderate or large, and the rationale for these thresholds and ranges (item #12). - If any adaptations to an existing tool or system to assess certainty were made, specify the adaptations. - Report how many reviewers assessed certainty in the body of evidence for an outcome, whether multiple reviewers worked independently, and any processes used to resolve disagreements between assessors. - Report any processes used to obtain or confirm relevant information from investigators. - If an automation tool was used to support the assessment of certainty, report how the automation tool was used, how the tool was trained, and details on the tool's performance and internal validation. - Describe methods for reporting the results of assessments of certainty, such as the use of Summary of Findings tables. - If standard phrases that incorporate the certainty of evidence were used (e.g. "hip protectors <i>probably</i> reduce the risk of hip fracture slightly"), report the intended interpretation of each phrase and the reference for the source guidance.
RESULTS		
STUDY SELECTION (flow of studies)	16a	- Report, ideally using a flow diagram, the number of: records identified; records excluded before screening; records screened; records excluded after screening titles or titles and abstracts; reports retrieved for detailed evaluation; potentially eligible reports that were not retrievable; retrieved reports that did not meet inclusion criteria and the primary reasons for exclusion; and the number of studies and reports included in the review. If applicable, also report the number of ongoing studies and associated reports identified. - If the review is an update of a previous review, report results of the search and selection process for the current review and specify the number of studies included in the previous review. - If applicable, indicate in the PRISMA flow diagram how many records were excluded by a human and how many by automation tools.
STUDY SELECTION (excluded studies)	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.

Section and Topic	Item #	Elements recommended for reporting
STUDY CHARACTERISTICS	17	- Cite each included study. - Present the key characteristics of each study in a table or figure (considering a format that will facilitate comparison of characteristics across the studies). <i>If the review examines the effects of interventions, consider presenting an additional table that summarises the intervention details for each study.</i>
RISK OF BIAS IN STUDIES	18	- Present tables or figures indicating for each study the risk of bias in each domain/component/item assessed (e.g. blinding of outcome assessors, missing outcome data) and overall study-level risk of bias. - Present justification for each risk of bias judgement, for example in the form of relevant quotations from reports of included studies. <i>If assessments of risk of bias were done for specific outcomes or results in each study, consider displaying risk of bias judgements on a forest plot, next to the study results.</i>
RESULTS OF INDIVIDUAL STUDIES	19	- For all outcomes, irrespective of whether statistical synthesis was undertaken, present for each study summary statistics for each group (where appropriate). For dichotomous outcomes, report the number of participants with and without the events for each group; or the number with the event and the total for each group (e.g. 12/45). For continuous outcomes, report the mean, standard deviation and sample size of each group. - For all outcomes, irrespective of whether statistical synthesis was undertaken, present for each study an effect estimate and its precision (e.g. standard error or 95% confidence/credible interval). For example, for time-to-event outcomes, present a hazard ratio and its confidence interval. - If study-level data is presented visually or reported in the text (or both), also present a tabular display of the results. - If results were obtained from multiple data sources (e.g. journal article, study register entry, clinical study report, correspondence with authors), report the source of the data. - If applicable, indicate which results were not reported directly and had to be computed or estimated from other information.
RESULTS OF SYNTHESSES (characteristics of contributing studies)	20a	- Provide a brief summary of the characteristics and risk of bias among studies contributing to each synthesis (meta-analysis or other). The summary should focus only on study characteristics that help in interpreting the results (especially those that suggest the evidence addresses only a restricted part of the review question, or indirectly addresses the question). - Indicate which studies were included in each synthesis (e.g. by listing each study in a forest plot or table or citing studies in the text).
RESULTS OF SYNTHESSES (results of statistical syntheses)	20b	- Report results of all statistical syntheses described in the protocol and all syntheses conducted that were not pre-specified. - If meta-analysis was conducted, report for each: - the summary estimate and its precision (e.g. standard error or 95% confidence/credible interval) - measures of statistical heterogeneity (e.g. τ^2, I^2, prediction interval) - If other statistical synthesis methods were used (e.g. summarising effect estimates, combining P values), report the synthesized result and a measure of precision (or equivalent information, for example, the number of studies and total sample size). - If the statistical synthesis method does not yield an estimate of effect (e.g. as is the case when P values are combined), report the relevant statistics (e.g. P value from the statistical test), along with an interpretation of the result that is consistent with the question addressed by the synthesis method. - If comparing groups, describe the direction of effect (e.g. fewer events in the intervention group, or higher pain in the comparator group). - If synthesising mean differences, specify for each synthesis, where applicable, the unit of measurement (e.g. kilograms or pounds for weight), the upper and lower limits of the measurement scale (e.g. anchors range from 0 to 10), direction of benefit (e.g. higher scores denote higher severity of pain), and the minimally important difference, if known. If synthesising standardised mean differences, and the effect estimate is being re-expressed to a particular instrument, specify details of the instrument, as per the mean difference.
RESULTS OF SYNTHESSES (results of investigations of heterogeneity)	20c	- If investigations of possible causes of heterogeneity were conducted: - present results regardless of the statistical significance, magnitude, or direction of effect modification. - identify the studies contributing to each subgroup. - report results with due consideration to the observational nature of the analysis and risk of confounding due to other factors. - If subgroup analysis was conducted:

Section and Topic	Item #	Elements recommended for reporting
		- o report for each analysis the exact P value for a test for interaction, as well as, within each subgroup, the summary estimates, their precision (e.g. standard error or 95% confidence/credible interval) and measures of heterogeneity. - o <i>consider presenting the estimate for the difference between subgroups and its precision.</i> • If meta-regression was conducted: - o report for each analysis the exact P value for the regression coefficient and its precision. - o <i>consider presenting a meta-regression scatterplot with the study effect estimates plotted against the potential effect modifier.</i> • If informal methods (i.e. those that do not involve a formal statistical test) were used to investigate heterogeneity, describe the results observed.
RESULTS OF SYNTHESSES (results of sensitivity analyses)	20d	• If any sensitivity analyses were conducted: - o report the results for each sensitivity analysis. - o comment on how robust the main analysis was given the results of all corresponding sensitivity analyses. - o <i>consider presenting results in tables that indicate: (i) the summary effect estimate, a measure of precision (and potentially other relevant statistics, for example, I^2 statistic) and contributing studies for the original meta-analysis; (ii) the same information for the sensitivity analysis; and (iii) details of the original and sensitivity analysis assumptions.</i> - o <i>consider presenting results of sensitivity analyses visually using forest plots.</i>
REPORTING BIASES	21	• Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. • If a tool was used to assess risk of bias due to missing results in a synthesis, present responses to questions in the tool, judgements about risk of bias and any information used to support such judgements. • If a funnel plot was generated to evaluate small-study effects (one cause of which is reporting biases), present the plot and specify the effect estimate and measure of precision used in the plot. If a contour-enhanced funnel plot was generated, specify the 'milestones' of statistical significance that the plotted contour lines represent (P = 0.01, 0.05, 0.1, etc.) • If a test for funnel plot asymmetry was used, report the exact P value observed for the test, and potentially other relevant statistics, for example the standardised normal deviate, from which the P value is derived. • If any sensitivity analyses seeking to explore the potential impact of missing results on the synthesis were conducted, present results of each analysis (see item #20d), compare them with results of the primary analysis, and report results with due consideration of the limitations of the statistical method. • <i>If studies were assessed for selective non-reporting of results by comparing outcomes and analyses pre-specified in study registers, protocols, and statistical analysis plans with results that were available in study reports, consider presenting a matrix (with rows as studies and columns as syntheses) to present the availability of study results.</i> • <i>If an assessment of selective non-reporting of results reveals that some studies are missing from the synthesis, consider displaying the studies with missing results underneath a forest plot or including a table with the available study results.</i>
CERTAINTY OF EVIDENCE	22	• Report the overall level of certainty (or confidence) in the body of evidence for each important outcome. • Provide an explanation of reasons for rating down (or rating up) the certainty of evidence (e.g. in footnotes to an evidence summary table). • Communicate certainty in the evidence wherever results are reported (i.e. abstract, evidence summary tables, results, conclusions), using a format appropriate for the section of the review. • <i>Consider including evidence summary tables, such as GRADE Summary of Findings tables.</i>
DISCUSSION		
DISCUSSION (interpretation)	23a	• Provide a general interpretation of the results in the context of other evidence.
DISCUSSION (limitations of evidence)	23b	• Discuss any limitations of the evidence included in the review.

Section and Topic	Item #	Elements recommended for reporting
DISCUSSION (limitations of review processes)	23c	Discuss any limitations of the review processes used, and comment on the potential impact of each limitation.
DISCUSSION (implications)	23d	- Discuss implications of the results for practice and policy. - Make explicit recommendations for future research.
OTHER INFORMATION		
REGISTRATION AND PROTOCOL (registration)	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.
REGISTRATION AND PROTOCOL (protocol)	24b	Indicate where the review protocol can be accessed (e.g. by providing a citation, DOI or link), or state that a protocol was not prepared.
REGISTRATION AND PROTOCOL (amendments)	24c	Report details of any amendments to information provided at registration or in the protocol, noting: (a) the amendment itself; (b) the reason for the amendment; and (c) the stage of the review process at which the amendment was implemented.
SUPPORT	25	- Describe sources of financial or non-financial support for the review, specifying relevant grant ID numbers for each funder. If no specific financial or non-financial support was received, this should be stated. - Describe the role of the funders or sponsors (or both) in the review. If funders or sponsors had no role in the review, this should be declared.
COMPETING INTERESTS	26	- Disclose any of the authors' relationships or activities that readers could consider pertinent or to have influenced the review. - If any authors had competing interests, report how they were managed for particular review processes.
AVAILABILITY OF DATA, CODE, AND OTHER MATERIALS	27	- Report which of the following are publicly available: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. - If any of the above materials are publicly available, report where they can be found (e.g. provide a link to files deposited in a public repository). - If data, analytic code, or other materials will be made available upon request, provide the contact details of the author responsible for sharing the materials and describe the circumstances under which such materials will be shared.

**Search strategies for
Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic
review with meta-analysis and Trial Sequential Analysis
(S Juul)
Search performed 27 June 2022**

Total number of records identified: 47237 records
Number of duplicates excluded: 15568 records
Number of records in final list: 31669 records

Cochrane Central Register of Controlled Trials (CENTRAL; 2022, Issue 6) in the Cochrane Library (17247 hit)

- #1 MeSH descriptor: [Attention Deficit Disorder with Hyperactivity] explode all trees
- #2 ("attention deficit hyperactivity disorder" or adhd):ti,ab
- #3 #1 or #2
- #4 MeSH descriptor: [Psychotic Disorders] explode all trees
- #5 MeSH descriptor: [Schizophrenia] explode all trees
- #6 (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) NEAR (think* or motor*)) or schizophreni*):ti,ab
- #7 #4 or #5 or #6
- #8 MeSH descriptor: [Bipolar Disorder] explode all trees
- #9 (bipolar or "mood elevation" or mania or hypomania or depress*):ti,ab
- #10 #8 or #9
- #11 MeSH descriptor: [Depressive Disorder] explode all trees
- #12 (depressi* or mood or unipolar):ti,ab
- #13 #11 or #12
- #14 MeSH descriptor: [Anxiety Disorders] explode all trees
- #15 (anxiet* or fear or "avoidance behavior*" or phobia* or panic* or agoraphobia):ti,ab
- #16 #14 or #15
- #17 MeSH descriptor: [Obsessive-Compulsive Disorder] explode all trees
- #18 ("obsessive compulsive disorder" or OCD or urge* or obsessi* or (repetitive NEAR (mental or behavior*))) :ti,ab
- #19 #17 or #18
- #20 MeSH descriptor: [Stress Disorders, Post-Traumatic] explode all trees
- #21 (post-trauma* or trauma*):ti,ab
- #22 #20 or #21
- #23 MeSH descriptor: [Feeding and Eating Disorders] explode all trees
- #24 ("eating behavior" or anorexia* or bulimia* or binge-eating*):ti,ab
- #25 #23 or #24
- #26 MeSH descriptor: [Personality Disorders] explode all trees
- #27 (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*):ti,ab
- #28 #26 or #27
- #29 3 or 7 or 10 or 13 or 16 or 19 or 22 or 25 or 28
- #30 MeSH descriptor: [Psychotherapy] explode all trees
- #31 (((psycho* or cognitive or behavior* or humanistic or systemic) NEAR therap*) or psychotherap* or "self care" or self-care):ti,ab
- #32 #30 or #31
- #33 (brief or extended or standard or intensiv* or ((short* or long*) NEXT term)):ti,ab
- #34 #32 and #33
- #35 #29 and #34

MEDLINE Ovid (1946 to 27 June 2022) (6964 hits)

1. exp Attention Deficit Disorder with Hyperactivity/
2. (attention deficit hyperactivity disorder or adhd).ti,ab.
3. 1 or 2
4. exp Psychotic Disorders/
5. exp Schizophrenia/
6. (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) adj (think* or motor*)) or schizophreni*):ti,ab.
7. 4 or 5 or 6

8. exp Bipolar Disorder/
9. (bipolar or mood elevation or mania or hypomania or depress*).ti,ab.
10. 8 or 9
11. exp Depressive Disorder/
12. (depressi* or mood or unipolar).ti,ab.
13. 11 or 12
14. exp Anxiety Disorders/
15. (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia).ti,ab.
16. 14 or 15
17. exp Obsessive-Compulsive Disorder/
18. (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive adj (mental or behavior*))).ti,ab.
19. 17 or 18
20. exp Stress Disorders, Post-Traumatic/
21. (post-trauma* or trauma*).ti,ab.
22. 20 or 21
23. exp "Feeding and Eating Disorders"/
24. (eating behavior or anorexia* or bulimia* or binge-eating*).ti,ab.
25. 23 or 24
26. exp Personality Disorders/
27. (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*).ti,ab.
28. 26 or 27
29. 3 or 7 or 10 or 13 or 16 or 19 or 22 or 25 or 28
30. exp psychotherapy/
31. (((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care).ti,ab.
32. 30 or 31
33. (brief or extended or standard or intensiv* or ((short* or long*) and term)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
34. 32 and 33
35. 29 and 34
36. limit 35 to ("all adult (19 plus years)" or "adolescent (13 to 18 years)" or "young adult (19 to 24 years)" or "adult (19 to 44 years)" or "young adult and adult (19-24 and 19-44)" or "middle age (45 to 64 years)" or "middle aged (45 plus years)" or "all aged (65 and over)" or "aged (80 and over)")
37. (random* or blind* or placebo* or meta-analys*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
38. 36 and 37

Embase Ovid (1974 to 27 June 2022) (8773 hits)

1. exp attention deficit disorder/
2. (attention deficit hyperactivity disorder or adhd).ti,ab.
3. 1 or 2
4. exp psychosis/
5. exp schizophrenia/
6. (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) adj (think* or motor*)) or schizophreni*).ti,ab.
7. 4 or 5 or 6
8. exp bipolar disorder/
9. (bipolar or mood elevation or mania or hypomania or depress*).ti,ab.
10. 8 or 9
11. exp depression/
12. (depressi* or mood or unipolar).ti,ab.
13. 11 or 12
14. exp anxiety disorder/
15. (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia).ti,ab.
16. 14 or 15
17. exp obsessive compulsive disorder/

18. (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive adj (mental or behavior*))).ti,ab.
19. 17 or 18
20. exp posttraumatic stress disorder/
21. (post-trauma* or trauma*).ti,ab.
22. 20 or 21
23. exp eating disorder/
24. (eating behavior or anorexia* or bulimia* or binge-eating*).ti,ab.
25. 23 or 24
26. exp personality disorder/
27. (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*).ti,ab.
28. 26 or 27
29. 3 or 7 or 10 or 13 or 16 or 19 or 22 or 25 or 28
30. exp psychotherapy/
31. (((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care).ti,ab.
32. 30 or 31
33. (brief or extended or standard or intensiv* or ((short* or long*) and term)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
34. 32 and 33
35. 29 and 34
36. limit 35 to (adult <18 to 64 years> or aged <65+ years>)
37. (random* or blind* or placebo* or meta-analys*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]
38. 36 and 37

LILACS (Bireme; 1982 to 27 June 2022) (1578 hits)

(attention deficit hyperactivity disorder or adhd) or (psychotic or delusion\$ or hallucination\$ or ((disorganised or abnormal) and (think\$ or motor\$)) or schizophreni\$) or (bipolar or mood elevation or mania or hypomania or depress\$) or (depressi\$ or mood or unipolar) or (anxiet\$ or fear or avoidance behavior\$ or phobia\$ or panic\$ or agoraphobia) or (obsessive compulsive disorder or OCD or urge\$ or obsessi\$ or (repetitive and (mental or behavior\$))) or (post-trauma\$ or trauma\$) or (eating behavior or anorexia\$ or bulimia\$ or binge-eating\$) or (schizotypal\$ or paranoid\$ or schizoid\$ or histrionic\$ or narcissistic\$ or antisocial\$ or borderline\$ or avoidant\$ or dependent or obsessive-compulsi\$) [Words] and (((psycho\$ or cognitive or behavior\$ or humanistic or systemic) and therap\$) or psychotherap\$ or self care or self-care) [Words] and (brief or extended or standard or intensiv\$ or ((short\$ or long\$) and term)) [Words]

PsycINFO (EBSCOhost; 1806 to 27 June 2022) (5877 hits)

S37 S35 AND S36
 S36 TX (random* or blind* or placebo* or meta-analys*)
 S35 S29 AND S34
 S34 S32 AND S33
 S33 TI ((brief or extended or standard or intensiv* or ((short* or long*) and term))) OR AB ((brief or extended or standard or intensiv* or ((short* or long*) and term)))
 S32 S30 OR S31
 S31 TI ((((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care)) OR AB ((((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care))
 S30 MA Psychotherapy
 S29 S3 OR S7 OR S10 OR S13 OR S16 OR S19 OR S22 OR S25 OR S28
 S28 S26 OR S27
 S27 TI ((schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*)) OR AB ((schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*))
 S26 MA Personality Disorders
 S25 S23 OR S24
 S24 TI ((eating behavior or anorexia* or bulimia* or binge-eating*)) OR AB ((eating behavior or anorexia* or bulimia* or binge-eating*))

S23 MA Feeding and Eating Disorders
 S22 S20 OR S21
 S21 TI ((post-trauma* or trauma*)) OR AB ((post-trauma* or trauma*))
 S20 MA Stress Disorders, Post-Traumatic
 S19 S17 OR S18
 S18 TI ((obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive near (mental or behavior*)))) OR AB ((obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive near (mental or behavior*))))
 S17 MA Obsessive-Compulsive Disorder
 S16 S14 OR S15
 S15 TI ((anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia)) OR AB ((anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia))
 S14 MA Anxiety Disorders
 S13 S11 OR S12
 S12 TI ((depressi* or mood or unipolar)) OR AB ((depressi* or mood or unipolar))
 S11 MA Depressive Disorder
 S10 S8 OR S9
 S9 TI ((bipolar or mood elevation or mania or hypomania or depress*)) OR AB ((bipolar or mood elevation or mania or hypomania or depress*))
 S8 MA Bipolar Disorder
 S7 S4 OR S5 OR S6
 S6 TI ((psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) near (think* or motor*)) or schizophreni*)) OR AB ((psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) near (think* or motor*)) or schizophreni*))
 S5 MA Schizophrenia
 S4 MA Psychotic Disorders
 S3 S1 OR S2
 S2 TI ((attention deficit hyperactivity disorder or adhd)) OR AB ((attention deficit hyperactivity disorder or adhd))
 S1 MA attention deficit disorder with hyperactivity

Science Citation Index Expanded (SCI-EXPANDED) (1900 to 27 June 2022); Social Sciences Citation Index (SSCI) (1956 to 27 June 2022); Conference Proceedings Citation Index- Science (CPCI-S) (1990 to 27 June 2022); and Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) (1990 to 27 June 2022) (Web of Science) (6798 hits)

#7 #6 AND #5

#6 TS=(random* or blind* or placebo* or meta-analys*)

#5 #4 AND #1

#4 #3 AND #2

#3 TS=(brief or extended or standard or intensiv* or ((short* or long*) and term))

#2 TS=(((psycho* or cognitive or behavior* or humanistic or systemic) and therap*) or psychotherap* or self care or self-care)

#1 TI=((attention deficit hyperactivity disorder or adhd) or (psychotic or delusion* or hallucination* or ((disorgani*ed or abnormal) and (think* or motor*)) or schizophreni*) or (bipolar or mood elevation or mania or hypomania or depress*) or (depressi* or mood or unipolar) or (anxiet* or fear or avoidance behavior* or phobia* or panic* or agoraphobia) or (obsessive compulsive disorder or OCD or urge* or obsessi* or (repetitive near (mental or behavior*))) or (post-trauma* or trauma*) or (eating behavior or anorexia* or bulimia* or binge-eating*) or (schizotypal* or paranoid* or schizoid* or histrionic* or narcissistic* or antisocial* or borderline* or avoidant* or dependent or obsessive-compulsi*))

List of included papers

(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19)

1. Knekt P, Lindfors O, Harkanen T, Valikoski M, Virtala E, Laaksonen MA, et al. Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*. 2008;38(5):689-703.
2. Lorentzen S, Ruud T, Fjeldstad A, Hoglend P. Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial. *Br J Psychiatry*. 2013;203(3):280-7.
3. Herbert JD, Rheingold AA, Gaudiano BA, Myers VH. Standard Versus Extended Cognitive Behavior Therapy for Social Anxiety Disorder: A Randomized-Controlled Trial. *Behavioural and Cognitive Psychotherapy*. 2004;32(2):131-47.
4. Clark DM, Salkovskis PM, Hackmann A, A. W, Ludgate J, Gelder M. Brief Cognitive Therapy for Panic Disorder: A Randomized Controlled Trial. *Journal of Consulting and Clinical Psychology*. 1999;67(4):583-9.
5. Roberge P, Marchand A, Reinharz D, Savard P. Cognitive-Behavioral Treatment for Panic Disorder With Agoraphobia A Randomized, Controlled Trial and Cost-Effectiveness Analysis. *Journal of Behavior Modification*. 2008;32(3):333-51.
6. Bruijnicks SJE, Lemmens L, Hollon SD, Peeters F, Cuijpers P, Arntz A, et al. The effects of once- versus twice-weekly sessions on psychotherapy outcomes in depressed patients. *Br J Psychiatry*. 2020;216(4):222-30.
7. Barkham M, Rees A, Shapiro DA, Stiles WB, Agnew RM, Halstead J, et al. Outcomes of time-limited psychotherapy in applied settings: replicating the Second Sheffield Psychotherapy Project. *J Consult Clin Psychol*. 1996;64(5):1079.
8. Shapiro DA, Barkham M, Rees A, Hardy GE, Reynolds S, Startup M. Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *J Consult Clin Psychol*. 1994;62(3):522.
9. Bohni M, Spindler H, Arendt M, Hougaard E, Rosenberg N. A randomized study of massed three-week cognitive behavioural therapy schedule for panic disorder. *Acta Psychiatr Scand*. 2009;120(3):187-95.
10. Christensen H, Griffiths K, Mackinnon A, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. *Psychol Med*. 2006;36(12):1737.
11. Dekker J, Molenaar PJ, Kool S, Van Aalst G, Peen J, de Jonghe F. Dose-effect relations in time-limited combined psycho-pharmacological treatment for depression. *Psychol Med*. 2005;35(1):47-58.
12. Ehlers A, Hackmann A, Grey N, Wild J, Liness S, Albert I, et al. A randomized controlled trial of 7-day intensive and standard weekly cognitive therapy for PTSD and emotion-focused supportive therapy. *Am J Psychiatry*. 2014;171(3):294-304.
13. Foa EB, McLean CP, Zang Y, Rosenfield D, Yadin E, Yarvis JS, et al. Effect of prolonged exposure therapy delivered over 2 weeks vs 8 weeks vs present-centered therapy on PTSD symptom severity in military personnel: A randomized clinical trial. *JAMA*. 2018;319(4):354-64.
14. Nacasch N, Huppert JD, Su Y-J, Kivity Y, Dinshtein Y, Yeh R, et al. Are 60-Minute Prolonged Exposure Sessions With 20-Minute Imaginal Exposure to Traumatic Memories Sufficient to Successfully Treat PTSD? A Randomized Noninferiority Clinical Trial. *Behav Ther*. 2015;46(3):328-41.

15. Kenardy JA, Dow MG, Johnston DW, Newman MG, Thomson A, Taylor CB. A comparison of delivery methods of cognitive-behavioral therapy for panic disorder: an international multicenter trial. *J Consult Clin Psychol*. 2003;71(6):1068.
16. Böttche M, Wagner B, Vöhringer M, Heinrich M, Stein J, Selmo P, et al. Is only one cognitive technique also effective? Results from a randomized controlled trial of two different versions of an internet-based cognitive behavioural intervention for post-traumatic stress disorder in Arabic-speaking countries. *Eur J Psychotraumatol*. 2021;12(1):1943870.
17. Hadjistavropoulos H, Peynenburg V, Thiessen D, Nugent M, Karin E, Dear B, et al. A randomized factorial trial of internet-delivered cognitive behavioural therapy: An 8-week program with or without extended support and booster lesson. *Internet Interventions*. 2022;27:100499.
18. Dell L, Sbisà AM, Forbes A, O'Donnell M, Bryant R, Hodson S, et al. Effect of massed v. standard prolonged exposure therapy on PTSD in military personnel and veterans: a non-inferiority randomised controlled trial. *Psychol Med*. 2022:1-8.
19. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022:1-16.

Excluded studies

1. Laaksonen, M. A., Knekt, P., & Lindfors, O. (2013). Psychological predictors of the recovery from mood or anxiety disorder in short-term and long-term psychotherapy during a 3-year follow-up. *Psychiatry Research*, 208(2), 162-173. **(Double or serial publication)**
2. Knekt, P., Laaksonen, M. A., Raitasalo, R., Haaramo, P., & Lindfors, O. (2010). Changes in lifestyle for psychiatric patients three years after the start of short-and long-term psychodynamic psychotherapy and solution-focused therapy. *European psychiatry*, 25(1), 1-7. **(Double or serial publication)**
3. Knekt, P., Heinonen, E., Härkäpää, K., Järvikoski, A., Virtala, E., Rissanen, J., ... & Helsinki Psychotherapy Study Group. (2015). Randomized trial on the effectiveness of long-and short-term psychotherapy on psychosocial functioning and quality of life during a 5-year follow-up. *Psychiatry Research*, 229(1-2), 381-388. **(Double or serial publication)**
4. Baker, A., Turner, A., Kay-Lambkin, F. J., & Lewin, T. J. (2009). The long and the short of treatments for alcohol or cannabis misuse among people with severe mental disorders. *Addictive behaviors*, 34(10), 852-858. **(Different psychotherapy types)**
5. Albert, N., Melau, M., Jensen, H., Emborg, C., Jepsen, J. R. M., Fagerlund, B., ... & Nordentoft, M. (2017). Five years of specialised early intervention versus two years of specialised early intervention followed by three years of standard treatment for patients with a first episode psychosis: randomised, superiority, parallel group trial in Denmark (OPUS II). *Bmj*, 356. **(Not psychotherapy)**
6. Tecic, T., Schneider, A., Althaus, A., Schmidt, Y., Bierbaum, C., Lefering, R., ... & Neugebauer, E. A. (2011). Early short-term inpatient psychotherapeutic treatment versus continued outpatient psychotherapy on psychosocial outcome: a randomized controlled trial in trauma patients. *Journal of Trauma and Acute Care Surgery*, 70(2), 433-441. **(Wrong type of participants)**
7. Lorentzen, S., Fjeldstad, A., Ruud, T., & Høglend, P. A. (2015). Comparing short-and long-term group therapy: Seven-year follow-up of a randomized clinical trial. *Psychotherapy and psychosomatics*, 84(5), 320-322. **(Double or serial publication)**
8. Heinonen, E., Knekt, P., Härkänen, T., Virtala, E., & Lindfors, O. (2018). Childhood adversities as predictors of improvement in psychiatric symptoms and global functioning in solution-focused and short-and long-term psychodynamic psychotherapy during a 5-year follow-up. *Journal of affective disorders*, 235, 525-534. **(Double or serial publication)**

9. Knekt, P., Virtala, E., Härkänen, T., Vaarama, M., Lehtonen, J., & Lindfors, O. (2016). The outcome of short-and long-term psychotherapy 10 years after start of treatment. *Psychological medicine*, 46(6), 1175-1188. **(Double or serial publication)**
10. Dow, M., Kenardy, J., Johnston, D., Newman, M., Taylor, C., & Thomson, A. (2007). Prognostic indices with brief and standard CBT for panic disorder: II. Moderators of outcome. *Psychological Medicine*, 37(10), 1503-1509. **(Double or serial publication)**
11. Knekt, P., Lindfors, O., Sares-Jäske, L., Virtala, E., & Härkänen, T. (2013). Randomized trial on the effectiveness of long-and short-term psychotherapy on psychiatric symptoms and working ability during a 5-year follow-up. *Nordic journal of psychiatry*, 67(1), 59-68. **(Double or serial publication)**
12. Laurensen, E. M., Westra, D., Kikkert, M. J., Noom, M. J., Eeren, H. V., van Broekhuizen, A. J., ... & Dekker, J. J. (2014). Day Hospital Mentalization-Based Treatment (MBT-DH) versus treatment as usual in the treatment of severe borderline personality disorder: protocol of a randomized controlled trial. *BMC psychiatry*, 14(1), 1-13. **(Protocol)**
13. Ordman, A. M., & Kirschenbaum, D. S. (1985). Cognitive-behavioral therapy for bulimia: An initial outcome study. *Journal of Consulting and Clinical Psychology*, 53(3), 305–313. **(Different psychotherapy types)**
14. Kool, M., Van, H. L., Bartak, A., de Maat, S. C., Arntz, A., van den Eshof, J. W., ... & Dekker, J. J. (2018). Optimizing psychotherapy dosage for comorbid depression and personality disorders (PsyDos): a pragmatic randomized factorial trial using schema therapy and short-term psychodynamic psychotherapy. *BMC psychiatry*, 18, 1-15. **(Protocol)**
15. Ben-Itzhak, S., Bluvstein, I., Schreiber, S., Aharonov-Zaig, I., Maor, M., Lipnik, R., & Bloch, M. (2012). The effectiveness of brief versus intermediate duration psychodynamic psychotherapy in the treatment of adjustment disorder. *Journal of Contemporary Psychotherapy*, 42, 249-256. **(Wrong type of participants)**
16. Botella, C., & García-Palacios, A. (1999). The possibility of reducing therapist contact and total length of therapy in the treatment of panic disorder. *Behavioural and Cognitive Psychotherapy*, 27(3), 231-247. **(Different psychotherapy types)**
17. Juul, S., Lunn, S., Poulsen, S., Sørensen, P., Salimi, M., Jakobsen, J. C., ... & Simonsen, S. (2019). Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*, 20, 1-10. **(Protocol)**

18. De Jong, K., Timman, R., Hakkaart-Van Roijen, L., Vermeulen, P., Kooiman, K., Passchier, J., & Busschbach, J. V. (2014). The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: A randomized controlled trial. *Psychotherapy Research*, 24(6), 629-639. **(Wrong comparison)**
19. Murray, L. K., Haroz, E. E., Doty, B., Singh, N. S., Bogdanov, S., Bass, J., ... & Bolton, P. (2018). Testing the effectiveness and implementation of a brief version of the Common Elements Treatment Approach (CETA) in Ukraine: a study protocol for a randomized controlled trial. *Trials*, 19(1), 1-16. **(Protocol)**
20. Gottheil, E., Weinstein, S. P., Sterling, R. C., Lundy, A., & Serota, R. D. (1998). A randomized controlled study of the effectiveness of intensive outpatient treatment for cocaine dependence. *Psychiatric Services*, 49(6), 782-787. **(Different psychotherapy types)**
21. Fjeldstad, A., Høglend, P., & Lorentzen, S. (2017). Patterns of change in interpersonal problems during and after short-term and long-term psychodynamic group therapy: A randomized clinical trial. *Psychotherapy Research*, 27(3), 350-361. **(Double or serial publication)**
22. Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose–effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, 64(5), 927–935. **(Wrong study design)**
23. Glick, I. D., Hargreaves, W. A., Drues, J., & Showstack, J. A. (1976). Short versus long hospitalization: A prospective controlled study: IV. One-year follow-up results for schizophrenic patients. *The American Journal of Psychiatry*. **(Not psychotherapy)**
24. Reneses, B., Figuera, D., Salcedo, G., Trujillo, M., López-Ibor, J. J., Galián, M., ... & Serrano, R. (2011). A controlled randomized study on the efficacy of short-term dynamic psychotherapy in borderline personality disorders (BPD). Preliminary results. *European Psychiatry*, 26(S2), 1040-1040. **(Different psychotherapy types)**
25. Covi, L., Hess, J. M., Schroeder, J. R., & Preston, K. L. (2002). A dose response study of cognitive behavioral therapy in cocaine abusers. *Journal of substance abuse treatment*, 23(3), 191-197. **(Wrong type of participants)**
26. Naeem, F., Farooq, S., & Kingdon, D. (2014). Cognitive behavioral therapy (brief vs standard duration) for schizophrenia. *Schizophrenia bulletin*, 40(5), 958-959. **(Systematic review)**

27. Buonocore, M., Bosia, M., Bechi, M., Spangaro, M., Cavedoni, S., Cocchi, F., ... & Cavallaro, R. (2017). Is longer treatment better? A comparison study of 3 versus 6 months cognitive remediation in schizophrenia. *Neuropsychology*, 31(4), 467. **(Not psychotherapy)**
28. Marchand, A., Roberge, P., Primiano, S., & Germain, V. (2009). A randomized, controlled clinical trial of standard, group and brief cognitive-behavioral therapy for panic disorder with agoraphobia: A two-year follow-up. *Journal of Anxiety Disorders*, 23(8), 1139-1147. **(Double or serial publication)**
29. McMain, S. F., Chapman, A. L., Kuo, J. R., Guimond, T., Streiner, D. L., Dixon-Gordon, K. L., ... & Hoch, J. S. (2018). The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial protocol. *BMC psychiatry*, 18(1), 1-16. **(Protocol)**
30. Kashner, T. M., Henley, S. S., Golden, R. M., Rush, A. J., & Jarrett, R. B. (2007). Assessing the preventive effects of cognitive therapy following relief of depression: a methodological innovation. *Journal of affective disorders*, 104(1-3), 251-261. **(Different psychotherapy types)**
31. Lorentzen, S., Ruud, T., Fjeldstad, A., & Høglend, P. A. (2015). Personality disorder moderates outcome in short-and long-term group analytic psychotherapy: A randomized clinical trial. *British Journal of Clinical Psychology*, 54(2), 129-146. **(Double or serial publication)**
32. Knekt, P., Lindfors, O., Laaksonen, M. A., Raitasalo, R., Haaramo, P., Järvikoski, A., & Helsinki Psychotherapy Study Group. (2008). Effectiveness of short-term and long-term psychotherapy on work ability and functional capacity—a randomized clinical trial on depressive and anxiety disorders. *Journal of affective disorders*, 107(1-3), 95-106. **(Double or serial publication)**
33. Knekt, P., Lindfors, O., Laaksonen, M. A., Renlund, C., Haaramo, P., Härkänen, T., ... & Helsinki Psychotherapy Study Group. (2011). Quasi-experimental study on the effectiveness of psychoanalysis, long-term and short-term psychotherapy on psychiatric symptoms, work ability and functional capacity during a 5-year follow-up. *Journal of Affective Disorders*, 132(1-2), 37-47. **(Double or serial publication)**
34. Barrowclough, C., Marshall, M., Gregg, L., Fitzsimmons, M., Tomenson, B., Warburton, J., & Lobban, F. (2014). A phase-specific psychological therapy for people with problematic

- cannabis use following a first episode of psychosis: a randomized controlled trial. *Psychological Medicine*, 44(13), 2749-2761. **(Wrong type of participants)**
35. Hargreaves, W. A., Glick, I. D., Druess, J., Showstack, J. A., & Feigenbaum, E. (1977). Short vs long hospitalization: A prospective controlled study: VI. Two-year follow-up results for schizophrenics. *Archives of General Psychiatry*, 34(3), 305-311. **(Not psychotherapy)**
 36. Kroska, E. B., Roche, A. I., & O'Hara, M. W. (2020). How much is enough in brief Acceptance and Commitment Therapy? A randomized trial. *Journal of Contextual Behavioral Science*, 15, 235-244. **(Wrong type of participants)**
 37. Johnson, D. R., & Lubin, H. (2002). Effect of brief versus long-term inpatient treatment on homecoming stress in combat-related posttraumatic stress disorder: three-year follow-up. *The Journal of nervous and mental disease*, 190(1), 47-51. **(Non-randomised)**
 38. Joutsenniemi, K., Laaksonen, M. A., Knekt, P., Haaramo, P., & Lindfors, O. (2012). Prediction of the outcome of short-and long-term psychotherapy based on socio-demographic factors. *Journal of affective disorders*, 141(2-3), 331-342. **(Double or serial publication)**
 39. Chang, W. C., Chan, G. H. K., Jim, O. T. T., Lau, E. S. K., Hui, C. L. M., Chan, S. K. W., ... & Chen, E. Y. H. (2015). Optimal duration of an early intervention programme for first-episode psychosis: randomised controlled trial. *The British Journal of Psychiatry*, 206(6), 492-500. **(Not psychotherapy)**
 40. Dow, M. G., Kenardy, J. A., Johnston, D. W., Newman, M. G., Taylor, C. B., & Thomson, A. (2007). Prognostic indices with brief and standard CBT for panic disorder: I. Predictors of outcome. *Psychological Medicine*, 37(10), 1493-1502. **(Double or serial publication)**

Supplementary material 5. Risk of bias table

	Risk of bias							Overall
	D1	D2	D3	D4	D5	D6	D7	
Barkham et al. 1996	-	-	-	-	-	-	+	X
Bohni et al. 2009	-	-	-	X	X	-	+	X
Böttche et al. 2021	+	-	-	-	X	+	+	X
Bruijnicks et al. 2020	+	+	-	-	X	X	+	X
Christensen et al. 2006	-	-	-	-	-	-	+	X
Clark et al. 1999	-	-	-	-	-	-	X	X
Dekker et al. 2004	-	-	X	+	X	-	+	X
Dell et al. 2022	+	+	X	+	+	+	+	X
Ehlers et al. 2014	-	-	X	+	+	-	+	X
Foa et al. 2018	-	-	-	+	X	X	+	X
Hadjistavropoulos et al. 2022	+	-	X	X	X	+	+	X
Herbert et al. 2004	-	-	X	X	X	-	+	X
Kenardy et al. 2003	-	-	-	-	-	X	X	X
Knekt et al. 2008	+	+	-	X	X	-	+	X
Lorentzen et al. 2013	X	-	-	-	X	X	+	X
McMain et al. 2022	+	+	X	+	X	+	+	X
Nacasch et al. 2015	+	+	-	+	X	-	+	X
Roberge et al. 2008	-	-	-	-	-	-	+	X
Shapiro et al. 1990	-	-	-	-	X	-	+	X

Study

D1: Random sequence generation
D2: Allocation concealment
D3: Blinding of participants and personnel
D4: Blinding of outcome assessment
D5: Incomplete outcome data
D6: Selective reporting
D7: Other sources of bias

Judgement
X High
- Unclear
+ Low

Supplementary material: Required information sizes for the predefined review outcomes

Trial	Number randomized	Predefined review outcomes								
		Quality of life		Serious adverse events	Symptom severity		Suicide/ suicide-attempts	Self-harm	Level of functioning	
Barkham et al. 1996	54	-		-	-		-	-	-	
Böttche et al. 2021	224	RIS:	176		RIS	204				
Bohni et al. 2009*	48	-		-	-		-	-	-	
Bruijniks et al. 2020	CBT: 100	CBT:		-	CBT:		-	-	-	
		RIS:	188		RIS:	190				
	IPT: 100	IPT			IPT:					
		RIS:	186		RIS:	178				
Christensen et al. 2006	931	-		-	-		-	-	-	
Clark et al. 1999	29	-		-	RIS:	148	-	-	-	
Dekker et al. 2004	103	RIS:	174	-	RIS:	154	-	-	-	
Dell et al. 2021	135				RIS:	170				
Ehlers et al. 2014	61	RIS:	174	-	RIS:	188	-	-	RIS:	170
Foa et al. 2018	219	-		-	N/A		-	-	-	
Hadjistavropoulos et al. 2022	ICBT no booster versus booster: 236	RIS:	202	-	RIS PTSD:	178	-		RIS:	210
					RIS GAD:	170				
	ICBT- extention no booster versus booster: 233	RIS:	182		RIS PTSD:	174			RIS:	174
					RIS GAD:	172				
Herbert et al. 2004	34	-		-	RIS:	178	-	-	-	
Kenardy et al. 2003	81	-		-	RIS:	188	-	-	-	
Knekt et al. 2008	229	-		-	RIS:	178	-	-	RIS:	218
Lorentzen et al. 2013	167	-		-	-		N/A	-	RIS:	188

Supplementary material: Required information sizes for the predefined review outcomes

McMain et al. (unpublished data)	240	RIS:	170	RIS:	49,084	RIS:	172	RIS:	49,084	RIS:	1,596	RIS:	172
Nacasch et al. 2015	40	-		-		RIS:	162	-		-		-	
Roberge et al. 2008	65	-		-		RIS:	232	-		-		-	
Shapiro et al. 1990	150	-		-		-		-		-		-	

CBT; Cognitive Behavioural Therapy: IPT; Interpersonal Therapy: RIS; Required Information Size

Summary of Findings table

Short-term versus long-term dialectical behavioral therapy for borderline personality disorder at end of treatment						
Patients or population: Borderline personality disorder Setting: Any setting Intervention: Short-term dialectical behavioral therapy Comparison: Long-term dialectical behavioral therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term DBT	Risk with Short-term DBT				
Quality of life Follow-up: 12 months	-	-	MD: -0.71 (-7.25 to 5.83)	181 (1 RCT)	⊕○○○ VERY LOW a,b	
Serious adverse events Follow-up: 12 months	22 per 1,000	22 per 1,000	-	183 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Symptom severity Follow-up: 12 months	-	-	MD: -0.7 (-7.24 to 5.84)	181 (1 RCT)	⊕○○○ VERY LOW a,b	
Suicide or suicide attempts Follow-up: 12 months	22 per 1,000	22 per 1,000	-	183 (1 RCT)	⊕○○○ VERY LOW a,b, c	
Self-harm Follow-up: 12 months	397 per 1,000	311 per 1,000	-	183 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Level of functioning Follow-up: 12 months	-	-	MD: 0.73 (-0.2 to 0.14)	181 (1 RCT)	⊕○○○ VERY LOW a,b	

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence; **MD:** Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

c. Downgraded 2 for imprecision due to low number of participants

Summary of Findings table

Short-term versus long-term psychodynamic psychotherapy for mood- and anxiety disorders at end of treatment						
Patients or population: Mood- and anxiety disorders Setting: Any setting Intervention: Short-term psychodynamic psychotherapy Comparison: Long-term psychodynamic psychotherapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term psychodynamic	Risk with Short-term psychodynamic				
Quality of life	-	-	-	-	-	Outcome not yet measured or reported
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 36 months	-	-	MD: 1.8 (0.112 to 3.48)	190 (1 RCT)	⊕○○○ VERY LOW a,b	
Suicide or suicide attempts Follow-up: 36 months	0 per 1,000	0 per 1,000	-	167 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning Follow-up: mean 36 months	-	-	SMD -0.13; (-0.42 to 0.15)	340 (2 RCTs)	⊕○○○ VERY LOW a,c	

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence; **SMD:** Standardized Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for indirectness as the results are from a single trial from a single country, therefore results may not be generalizable to other settings.

c. Downgraded 1 for imprecision due to low number of participants

Summary of Findings table

Short-term versus long-term internet-based cognitive behavioral therapy at end of treatment						
Patients or population: Depression and anxiety						
Setting: Any setting						
Intervention: Short-term (8 weeks) internet-based cognitive behavioral therapy						
Comparison: Long-term (8 weeks + 3 booster sessions) internet-based cognitive behavioral therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term interpersonal therapy	Risk with Short-term interpersonal therapy				
Quality of life Follow-up: 10 weeks	-	-	MD: 0.56 (-2.71 to 3.83)	166 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 10 weeks	-	-	Depression MD: 2.09 (1.31 to 2.86) Anxiety MD: 2 (1.29 to 2.7)	166 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning Follow-up: 10 weeks	-	-	MD: 3.2 (1.86 to 4.53)	166 (1 RCT)	⊕○○○ VERY LOW a,b,c	

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence, **MD:** Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Short-term versus long-term internet-based cognitive behavioral therapy at end of treatment						
Patients or population: Depression and anxiety						
Setting: Any setting						
Intervention: Short-term (12 weeks) internet-based cognitive behavioral therapy						
Comparison: Long-term (12 weeks + 3 booster sessions) internet-based cognitive behavioral therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term interpersonal therapy	Risk with Short-term interpersonal therapy				
Quality of life Follow-up: 15 weeks	-	-	MD: 4.93 (1.68 to 8.17)	178 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 15 weeks	-	-	Depression MD: 2.09 (1.31 to 2.86) Anxiety MD: 2 (1.29 to 2.7)	178 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning Follow-up: 15 weeks	-	-	MD: 3.2 (1.86 to 4.53)	178 (1 RCT)	⊕○○○ VERY LOW a,b,c	

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence, **MD:** Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Short-term versus long-term cognitive behavioral therapy for major depressive disorder at end of treatment						
Patients or population: Major depressive disorder Setting: Any setting Intervention: Short-term cognitive behavioral therapy Comparison: Long-term cognitive behavioral therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term CBT	Risk with Short-term CBT				
Quality of life Follow-up: 6 months	-	-	MD -1.4 (-10.33 to 7.53)	98 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 6 months	-	-	MD: 2.91 (-2.71 to 8.5)	98 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning	-	-	-	-	-	Outcome not yet measured or reported

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence; **MD:** Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Short-term versus long-term interpersonal therapy for major depressive disorder at end of treatment						
Patients or population: Major depressive disorder Setting: Any setting Intervention: Short-term interpersonal therapy Comparison: Long-term interpersonal therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term interpersonal therapy	Risk with Short-term interpersonal therapy				
Quality of life Follow-up: 6 months	-	-	MD: -6.66 (-14.76 to 1.44)	102 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 6 months	-	-	MD: 2.89 (-3.16 to 8.94)	102 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning	-	-	-	-	-	Outcome not yet measured or reported

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence, **MD:** Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

- a. Downgraded 2 for risk of bias
- b. Downgraded 1 for imprecision due to low number of participants
- c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Short-term versus long-term psychodynamic psychotherapy for major depressive disorder at end of treatment						
Patients or population: Major depressive disorder Setting: Any setting Intervention: Short-term psychodynamic psychotherapy Comparison: Long-term psychodynamic psychotherapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term psychodynamic psychotherapy	Risk with Short-term psychodynamic psychotherapy				
Quality of life Follow-up: 24 weeks	-	-	MD: -0.2 (-3.74 to 3.34)	90 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 24 weeks	-	-	MD: -1 (-4.02 to 2.02)	90 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning	-	-	-	-	-	Outcome not yet measured or reported
CI: Confidence interval; GRADE: GRADE Working Group grades of evidence, MD: Mean Difference						
GRADE Working Group grades of evidence High certainty: We are very confident that the true effect lies close to that of the estimate of the effect Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect						

Explanations

- a. Downgraded 2 for risk of bias
- b. Downgraded 1 for imprecision due to low number of participants
- c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Short-term versus long-term cognitive behavioral therapy for post-traumatic stress disorder at end of treatment						
Patients or population: Post-traumatic stress disorder Setting: Any setting Intervention: Short-term cognitive behavioral therapy Comparison: Long-term cognitive behavioral therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term CBT	Risk with Short-term CBT				
Quality of life Follow-up: 14 weeks	-	-	MD: -10 (-21.01 to 0.49)	61 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: 14 weeks	-	-	MD: 5.25 (-9.07 to 19.57)	61 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning Follow-up: 14 weeks	-	-	MD: -0.72 (-5.34 to 3.90)	61 (1 RCT)	⊕○○○ VERY LOW a,b,c	

CBT: cognitive behavioral therapy; CI: Confidence interval; GRADE: GRADE Working Group grades of evidence, MD: Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

60 versus 90 minute sessions prolonged exposure therapy for post-traumatic stress disorder at end of treatment

Patients or population: Post-traumatic stress disorder

Setting: Any setting

Intervention: Short-term prolonged exposure therapy

Comparison: Long-term prolonged exposure therapy

Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term CBT	Risk with Short-term CBT				
Quality of life	-	-	-	-	-	Outcome not yet measured or reported
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: posttreatment (unclear)	-	-	MD: 1.06 (-4.87 to 6.99)	37 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning	-	-	-	-	-	Outcome not yet measured or reported

CI: Confidence interval; GRADE: GRADE Working Group grades of evidence; MD: Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 2 for risk of bias

b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Massed versus standard prolonged exposure therapy for post-traumatic stress disorder at end of treatment						
Patients or population: Post-traumatic stress disorder Setting: Any setting Intervention: Short-term prolonged exposure therapy Comparison: Long-term prolonged exposure therapy						
Outcomes	Anticipated absolute effects		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term CBT	Risk with Short-term CBT				
Quality of life	-	-	-	-	-	Outcome not yet measured or reported
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: posttreatment (10 weeks)	-	-	MD: 2.01 (-.82 to 4.84)	137 (1 RCT)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning	-	-	-	-	-	Outcome not yet measured or reported

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence; **MD:** Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. Downgraded 1 for risk of bias

b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 1 for indirectness because the results are from a single trial from a single country, therefore results may not be generalizable to other settings

Summary of Findings table

Short-term versus long-term cognitive behavioral therapy for anxiety disorders at end of treatment

Patients or population: Anxiety disorders

Setting: Any setting

Intervention: Short-term cognitive behavioral therapy

Comparison: Long-term cognitive behavioral therapy

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Long-term CBT	Risk with Short-term CBT				
Quality of life	-	-	-	-	-	Outcome not yet measured or reported
Serious adverse events	-	-	-	-	-	Outcome not yet measured or reported
Symptom severity Follow-up: mean 28.5 weeks	-	-	SMD: 0.08 (-0.47 to 0.63)	209 (4 RCTs)	⊕○○○ VERY LOW a,b,c	
Suicide or suicide attempts	-	-	-	-	-	Outcome not yet measured or reported
Self-harm	-	-	-	-	-	Outcome not yet measured or reported
Level of functioning	-	-	-	-	-	Outcome not yet measured or reported

CI: Confidence interval; **GRADE:** GRADE Working Group grades of evidence; **SMD:** Standardized Mean Difference

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

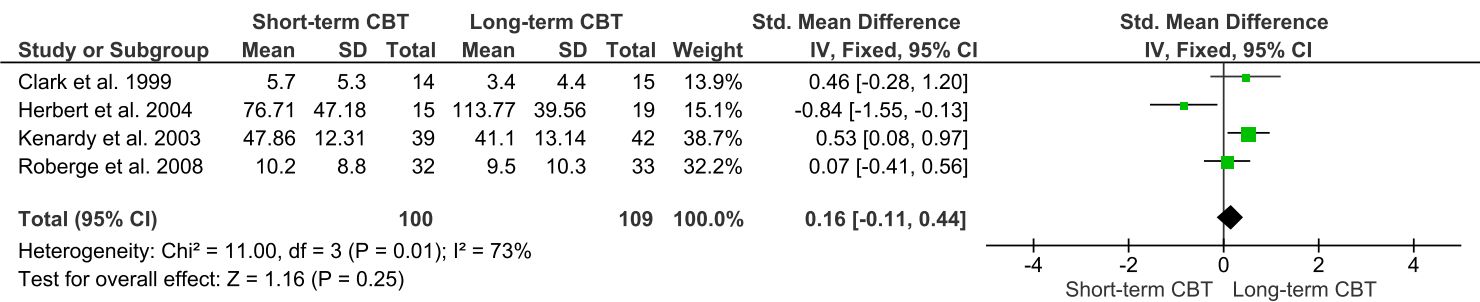
Explanations

a. Downgraded 2 for risk of bias

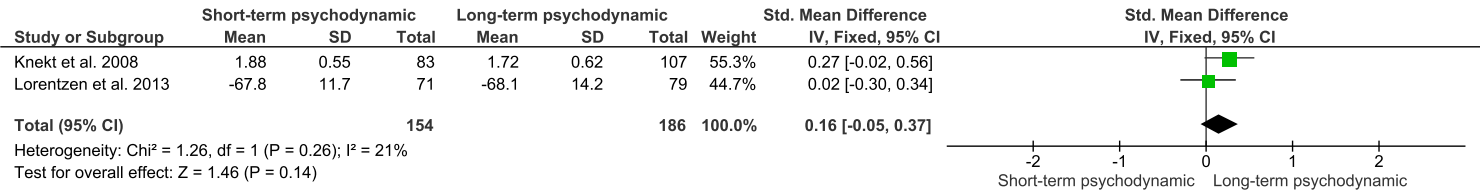
b. Downgraded 1 for imprecision due to low number of participants

c. Downgraded 2 for heterogeneity ($I^2 = 73\%$)

Supplementary material 18. Fixed-effect meta-analysis of short-term versus long-term cognitive-behavioral therapy for anxiety disorders on anxiety symptoms (sensitivity analysis)



Supplementary material 19. Fixed-effect meta-analysis of short-term versus long-term psychodynamic psychotherapy for mood- and anxiety disorders on level of functioning (sensitivity analysis)



Paper 3

STUDY PROTOCOL

Open Access



Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial

Sophie Juul^{1,2*} , Susanne Lunn², Stig Poulsen², Per Sørensen¹, Mehrak Salimi¹, Janus Christian Jakobsen³, Anthony Bateman⁴ and Sebastian Simonsen¹

Abstract

Background: Psychotherapy for borderline personality disorder is often lengthy and resource-intensive. However, the current length of outpatient treatments is arbitrary and based on trials that never tested if the treatment intensity could be reduced. As a result, there is insufficient evidence to inform the decision between short-term and long-term psychotherapy for borderline personality disorder. Mentalization-based therapy is one treatment option for borderline personality disorder and consists traditionally of an 18-month treatment program.

Methods/design: This trial is an investigator-initiated single-center randomized clinical superiority trial of short-term (20 weeks) compared to long-term (14 months) mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder. Participants will be recruited from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark. Participants will be included if they meet a minimum of four DSM-V criteria for borderline personality disorder. Participants will be assessed before randomization, and at 8, 16, and 24 months after randomization. The primary outcome is severity of borderline symptomatology assessed with the Zanarini Rating Scale for borderline personality disorder. Secondary outcomes include self-harm incidents, functional impairment (Work and Social Adjustment Scale, Global Assessment of Functioning) and quality of life (Short-Form Health Survey 36). Severity of psychiatric symptoms (Symptom Checklist 90-R) will be included as an exploratory outcome. Measures of personality functioning, attachment, borderline symptoms, group alliance, and mentalization skills will be included to explore potential predictors and mechanisms of change.

Discussion: This trial will provide evidence of the beneficial and harmful effects of short-term compared to long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder.

Trial registration: ClinicalTrials.gov, [NCT03677037](https://clinicaltrials.gov/ct2/show/study/NCT03677037). Registered on September 19, 2018.

Keywords: Mentalization-based therapy, Borderline personality disorder, Randomized clinical trial, Treatment intensity

* Correspondence: sophie.juul@regionh.dk

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Gentofte, Capital Region of Denmark, Denmark

²Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

Borderline personality disorder is a psychiatric condition characterized by a pervasive pattern of symptoms such as interpersonal conflicts, identity diffusion, impulsivity, and emotional dysregulation [1]. According to epidemiological studies, 1.6% of the general population suffer from borderline personality disorder [2]. In clinical populations, it is the most common personality disorder [2], with a prevalence of between 9% and 22% of all psychiatric outpatients [3–5]. Borderline personality disorder is associated with high levels of psychiatric comorbidity, particularly depression, anxiety disorders, eating disorders, substance abuse [6–8], and other personality disorders [9]. Together, these findings emphasize the need for the development of efficacious and cost-effective treatments for this severe and highly prevalent disorder.

While pharmacological treatment may reduce some borderline-related symptoms, there is still no convincing evidence that it is suitable for treating all diagnostic criteria [10]. Although further evidence is still warranted, psychotherapy continues to be the primary treatment of choice for borderline personality disorder [11]. During the last 10–15 years, studies have established the efficacy of different forms of intensive, specialized long-term psychotherapy modalities. These have recently been evaluated in a systematic review and meta-analysis exploring the efficacy of psychotherapies for borderline personality disorder, in which it was concluded that dialectical behavior therapy and psychodynamic therapies (transference focused therapy and mentalization-based therapy) significantly improved borderline-relevant outcomes [12]. However, no single treatment modality has been established as the primary treatment of choice.

Mentalization-based therapy (MBT) is a psychodynamic therapy rooted in attachment and cognitive theory [13], which was developed specifically for treating borderline personality disorder [14]. Mentalization refers to the capacity to understand one's own and others' mental states. The theoretical assumption is that patients with borderline personality disorder are more vulnerable to lose this capacity when experiencing emotional distress. The MBT manual offers therapeutic techniques to identify these shifts and to bring the patient back into a mentalizing mode [14, 15]. The therapy program consists of four basic components: (1) psycho-education, (2) case formulation, (3) group therapy, and (4) individual therapy. All of these aim to enhance the patient's capacity to mentalize. Increasing mentalization skills is assumed to minimize borderline-related symptoms such as emotional dysregulation, impulsivity, and suicidal ideation. However, information about the mechanisms that produce a change in MBT, or in psychotherapy in general, is still limited [16, 17].

MBT for adult borderline personality disorder has been tested in cohort studies [18, 19] and one randomized but uncontrolled trial [20]. Two forms of MBT have been tested in randomized controlled trials: day hospital MBT [21, 22] and intensive outpatient MBT [13], each lasting a maximum of 18 months. For a systematic review of the current evidence base of MBT for borderline personality disorder, see Vogt and Norman [16].

Bateman and Fonagy [13] assessed the effects of intensive outpatient MBT in a randomized clinical trial, in which 134 participants with a confirmed borderline personality disorder diagnosis were randomized either to 18 months of outpatient MBT, combining weekly group and individual sessions with different therapists, or to structured clinical management. In this trial, MBT was superior to structured clinical management in terms of its effects on suicide attempts, severe incidents of self-harm, and on self-reported measures. Treatment effects were sustained at the 5-year follow-up [23]. Nevertheless, only 134 participants were randomized, which questions whether the trial was powered to assess the chosen outcomes, and only 41 were assessed after 5 years. Further, the trial investigators were also the developers of MBT. Thus, the small sample size and the substantial problems with incomplete outcome data, especially at the long-term follow-up, are threats to the validity of the study.

However, while intensive outpatient MBT currently has empirical support as an 18-month program for borderline personality disorder, evidence that this is the optimal length of the intervention is not available. Consequently, MBT is now offered for different lengths of time (both shorter and longer) in outpatient settings around the world [14]. Various other short-term psychotherapies for borderline personality disorder have already been developed and tested in randomized clinical trials, e.g., emotion regulation group therapy [24], systems training for emotional predictability and problem-solving [25, 26], and brief dialectical behavior therapy skills training [27]. However, all the trials have either compared a short-term experimental group to a short-term control group or tested the short-term treatment as an adjunctive to treatment as usual. Thus, these trials do not provide guidance on evidence-based decisions regarding the optimal length of treatment for borderline patients. In addition, no empirical evidence is available to identify which subtypes of patients would benefit from short-term treatment and which would require more intensive treatment [28].

We performed a preliminary literature search (PubMed and Cochrane Library) for trials comparing different lengths of psychotherapy for borderline personality disorder. No such trials were found. When we expanded our search terms to all types of psychiatric disorders, only few trials were identified [29, 30]. We are

currently working on a protocol for a more comprehensive systematic review, including a full assessment of risk of bias and a trial sequential analysis of short-term compared with long-term psychotherapy for all psychiatric disorders. The systematic review will be submitted for publication before data collection is completed in this trial.

Methods/design

Objective

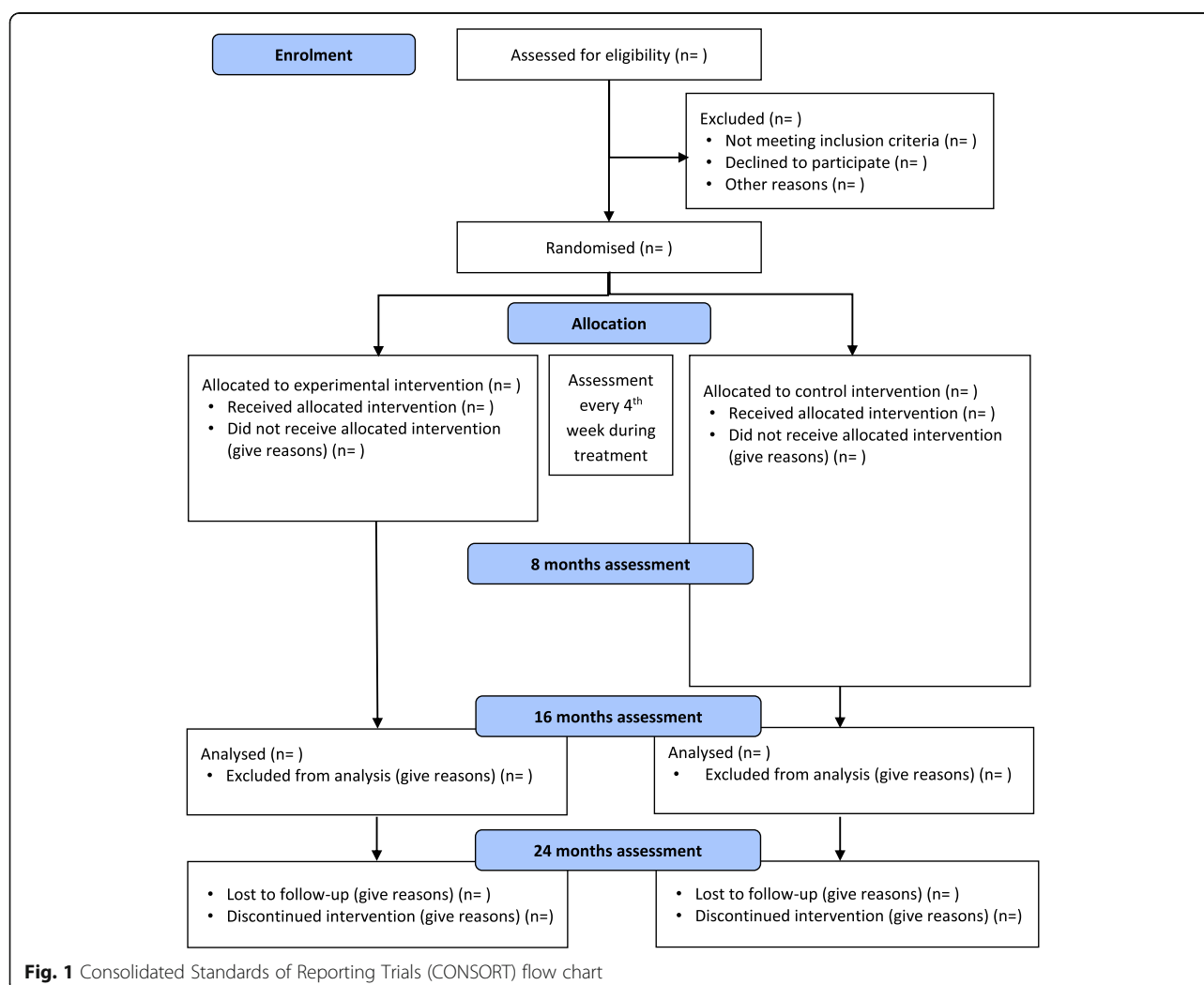
The primary objective of this trial will be to evaluate the beneficial and harmful effects of short-term (20 weeks) MBT compared with long-term (14 months) MBT for adult outpatients with subthreshold or diagnosed borderline personality disorder. We will evaluate the treatments on the primary outcome (borderline symptomatology), secondary outcomes (self-harm incidents, quality of life, and functional impairment, and exploratory outcomes (psychiatric symptoms). Measures of personality functioning, attachment, borderline psychopathology, group

cohesion, and mentalization skills will be included as predictor and mediator variables.

Design

We have designed an investigator-initiated parallel-group single-centre randomized clinical superiority trial of short-term versus long-term MBT for outpatients with subthreshold or diagnosed borderline personality disorder. The Consolidated Standards of Reporting Trials (CONSORT) flow chart for the trial is shown in Fig. 1. [31, 32]. The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) participant timeline is given in Fig. 2, and the SPIRIT checklist is given in Additional file 1 [33].

We will consider for participation all patients referred to the trial site. Patients will be included in the trial, if they comply with the eligibility criteria listed in Table 1. There are inclusion and exclusion criteria as part of the procedure for clinical intake at the trial site, and criteria specific to this trial. For a detailed overview of typical



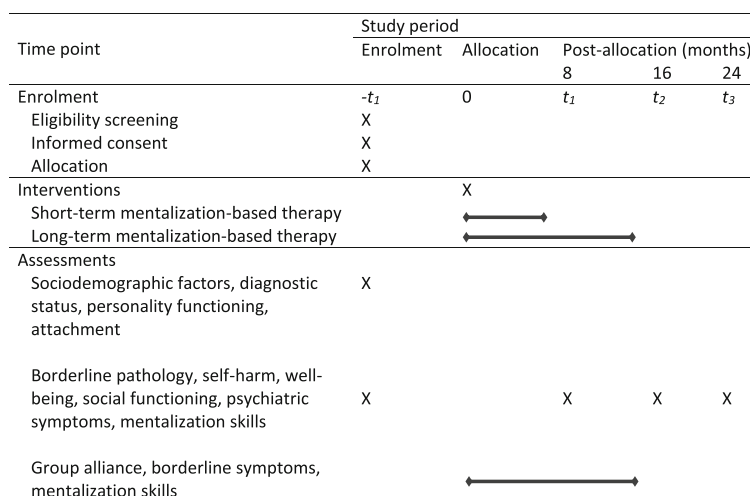


Fig. 2 Participant timeline for the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)

patient characteristics at the trial site, see Simonsen et al. [34].

We will include participants with at least subthreshold borderline personality disorder. According to the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V) [1], the threshold for a full diagnosis is five out of nine diagnostic criteria. However, there is increasing evidence that even four confirmed diagnostic criteria are not qualitatively different from diagnosed borderline personality disorder in terms of impairment, and that the diagnostic threshold should be more inclusive than established by the DSM system to reflect the dimensionality of the construct [35, 36]. For this reason, previous trials have included participants with at least a subthreshold diagnosis [24, 37], and we will do the same in this trial.

Trial site and personnel

The trial site is the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark (from now on referred to as “the clinic”). The clinic specializes in MBT for borderline personality disorder. Patients are

referred from the Capital Region of Denmark via a central visitation unit, where they are initially screened for eligibility before referral to the clinic. Once referred to the clinic, psychiatrists and attending physicians will perform the initial selection and screening of a participant to the trial and collect informed consent. The principal investigator, sponsor-investigator, or a trained research assistant will then conduct the baseline assessments of the participant. All post-baseline assessments will be carried out by trial investigators who are blind to treatment allocation.

Trial therapists provide therapy to both the short-term and long-term treatment groups. Before commencing the trial, all trial therapists at the clinic will have received training in the short-term MBT program by trial investigators and national and international MBT specialists. The training covers relevant topics like case formulations, termination of psychotherapy, and case-specific supervision. The training will continue throughout the trial period. Therapist treatment fidelity will be rated by an independent certified rater. This will allow us to investigate whether the delivered interventions adhere to the MBT manual.

Table 1 Eligibility criteria

	Criteria exclusive to the outpatient clinic	Criteria exclusive to the trial
Inclusion criteria	Aged 18–60 Personality disorder(s) considered to be the primary diagnosis/diagnoses	A minimum of four confirmed DSM-V diagnostic criteria for borderline personality disorder Written informed consent
Exclusion criteria	Possibility of a learning disability (IQ < 75) A full diagnosis of schizotypal personality disorder or antisocial personality disorder Presence of a comorbid psychiatric disorder that requires specialist treatment Current (past 2 months) substance dependence including alcohol Concurrent psychotherapeutic treatment outside the clinic	Unable to understand Danish Lack of informed consent

DSM-V Diagnostic and Statistical Manual of Mental Disorders, 5th edition

Randomization

Copenhagen Trial Unit, a Danish center for clinical intervention research, will be responsible for the central randomization. Trial investigators will call designated staff at Copenhagen Trial Unit using a central telephone to randomize eligible participants to either the experimental group or the control group with a 1:1 allocation according to a computer-generated allocation sequence with permuted blocks of various sizes generated by Copenhagen Trial Unit and unknown to the investigators, secretaries, and clinical staff at the trial site. This is done to eliminate any predictability in the random sequence. The randomization is stratified by (1) sex and (2) high/low scores on the primary outcome measure, the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) [38] at baseline.

Interventions

The short-term MBT program delivered in this trial is overall similar to the existing long-term program, but differs structurally in the following ways: (1) the short-term program is lower in treatment intensity (both duration and exposure), (2) the same therapists provide both group and individual sessions in the short-term program (conjoined psychotherapy), whereas the group therapy and individual therapy are provided by different therapists in the long-term program (combined psychotherapy), and (3) the short-term program is structured in closed groups, in which all participants start and finish the program together, whereas the long-term program is structured as slow-open groups, in which a new participant can enter a group when another finishes. Both interventions in this trial adhere to the treatment guidelines provided by the National Institute for Health and Care Excellence [39].

Experimental intervention

The short-term MBT program is designed as a 20-week psychotherapy program consisting of five sessions of introductory MBT (MBT-I) followed by 15 sessions of group MBT (MBT-G) accompanied by conjoined individual sessions every second week and two psycho-educative meetings with participants and their relatives. Seven to nine participants and two therapists will be included in each short-term MBT group. The groups are closed to enhance cohesion between group participants. A total of 11 short-term MBT groups will be included in this trial.

Originally, MBT-I was a 12-session introductory psycho-educative program covering relevant topics like personality disorders, attachment, and mentalization [40]. The original manual has been modified for our 5-week intervention. A copy of our modified manual is available upon request. After the completion of MBT-I,

the same group of participants will move on to MBT-G, which consists of 15 sessions of mentalization-based psychotherapy in groups, as manualized by Bateman and Fonagy [14]. In our short-term MBT program, group sessions will be accompanied by individual psychotherapy every second week with one of the two group therapists. As part of the individual therapy, a case formulation will be prepared and subsequently shared by the participants in the group. The overall purpose of the individual sessions is for the therapist and participant to develop a consensus of the participant's main difficulties and to establish psychotherapeutic focus points in the group therapy. Furthermore, participants and relatives will be invited to two psycho-educative meetings hosted by the therapists at the beginning of the treatment program to enhance the mentalization work at home. The participants in the short-term MBT program will furthermore be offered three individual follow-up sessions after the end of treatment.

Control intervention

Long-term MBT is organized as a 14-month program and has been implemented at the clinic for the past 7 years. All participants randomized to long-term MBT will initially enter a 6-week MBT-I program manualized by Karterud and Bateman [40] and modified for our 6-week intervention in collaboration with the authors. New MBT-I groups commence every time new participants are recruited and randomized to long-term MBT. A maximum of 12 participants can enter an MBT-I group. When MBT-I finishes, participants will be allocated to one of eight slow-open MBT treatment groups. MBT-G is then organized as 12 months of weekly group therapy sessions, also manualized by Bateman and Fonagy [14]. In the long-term MBT program, group sessions will be accompanied by combined individual MBT sessions every 2 weeks throughout the program. As part of the individual therapy, a case formulation will be developed and subsequently shared with the group by the participant. Furthermore, participants and relatives will also be invited to two psycho-educative meetings hosted by the therapists to enhance the mentalization work at home. When a participant drops out or completes MBT-G, a new participant can start in the group. This procedure continues until the target sample size has been randomized to the long-term MBT program. The participants in the long-term MBT program will also be offered three individual follow-up sessions after the end of treatment.

Concomitant interventions

Participants who are receiving psychotropic treatment will be allowed to continue their medical treatment while participating in the trial. The medical protocol will

follow national as well as international medical recommendations for the treatment of borderline personality disorder and comorbid disorders [39, 41]. Psychiatrists or attending physicians at the clinic will assess the need for additional psychotropic treatment and are asked to adhere to the guidelines. All participants, regardless of treatment condition, will be asked about their current medication by trial personnel during trial interviews to allow us to measure any potential differences in the use of psychotropic medication between the groups.

Baseline assessment at trial intake

Baseline assessments will be carried out prior to randomization by the principal investigator, the sponsor-investigator, and a trained research assistant, all of whom are also clinical psychologists. *General psychopathology* will be assessed with the Mini International Neuropsychiatric Interview (MINI) [42]. *Personality disorders* will be assessed with the Structured Clinical Interview for DSM-V Personality Disorders (SCID-5-PD), formerly known as SCID-II [43]. The SCID-5-PD is considered the gold standard for clinician-administered semi-structured interviews designed to assess personality disorders according to DSM-V criteria [44].

Outcomes

Primary outcome

The primary outcome is the severity of borderline symptomatology assessed with the ZAN-BPD [38], which is a clinician-administered scale for the assessment of change in borderline psychopathology over time. Each of the nine borderline personality disorder criteria are rated on a 0 to 4 anchored scale reflecting the severity of symptoms. The rating is intended to reflect both the frequency and the severity of borderline psychopathology. The interview provides a total score of borderline psychopathology ranging from 0 to 36. ZAN-BPD will be assessed by investigators blind to treatment allocation at baseline, and at the 8-, 16-, and 24-month follow-ups. We will video-record interviews to allow an assessment of inter-rater reliability based on the intraclass correlation coefficient. The results will be evaluated using the guidelines provided by Cicchetti [45].

Secondary outcomes

Functional impairment will be assessed with the Work and Social Adjustment Scale (WSAS) [46, 47]. This self-report scale will be assessed at baseline, and at the 8-, 16-, and 24-month follow-ups. *Quality of life* will be assessed with the Short-Form Health Survey (SF-36) [48], which consists of a mental and a physical component. We will use the mental component as a secondary outcome and the physical component as an exploratory outcome. This self-report scale will

be given at baseline, and at the 8-, 16-, and 24-month follow-ups. *Global functioning* will be measured with the Global Assessment of Functioning (GAF) split version [49]. GAF will be assessed by investigators blind to treatment allocation at baseline, and at the 8-, 16-, and 24-month follow-ups. Inter-rater reliability will be calculated using the previously mentioned guidelines. *Severe self-harm* (dichotomous data) will be measured as the proportion of participants with severe self-harm defined as deliberate acts of self-harm resulting in visible tissue damage. Self-harm will be assessed by investigators blind to treatment allocation using the Suicide and Self-harm Inventory (SSHI) (citation) at baseline, and at the 8-, 16-, and 24-month follow-ups.

Exploratory outcomes

Symptom distress will be measured with the Global Severity Index (GSI) of the Symptom Checklist 90-R (SCL-90-R) [50]. SCL-90-R will be given at baseline, and at the 8-, 16-, and 24-month follow-ups.

Potential predictors and mediators

Questionnaires are given at baseline and every fourth week throughout the intervention period for both intervention groups to allow us to explore predictors and mechanisms of change. In a separate statistical analysis plan, which will be submitted for publication before data collection is completed in this trial, we will describe how these exploratory analyses will be performed. The following predictors and mediators will be investigated.

Personality functioning will be assessed with the Levels of Personality Functioning Scale, Brief Form (LPFS-BF) [51], which is a newly developed brief 12-item self-report questionnaire assessing levels of personality functioning according to the DSM-V alternative model for personality disorders. *Attachment* will be assessed with the brief self-report Relationship Questionnaire (RQ), which gives continuous and categorical ratings of the four attachment styles [52]. *Borderline symptomatology* will be assessed using the Zanarini Rating Scale for Borderline Personality Disorder, Self-Report Version (ZAN-BPD-SRV) [53]. *Group alliance* will be assessed using the 12-item version of the Group Questionnaire (GQ) [54], which is a brief self-report measure of the three core components of group alliance: alliance to the other participants, alliance to the therapists, and group cohesion as a whole. *Mentalization skills* will be assessed with the 15-item Mentalization Questionnaire (MZQ) [55].

For an overview of all measures and the corresponding time of assessment, see Table 2.

Table 2 Assessments administered at baseline and each follow-up point throughout the trial

Assessment points	Self-report measures	Expert ratings
Baseline	LPFS-BF, RQ, SF-36, WSAS, SCL-90-R, MZQ	MINI, SCID-5-PD, ZAN-BPD, GAF, SSHI
Every 4 weeks	GQ, MZQ, ZAN-BPD-SRV	–
Follow-ups at 8, 16, and 24 months	SF-36, WSAS, SCL-90-R, MZQ	ZAN-BPD, SSHI, GAF

GAF Global Assessment of Functioning, GQ Group Questionnaire, LPFS-BF Level of Personality Functioning Scale, Brief Form, MINI Mini International Neuropsychiatric Interview, MZQ Mentalization Questionnaire, RQ Relationship Questionnaire, SCID-5-PD Structured Clinical Interview for DSM-V Personality Disorders, SCL-90-R Symptom Checklist 90-R, SF-36 Short-Form Health Survey 36, SSHI Suicide and Self-Harm Inventory, WSAS Work and Social Adjustment Scale, ZAN-BPD Zanzarini Rating Scale for Borderline Personality Disorder, ZAN-BPD-SRV Zanzarini Rating Scale for Borderline Personality Disorder, Self-Report Version

Blinding

Trial participants and therapists will not be blind to treatment allocation. This is due to the difficulties of implementing an efficient blinding procedure in psychotherapy trials. Baseline assessments will be done before allocation of participants, and the outcome assessments will be performed by blinded assessors. Participants will be instructed to withhold information of their allocation group when assessed. The statistical analyses will be conducted by blinded external statisticians from Copenhagen Trial Unit with the intervention groups coded as A and B. The steering committee will write and agree on two abstracts while the blinding is intact; one assuming the experimental intervention group is A and the control intervention group is B, and the other assuming the opposite. After this, the randomization code will be broken [56, 57].

Participant discontinuation and withdrawal

Participants can withdraw from the trial at any time without giving a reason and without consequences for future treatment at the clinic. To secure data for the trial, a trial investigator will contact the participant and ask what aspects of the trial the participant wishes to withdraw from: (1) the trial intervention or control group, (2) the assessment interviews, or (3) use of already collected data in analyses. If the participant specifies that they wish to withdraw fully and thereby withdraw from all the points above, their data will be deleted and not used in any analysis. The trial investigator will encourage the participant to continue attending the follow-up assessments.

Data management

The data in this trial will be collected using electronic case report forms developed in the data collection system REDCap. The system has been approved by the Danish Data Protection Agency and fulfills the requirements for data security. Data from the interviews will be entered directly into REDCap on a tablet, and all self-report measures will be collected from REDCap. For a detailed overview of outcome measures and data collection time points, see Table 2. The only source data are participants' signed consent forms.

Student assistants employed in the research department but not otherwise involved in the trial will make sure that all self-report measures are sent to participants at the right times, and that the data are complete for all participants enrolled in the trial. The REDCap database has an integrated audit trail to document any access to and changes of the data. The validated data will be exported to SAS for further statistical analyses by statisticians from Copenhagen Trial Unit.

Statistical plan and data analysis

Sample size

The sample size was determined by the predicted change in the primary outcome measure, ZAN-BPD. A 3.5-point superiority margin is considered to be the minimal important difference. Consistent with previous trials that have used ZAN-BPD as an outcome measure for a patient group like ours [24, 58], we expect a standard deviation of 8. With power set at 80% and alpha set at 5% two-tailed, a sample size of 83 participants is needed in each treatment group, corresponding to a total of 166 participants.

Statistical methods

All continuous outcomes will be assessed using linear regression and dichotomous outcomes will be analyzed using logistic regression. The analyses will be based on an intention-to-treat population and will primarily be adjusted for the baseline value of the outcome of interest and the stratification variables used in the randomization. We will secondly adjust all analyses for the following design variables: age (18–30 and 30–60) and functional impairment as assessed with the overall baseline GAF score (0–48 and 49–100) [34].

We will use a five-step procedure [59] to assess if the thresholds for statistical and clinical significance are crossed and we will handle missing data according to the procedure suggested by Jakobsen et al. [60]. A detailed statistical analysis plan will be published before the analyses commence, in which we will provide a detailed description of all primary, secondary, and exploratory analyses. All analyses will be performed blinded with the two intervention groups concealed as A and B.

Interim analyses

An external data safety monitoring committee will perform interim analyses when 50% of the data have been collected according to the good clinical practice guidelines of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use [61]. It will decide whether the trial should stop or carry on. Early stopping criteria will follow the recommendations of Jakobsen et al. [59].

Discussion

This trial will provide evidence of the beneficial and harmful effects of short-term compared to long-term MBT for outpatients with subthreshold or diagnosed borderline personality disorder. To the best of our knowledge, short-term MBT has never been tested before. Gaining more information on how different lengths of treatment work for specific subtypes of patients may help to minimize the potential burden from long-term psychotherapy for some, while at the same time it may identify subtypes of patients for whom short-term psychotherapy is contraindicated. This knowledge may enhance the cost-effectiveness of treatment options for borderline personality disorder. Further, this trial may provide information on the potential predictors and mediators of treatment response.

The present trial has several strengths. First, it has a high degree of external validity because of the relatively inclusive eligibility criteria. Second, the methodology is based on CONSORT, and was predefined and described in detail before randomization began, including, e.g., blinding of all possible parties and implementation of a central randomization system both for generating an allocation sequence and for concealing allocation. Third, the implementation of systematic treatment fidelity ratings allows us to investigate treatment fidelity in both groups.

Our trial also has limitations. First, no systematic review of the effects of short-term compared to long-term psychotherapy for psychiatric disorders is currently available. As mentioned earlier, we are currently performing such a review. Second, the long-term MBT intervention, which is 14 months of treatment in this trial, diverges in intensity (both duration and exposure) from the original 18-month program [13]. This is due to the fixed length of the treatment packages, which have been implemented in the Danish mental health care system. Third, we cannot account for any potential confounding variables because of the structural differences between the groups: the short-term MBT program is closed and conjoined, whereas the long-term program is slow-open and combined.

Dissemination policy

The Danish population will be informed of the trial as well as its final results through national media. The

results of the trial will be presented at all outpatient clinics treating borderline personality disorder in the Mental Health Services, Capital Region of Denmark, by the principal investigator or sponsor-investigator. The final and interim results will be presented at national and international conferences. Further, associations for patients and relatives will be informed about the results of the trial and its future implications. The trial results will be written up by the steering committee and will be published in international peer-reviewed journals. The government of Denmark will be informed of the results before a press release is issued but will have no influence on the reporting of results.

Trial status

The current protocol is version 1, dated 9 October 2018. The first participant was enrolled on 24 September 2018. Recruitment is expected to be completed by September 1, 2020

Additional file

Additional file 1: SPIRIT checklist. (DOCX 61 kb)

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials; DSM-V: Diagnostic and Statistical Manual of Mental Disorders, 5th edition; GAF: Global Assessment of Functioning; GQ: Group Questionnaire; GSI: Global Severity Index; LPFS-BF: Levels of Personality Functioning Scale, Brief Form; MBT: Mentalization-based therapy; MBT-G: Mentalization-based group therapy; MBT-I: Introduction to Mentalization-Based Therapy; MINI: Mini International Neuropsychiatric Interview; MQZ: Mentalization Questionnaire; RQ: Relationship Questionnaire; SCID-5-PD: Structured Clinical Interview for DSM-V Personality Disorders; SCL-90-R: Symptom Checklist 90-R; SF-36: Short-Form Health Survey 36; SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials; SSHI: Suicide and Self-Harm Inventory; WSAS: Work and Social Adjustment Scale; ZAN-BPD: Zangarini Rating Scale for Borderline Personality Disorder; ZAN-BPD-SRV: Zangarini Rating Scale for Borderline Personality Disorder, Self-Report Version

Acknowledgements

We are grateful to the participants and to our colleagues at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre for their hard work and good cooperation.

Funding

The trial is funded by research grants from TrygFonden and from the Mental Health Services Research Foundation, Capital Region of Denmark. Funding applications have undergone anonymous peer review. Neither the funding bodies nor the sponsor will be involved in the collection, analysis, or interpretation of the data, or in writing the manuscripts. The grants will be administered by the head of administration at Stolpegaard Psychotherapy Center. At the end of the trial, the budget will undergo external auditing independently of the trial sponsor and investigators.

Availability of data and materials

The trial investigators, the steering committee, and statisticians at the Copenhagen Trial Unit will have access to the data. After the end of the trial, approval for making the final dataset publicly available in a depersonalized format in the Danish Data Archive will be applied for through the Danish Data Protection Agency.

Trial sponsor

The sponsor of the trial is the Mental Health Services in the Capital Region of Denmark, Kristineberg 3, 2100 Copenhagen Ø.

Protocol amendments

No substantial deviations from the protocol will be implemented without the prior review and approval of the regulatory authorities, except where such may be necessary to eliminate an immediate hazard to the trial participants. In such case, the deviation will be reported to the regulatory authorities as soon as possible. Any deviation, however minor, will be documented and made available in the trial master file in a designated table developed for this purpose.

Authors' contributions

SS and SJ conceived the initial idea of a randomized controlled trial comparing short-term and long-term MBT for borderline personality disorder, along with MS, the acting chief psychologist at the clinic. PS approved of the idea and provided organizational support. SS and SJ developed the trial design with contributions and supervision from JCJ, SL, SP, PS, and AB. SJ wrote the protocol, which was carefully discussed with SS, JCJ, SL, SP, PS, and AB. JCJ designed the plan for the statistical analyses and wrote the sections on sample size and statistical approach. All authors read and approved the final manuscript.

Ethics approval and consent to participate

There are no ethical concerns regarding this trial. Both the experimental intervention and control intervention follow the treatment guidelines published by the National Institute for Health and Care Excellence and the Danish Health and Medicines Authority. Any participants can, at any time, withdraw their consent without any implications for future treatment at the clinic. Prior to commencing the trial, ethical approval was obtained from the regional research ethics committee (ID number H-18023136) and approval was obtained from the Danish Data Protection Agency (I-suite number 6553). During the trial period, data on adverse events will be collected at four time points and will be reported to the regional research ethics committee.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Gentofte, Capital Region of Denmark, Denmark. ²Department of Psychology, University of Copenhagen, Copenhagen, Denmark. ³Copenhagen Trial Unit, Center for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. ⁴St. Ann's Hospital, London, England.

Received: 11 October 2018 Accepted: 19 March 2019

Published online: 05 April 2019

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Arlington: American Psychiatric Pub; 2013.
- Togersen J. In: Oldham J, Skodol AE, Bender DS, editors. Prevalence, Sociodemographics, and Functional Impairment, in Textbook of personality disorders. Washington DC: American Psychiatric Publishing; 2014. p. 109–31.
- Fossati A, et al. Patterns of covariation of DSM-IV personality disorders in a mixed psychiatric sample. *Compr Psychiatry*. 2000;41(3):206–15.
- Zimmerman M, Rothschild L, Chelminski I. The prevalence of DSM-IV personality disorders in psychiatric outpatients. *Am J Psychiatry*. 2005;162(10):1911–8.
- Korzekwa MI, et al. Estimating the prevalence of borderline personality disorder in psychiatric outpatients using a two-phase procedure. *Compr Psychiatry*. 2008;49(4):380–6.
- Conklin CZ, Westen D. Borderline Personality Disorder in Clinical Practice. *Am J Psychiatry*. 2005;162:867–75.
- Zanarini MC, et al. Axis I comorbidity of borderline personality disorder. *Am J Psychiatry*. 1998;155:1733–9.
- Trull TJ, et al. Borderline personality disorder and substance use disorders A review and integration. *Clin Psychol Rev*. 2000;20(2):235–53.
- Barrachina J, et al. Axis II comorbidity in borderline personality disorder is influenced by sex, age, and clinical severity. *Compr Psychiatry*. 2011;52(6):725–30.
- Lieb K, et al. Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. *Br J Psychiatry*. 2010;196(1):4–12.
- Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385(9969):735–43.
- Cristea IA, et al. Efficacy of Psychotherapies for Borderline Personality Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. 2017;74(4):319–28.
- Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355–64.
- Bateman A, Fonagy P. Mentalization-based treatment for personality disorders: a practical guide. Oxford: Oxford University Press; 2016.
- Karterud S. Mentalization-Based Group Therapy (MBT-G): A theoretical, clinical, and research manual. Oxford: OUP Oxford; 2015.
- Vogt KS, Norman P. Is mentalization-based therapy effective in treating the symptoms of borderline personality disorder? A systematic review. *Psychol Psychother*. 2018. <https://doi.org/10.1111/papt.12194>
- Kazdin AE. Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol*. 2007;3:1–27.
- Kvarstein EH, et al. Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—does it make a difference? *Psychol Psychother*. 2015; 88(1):71–86.
- Bales D, et al. Treatment outcome of 18-month, day hospital mentalization-based treatment (MBT) in patients with severe borderline personality disorder in the Netherlands. *J Personal Disord*. 2012;26(4):568–82.
- Jorgensen CR, et al. Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. *Acta Psychiatr Scand*. 2013;127(4):305–17.
- Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563–9.
- Laurensen EM, Luyten P, Kikkert MJ, Westra D, Peen J, Soons MB, Dekker JJ. Day hospital mentalization-based treatment v. specialist treatment as usual in patients with borderline personality disorder: randomized controlled trial. *Psychol Med*. 2018;48(15):2522–29.
- Bateman A, Fonagy P. 8-year follow-up of patients treated for borderline personality disorder: mentalization-based treatment versus treatment as usual. *Am J Psychiatry*. 2008;165(5):631–8.
- Gratz KL, Tull MT, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099–112.
- Blum N, St. John D, Pfohl B, Stuart S, McCormick B, Allen J, Arndt S, Black DW. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for Outpatients With Borderline Personality Disorder: A Randomized Controlled Trial and 1-Year Follow-Up. *Am J Psychiatry*. 2008;165:468–78.
- Bos EH, van Wel EB, Appelo MT, Verbraak MJ. A randomized controlled trial of a Dutch version of systems training for emotional predictability and problem solving for borderline personality disorder. *J Nerv Ment Dis*. 2010; 198(4):299–304.
- McMain SF, et al. A randomized trial of brief dialectical behaviour therapy skills training in suicidal patients suffering from borderline disorder. *Acta Psychiatr Scand*. 2017;135(2):138–48.
- McMain SF. Advances in the Treatment of Borderline Personality Disorder: An Introduction to the Special Issue. *J Clin Psychol*. 2015;71(8):741–6.
- Knekt P, et al. Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*. 2008; 38(5):689–703.
- Lorentzen S, et al. Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial. *Br J Psychiatry*. 2013;203(3):280–7.
- Boutron I, et al. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med*. 2008;148(4):295–309.

32. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med Res Methodol*. 2001;1:2.
33. Chan A, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Ann Intern Med*. 2013;158:200–7.
34. Simonsen S, Heinskou T, Sørensen P, Folke S, Lau ME. Personality disorders: patient characteristics and level of outpatient treatment service. *Nord J Psychiatry*. 2017;71(5):325–31.
35. Clifton A, Pilkonis PA. Evidence for a single latent class of Diagnostic and Statistical Manual of Mental Disorders borderline personality pathology. *Compr Psychiatry*. 2007;48(1):70–8.
36. Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence, and effect of personality disorder. *Lancet*. 2015;385(9969):717–26.
37. Beck E, et al. Mentalization-based treatment in groups for adolescents with borderline personality disorder (BPD) or subthreshold BPD versus treatment as usual (M-GAB): study protocol for a randomized controlled trial. *Trials*. 2016;17(1):314.
38. Zanarini. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Personal Disord*. 2003;17(3):233–42.
39. National Institute for Health and Care Excellence (NICE). Borderline Personality Disorder: the NICE Guideline on Treatment and Management. *Br Psychol Soc R Coll Psychiatrists*. 2009.
40. Karterud S, Bateman A. Manual for mentaliseringsbasert psykoedukativ gruppeterapi (MBT-I). Oslo: Gyldendal akademisk; 2011.
41. Sundhedsstyrelsen, National klinisk retningslinje for behandling af emotionel ustabil personlighedsstruktur, borderline type. 2015.
42. Lecrubier Y, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry*. 1997;12:224–31.
43. First MB, et al. The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part II: Multi-site test-retest reliability study. *J Personal Disord*. 1995;9(2):92–104.
44. Lobbstaël J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clin Psychol Psychother*. 2011;18(1):75–9.
45. Cicchetti DV. Guidelines, Criteria, an Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychol Assess*. 1994;6(4):284–90.
46. Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: Psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personal Ment Health*. 2017;11(4):215–28.
47. Miles HH, Barrabee EL, Finesinger JE. Evaluation of psychotherapy: with a follow up study of 62 cases of anxiety neurosis. *Psychosom Med*. 1951 Mar-Apr;13(2):83–105.
48. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Med Care*. 1992;30(6):473–83.
49. Jones SH, et al. Reliability and Validity of the Global Assessment of Functioning (GAF). *Br J Psychiatry*. 1995;166:654–9.
50. Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: a study in construct validation. *J Clin Psychol*. 1977;33(4):981–9.
51. Hutsebaut J, Feensta D, Kamphuis J. Development and Preliminary Psychometric Evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM-5 Level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). *Personal Disord Theory Res Treat*. 2016;7(2):192–4.
52. Bartholomew K, Horowitz LM. Attachment Styles Among Young Adults: A Test of a Four-Category Model. *J Pers Soc Psychol*. 1991;61(2):226–44.
53. Zanarini, et al. Development of the self-report version of the Zanarini Rating Scale for Borderline Personality Disorder. *Personal Ment Health*. 2015;9(4):243–9.
54. Krogel J, et al. The Group Questionnaire: a clinical and empirically derived measure of group relationship. *Psychother Res*. 2013;23(3):344–54.
55. Hausberg M, Schulz H, Pieglar T, Happach C, Klöpper M, Brütt A, et al. Is a self-rated instrument appropriate to assess mentalization in patients with mental disorders? Development and first validation of the Mentalization Questionnaire (MZQ). *Psychother Res*. 2012;22(6):699–709.
56. Gotzsche PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials*. 1996;17(4):285–90 discussion 290–3.
57. Jarvinen TL, et al. Blinded interpretation of study results can feasibly and effectively diminish interpretation bias. *J Clin Epidemiol*. 2014;67(7):769–72.
58. Priebe S, 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.
59. Jakobsen JC, et al. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol*. 2014;14:34.
60. Jakobsen JC, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.
61. Englev E, Petersen KP. ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives. *Dan Med J*. 2003;165(16):1659–62.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	<u>1</u>
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	<u>2</u>
	2b	All items from the World Health Organization Trial Registration Data Set	<u>1-17</u>
Protocol version	3	Date and version identifier	<u>15</u>
Funding	4	Sources and types of financial, material, and other support	<u>16</u>
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	<u>1, 17</u>
	5b	Name and contact information for the trial sponsor	<u>16</u>
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	<u>16</u>
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	<u>14, 16</u>

Introduction

Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	<u>4-6</u>
	6b	Explanation for choice of comparators	<u>4-6</u>
Objectives	7	Specific objectives or hypotheses	<u>6</u>
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	<u>6</u>

Methods: Participants, interventions, and outcomes

Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	<u>7</u>
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	<u>7</u>
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	<u>8-10</u>
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	<u>13</u>
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	<u>7</u>
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	<u>10</u>
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	<u>10-12</u>
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	<u>7</u>

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	<u>13</u>
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	<u>7</u>

Methods: Assignment of interventions (for controlled trials)

Allocation:

Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	<u>8</u>
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	<u>8</u>
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	<u>8</u>
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	<u>12</u>
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	<u>12</u>

Methods: Data collection, management, and analysis

Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	<u>10-11</u>
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	<u>13</u>

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	<u>13</u>
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	<u>13-14</u>
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	<u>13-14</u>
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	<u>14</u>
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	<u>14</u>
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	<u>14</u>
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	<u>16</u>
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	<u>17</u>
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	<u>16</u>
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	<u>17</u>

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	<u>7</u>
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	<u>not relevant</u>
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	<u>13</u>
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	<u>16</u>
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	<u>16</u>
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	<u>not relevant</u>
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	<u>15</u>
	31b	Authorship eligibility guidelines and any intended use of professional writers	<u>15</u>
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	<u>16</u>
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	<u> </u>
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	<u>not relevant</u>

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons [“Attribution-NonCommercial-NoDerivs 3.0 Unported”](#) license.

Paper 4

UPDATE

Open Access



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)

Sophie Juul^{1,2,3*} , Sebastian Simonsen¹, Stig Poulsen², Susanne Lunn², Per Sørensen¹, Anthony Bateman⁴ and Janus Christian Jakobsen^{3,5}

Abstract

Background: Psychotherapy for borderline personality disorder is often extensive and resource-intensive. Mentalisation-based therapy is a psychodynamically oriented treatment option for borderline personality disorder, which includes a case formulation, psychoeducation, and group and individual therapy. The evidence on short-term compared with long-term mentalisation-based therapy is currently unknown.

Methods/design: The Short-Term MBT Project (MBT-RCT) is a single-centre, parallel-group, investigator-initiated, randomised clinical superiority trial in which short-term (20 weeks) will be compared with long-term (14 months) mentalisation-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder. Outcome assessors, data managers, the data safety and monitoring committee, statisticians, and decision-makers will be blinded to treatment allocation. Participants will be assessed before randomisation and at 8, 16, and 24 months after randomisation. The primary outcome will be the severity of borderline symptomatology assessed with the Zanarini Rating Scale for Borderline Personality Disorder. Secondary outcomes will be functional impairment (Work and Social Adjustment Scale), quality of life (Short-Form Health Survey 36—mental component), global functioning (Global Assessment of Functioning), and proportion of participants with severe self-harm. In this paper, we present a detailed statistical analysis plan including a comprehensive explanation of the planned statistical analyses, methods to handle missing data, and assessments of the underlying statistical assumptions. Final statistical analyses will be conducted independently by two statisticians following the present plan.

* Correspondence: sophie.juul@regionh.dk

¹Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Stolpegårdsvej 20, 2820 Gentofte, Denmark

²Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Discussion: We have developed this statistical analysis plan before unblinding of the trial results in line with the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines, which should increase the validity of the MBT-RCT trial by mitigation of analysis bias.

Trial registration: ClinicalTrials.gov [NCT03677037](https://clinicaltrials.gov/ct2/show/study/NCT03677037). Registered on 19 September 2018

Background

Mentalisation-based therapy (MBT) is a long-term, psychodynamically oriented psychotherapy developed specifically to treat patients with borderline personality disorder [1]. Even though more randomised clinical trials at low risk of bias are still needed, long-term MBT is considered one of the most evidence-based interventions currently available for patients with borderline personality disorder [2]. However, the optimal duration of MBT for borderline personality disorder is currently unclear.

The Short-Term MBT Project (MBT-RCT) is a single-centre, parallel-group, investigator-initiated, randomised clinical superiority trial with the objectives to assess the beneficial and harmful effects of short-term (20 weeks) compared with long-term (14 months) MBT for outpatients with subthreshold or diagnosed borderline personality disorder [3]. The Helsinki Declaration [4] and the International Conference on Harmonization of Good Clinical Practice (ICH-GCP) [5] Guidelines recommend that clinical trials should be analysed according to a pre-specified plan to prevent selective outcome reporting bias and data-driven analysis results [6–8].

In this publication, we describe the pre-planned statistical analyses of the primary and secondary outcomes in the MBT-RCT trial. The main publication of the trial results will adhere to this statistical analysis plan as approved by the steering group.

Methods

The design of the MBT-RCT trial has been described in detail previously [3]. The trial population will be adults (18 years of age or older) with subthreshold or diagnosed borderline personality disorder assessed with the Structural Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD). Participants will be eligible for enrolment

if they meet all of the following inclusion criteria and none of the exclusion criteria as presented in Table 1.

The MBT-RCT trial is registered at ClinicalTrials.gov (identifier: NCT03677037), is carried out in compliance with the Declaration of Helsinki [4], and is approved by the Danish Data Protection Agency (approval number: 6553) and by the Regional Research Ethics Committee of the Capital Region of Denmark (approval number: H-18023136).

Randomisation and blinding

Copenhagen Trial Unit, a Danish centre for clinical intervention research, will be responsible for the central randomisation. Randomisation will be performed with a 1:1 allocation according to a computer-generated allocation sequence with permuted blocks of various sizes. The allocation sequence will be concealed from all trial investigators. The randomisation is stratified by (1) sex and (2) high/low baseline scores on the primary outcome measure, the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) [9].

Outcome assessors, data managers, statisticians, the data safety and monitoring committee, and decision-makers will be blinded to treatment allocation [10]. Trial participants and therapists will not be blind to the treatment allocation. This is due to the difficulties of implementing an efficient blinding procedure in trials assessing psychological interventions [10].

Trial interventions

Experimental intervention

The short-term MBT programme is organised as a 20-week programme consisting of 20 weeks of group therapy in closed groups commencing with five sessions of psychoeducative introduction to MBT [11] followed by 15 sessions of group MBT group therapy accompanied

Table 1 Inclusion and exclusion criteria

	General criteria of the outpatient clinic	Criteria exclusive to the trial
Inclusion criteria	- Aged 18–60 - Personality disorder(s) considered to be the primary diagnosis	- A minimum of four confirmed DSM-5 diagnostic criteria for borderline personality disorder - Written informed consent
Exclusion criteria	- Learning disability (IQ < 75) - A full diagnosis of schizotypal personality disorder or antisocial personality disorder - Presence of a comorbid psychiatric disorder that requires specialist treatment - Current (past 2 months) substance dependence including alcohol - Concurrent psychotherapeutic treatment outside the clinic	- Unable to understand Danish - Lack of informed consent

by conjoined individual sessions every second week. All participants are furthermore invited to two psychoeducative meetings with other participants and their relatives. The participants will be treated by two group therapists, one of them also being the individual therapist. Both group and individual therapies are manualised by Bateman and Fonagy [1].

Control intervention

Long-term MBT is organised as a 14-month programme and has been implemented at the clinic for the past 10 years. All participants randomised to long-term MBT will initially enter a 6-week psychoeducative introduction to MBT [11]. New psychoeducative MBT groups commence every time new participants are recruited and randomised to long-term MBT. When the psychoeducative group finishes, participants will be allocated to one of eight slow-open MBT treatment groups. Treatment is then organised as 12 months of weekly group therapy sessions combined with individual therapy every second week. All participants are furthermore invited to two psychoeducative meetings with other participants and their relatives. The participants will be treated by two group therapists and a third individual therapist. Both group and individual therapies are manualised by Bateman and Fonagy [1].

Baseline characteristics

The baseline characteristics will be assessed from inclusion in the trial. A mock table of the complete predefined baseline table can be found in Supplementary Material 1. The baseline characteristics will be as follows:

1. Demographic characteristics:
 - (a) Age
 - (b) Sex
 - (c) Civil status
 - (d) Living situation (alone/with others)
 - (e) Education level
 - (f) Employment status
2. Clinical characteristics:
 - (a) Psychiatric comorbidity, e.g. a diagnosis of an anxiety disorder, major depressive disorder, post-traumatic stress disorder (reported if the frequency is above or equal to 10% in any of the intervention groups)
 - (b) Proportion of participants with subthreshold borderline personality disorder
 - (c) Mean number of borderline personality disorder diagnostic criteria
 - (d) Personality disorder comorbidity (reported if the frequency is above or equal to 10% in any of the intervention groups)

- (e) Proportion of participants with one or more suicide attempts 8 months prior to randomisation
 - (f) Proportion of participants with severe self-harm incidents defined as deliberate acts of self-harm resulting in visible tissue damage 8 months prior to randomisation
 - (g) Proportion of participants on psychopharmacological medication (e.g. antidepressants, antipsychotics) at baseline
3. Mean (SD) number of days from randomisation to assessment time point for all time points (8, 16, and 24 months post-randomisation)

Outcomes

The outcomes were predefined as primary, secondary, and exploratory [3]. This publication describes the statistical analysis plan of the primary and secondary outcomes only.

Primary outcome

- Severity of borderline symptomatology assessed with the Zanarini Rating Scale for Borderline Personality Disorder [9]

Secondary outcomes

- Functional impairment (assessed with the Work and Social Adjustment Scale (WSAS) [12])
- Quality of life (assessed with the Short-Form Health Survey (SF-36) mental component) [13]
- Global functioning (assessed with the Global Assessment of Functioning scale (GAF)) [14]
- Severe self-harm (defined as the proportion of participants with one or more deliberate acts of self-harm resulting in visible tissue damage)

Exploratory outcomes

- Symptom distress (assessed with the Symptom Checklist 90 (SCL-90)) [15]
- Quality of life (assessed with the Short-Form Health Survey (SF-36) physical component [13])

Assessment time points

All outcomes will be assessed at baseline and at 8, 16, and 24 months after randomisation. Investigator-administered outcomes (severity of borderline symptomatology, severe self-harm, and global functioning) will be assessed by blinded assessors at all time points. We will use the 16-month time point as the primary time point of interest, as it is the time point closest to the end of treatment in the long-term MBT group. In an

exploratory analysis, we will consider reporting the results of the comparison between the end of treatment in both groups (i.e. data from the 8 months time point in the short-term group compared with data from the 16 months time point in the long-term group). Data from the 24-month time point, as well as results of the exploratory outcomes, will be analysed using the same principles as described in this statistical analysis plan and published in a separate publication.

Safety

We will report the proportion of participants with one or more serious adverse events in both groups. We will use the International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use—Good Clinical Practice (ICH-GCP) definition of a serious adverse event, which is any untoward medical occurrence that resulted in death, was life-threatening, required hospitalisation or prolonging of existing hospitalisation, and resulted in persistent or significant disability or jeopardised the participant [5]. Two investigators will independently go through the participants' medical journals and assess possible serious adverse events at the 16 and 24 months time point of assessment according to the ICH-GCP definition.

Sample size and power estimations

The sample size estimation was based on the primary outcome, and our primary conclusions will be based on the results of the primary outcome. The outcomes in our outcome hierarchy were ranked according to clinical relevance, and we estimated the power of each non-primary outcome to ensure that we had sufficient power to confirm or reject minimally important intervention effects [16].

Sample size estimation

The sample size was determined by the predicted change in the primary outcome measure, ZAN-BPD. We considered a 3.5-point superiority margin to be the minimal important difference. Consistent with previous trials that have used ZAN-BPD as an outcome measure for a group of participants similar to ours [17, 18] we expect a standard deviation of 8. With power set at 80% and alpha set at 5% two-tailed, a sample size of 83 participants will be needed in each intervention group, corresponding to a total of 166 participants.

We did not adjust the sample size according to missing data. We plan to have as close to 0% missing data as possible. However, if missing data occur, we will use multiple imputations which will limit the loss of power (please see the '[Handling of missing data](#)' section).

Power estimation for secondary outcomes

For the secondary outcomes, we have performed power calculations as presented in Tables 2 and 3 [16].

General analysis principles

Statistical analyses will be performed in Stata [24]. All analyses will be conducted according to the intention-to-treat principle (ITT). The intention-to-treat population will include all randomised participants, regardless of missing data, lost to follow-up, or adherence to the intervention. Thus, by performing an intention-to-treat analysis, we will assess the effects of being randomised to the interventions. We will consider performing a per-protocol analysis, if the number of participants who prematurely drops out of treatment exceeds 5% of the total trial population. By performing a per-protocol analysis, we will assess the effects of adhering to the intervention, which must be considered hypothesis-generating only.

It is generally recommended that regression analyses should be adjusted for the stratification variables used in the randomisation [25–27]. Thus, all analyses will primarily be adjusted for the stratification variables used in the randomisation (and the baseline value of the outcome of interest when assessing continuous outcomes). We will secondly adjust all analyses for the following adjustment variables: age (18–30/31–60), baseline global functioning as assessed with the GAF score (0–48/49–100), baseline proportion of participants with severe self-harm 8 months prior to randomisation (participants with one or more events/participants with no events), and proportion of participants who had their group therapy temporarily paused due to COVID-19 in March 2020 and January 2021 compared to the proportion of participants who did not have their group therapy temporarily paused.

We will perform the following subgroup analyses (test of interaction):

- Baseline severity of borderline symptomatology (ZAN-BPD scores 0–11/12–36)
- Sex (male/female)
- Age (18–30/31–60)
- Baseline global functioning (GAF scores 0–48/49–100)
- Baseline proportion of participants with severe self-harm incidents 8 months prior to randomisation (participants with one or more events/participants with no events)
- Proportion of participants who had their group therapy temporarily paused due to COVID-19 in March 2020 and January 2021 compared to participants who did not have their group therapy temporarily paused

Table 2 Power estimations for the secondary continuous outcomes

Outcome	Minimal clinically important difference	Expected standard deviation	Alpha	Sample size	Power	Reference
	δ	σ	α	n	$1-\beta$	
Functional impairment (WSAS)	4.5	10	5	166	82	Phillips et al. [19]
Quality of life (SF-36—mental component)	5	11	5	166	83	Rollman et al. [20]
Global functioning (GAF)	4	8.5	5	166	85	Jørgensen et al. [21]

GAF Global Assessment of Functioning, SF-36 Short-Form Health Survey 36, WSAS Work and Social Adjustment Scale

We will present the results of the subgroup analyses in forest plots.

Trial profile

The flow of trial participants will be displayed in a Consolidated Standards of Reporting Trials (CONSORT) diagram [28]. The number of screened patients who were assessed for eligibility, and the number included in the primary and secondary analyses, as well as all reasons for exclusions in the primary and secondary analyses, will be reported.

Statistical analyses

Analysis of continuous data

Continuous outcomes will be presented as means and standard deviations for each group together with 95% confidence intervals for the means of the groups and the mean differences between the groups. We will analyse the continuous outcomes using linear regression. All variables will be included as fixed effects.

Analysis of dichotomous data

Dichotomous outcomes will be presented as proportions of participants in each group with the event, together with risk ratios with 95% confidence intervals. We will analyse the dichotomous outcomes using logistic regression. All variables will be included as fixed effects. Odds ratios will be transformed to risk ratios estimating marginal effects using the NLCOM command in Stata [24].

Level of significance

The threshold for significance will be assessed according to a five-step procedure, suggested by Jakobsen and colleagues [29].

The first step will be to calculate and report confidence intervals and p -values for the primary and secondary outcomes. All confidence intervals will be 95% and

two-sided. We will use a p -value of less than 0.05 as the threshold for statistical significance for our primary outcome (see the ‘Sample size estimation’ section) since we plan to report on only one primary outcome. Since our primary conclusions will be based on one outcome result at one time point (16 months post-randomisation), we will limit problems associated with multiple testing due to multiple outcome comparisons [30, 31]. All remaining outcome results and assessment time points will be considered hypothesis-generating only.

The second step will be to calculate and report the Bayes factor [32] for primary and secondary outcomes. The Bayes factor is the ratio between the probability of the results given that the null hypothesis (H_0) is true divided by the probability of the results given that the alternative hypothesis (H_A) is true [32]. Calculating and reporting the Bayes factor will allow us to interpret the results of the primary outcome in relation to former trial results [17, 18].

The third step will be to use Lan-DeMets monitoring boundaries if the trial is stopped before the sample size is reached [33]. This is done to avoid a potential false rejection of the null hypothesis caused by an insufficient sample size [34].

The fourth step regarding adjustment of p -values based on multiple testing of the primary outcome is not applicable to our trial. We only have one single primary outcome, primarily assessed at one time point (16 months post-randomisation) [29].

The fifth step is the assessment of the clinical significance. The assessment of the clinical significance of our trial results will be based on the intervention effects we predefined in the sample size and power estimations.

Interim analyses

We have pre-planned one interim analysis, which will be conducted after half of the trial participants have been

Table 3 Power estimations of secondary dichotomous outcomes

Outcome	Expected proportion in the control group	Alpha significance level	Power	Reference
Severe self-harm incidents ^a	15%	5	14%	Simonsen et al. [22]; Bateman and Fonagy [23]

^aEven though the expected power is only 14%, we define severe self-harm incidents as a secondary outcome, because it is considered very important for this population

assessed at the 8 months post-randomisation time point. The timing and prevalence of any additional interim analyses will be decided exclusively by the members of the data monitoring and safety committee. The role of the data monitoring and safety committee will be to make recommendations to the steering group to either continue, change, hold, or terminate the trial. This recommendation will primarily be based on safety considerations. The data monitoring and safety committee will be provided with the following trial data: number of participants randomised, number of participants per intervention group, baseline ZAN-BPD scores for all participants, ZAN-BPD scores at the 8-month post-randomisation time point for participants in both intervention groups with available data at that time point, proportion of participants with one or more deliberate acts of self-harm at the 8-month post-randomisation time point, and serious adverse events. Based on evaluations of these data, the data monitoring and safety committee will decide whether they want further data from the principal investigator and when next to perform analyses on data.

Handling of missing data

Missing data will be handled according to the recommendations of Jakobsen and colleagues [35]. In short, if we experience missing data, we will consider to use multiple imputations and use best-worst/worst-best case scenarios to assess the potential impact of the missing data [35].

All randomised participants (the intention-to-treat population) will be included in the primary analysis of all outcomes. If it is not valid to ignore missing data (that is, if the missing data exceeds 5%), we will consider using multiple imputations and use best-worst/worst-best case scenarios to assess the potential impact of the missing data [35]. Best-worst and worst-best case scenarios assess the potential range of impact of the missing data for the trial results [35]. In the 'best-worst' case scenario, it is assumed that all participants lost to follow-up in the short-term group have had a beneficial outcome (e.g. had no self-harm incidents), and all those in the long-term group have had a harmful outcome (e.g. had one or more self-harm incidents). Conversely, in the 'worst-best' case scenario, it is assumed that all participants who were lost to follow-up in the short-term group have had a harmful outcome and that all those lost to follow-up in the long-term group have had a beneficial outcome [35]. When continuous outcomes are used, a 'beneficial outcome' will be defined as the group mean plus two SDs of the group mean (fixed imputation), and a 'harmful outcome' will be defined as the group mean minus two SDs of the group mean (fixed imputation) [35]. We do not expect any missing baseline

data, as participants will only be randomised once they have a complete baseline dataset.

Assessments of underlying statistical assumptions

We will assess the underlying statistical assumptions for all statistical analyses [36, 37]. We will test for major interactions between each covariate and the intervention variable for all regression analyses. We will, in turn, include each possible first-order interaction between included covariates and the intervention variable. For each combination, we will test if the interaction term is significant and we will assess the effect size. We will only consider concluding that there is evidence of an interaction if (1) the interaction is statistically significant following the Bonferroni-adjusted thresholds (0.05 divided by the number of possible interactions) and (2) if the interaction shows a clinically significant effect. If we conclude that the interaction is statistically significant, we will consider both presenting a separate analysis for each interaction as well as an overall analysis including the interaction term in the model [36].

Assessments of underlying statistical assumptions for linear regression

We will visually inspect quantile-quantile plots of the residuals [38, 39] to assess if the residuals are normally distributed, and we will use residuals plotted against covariates and fitted values [38, 39] to assess for homogeneity of variances. If the plots show deviations from the model assumptions, we will consider transforming the outcome, i.e. by using log transformation or square root and/or use robust standard errors [38, 39].

Assessments of underlying statistical assumptions for logistic regression

We will assess if the deviance divided by the degrees of freedom is significantly larger than 1 to assess for relevant overdispersion. If that is the case, we will consider using a maximum likelihood estimate of the dispersion parameter.

Statistical reports

Two independent statisticians will analyse blinded data on all outcomes with intervention groups concealed as, e.g., 'A' and 'B'. Two independent statistical reports will be delivered to the principal investigator (SJ) and will be shared with the steering group. If there are discrepancies between the two primary statistical reports, these will be identified and we will then consider which is the most correct result. A final statistical report will be prepared, and all two (or three, if anything is to be corrected) statistical reports will be published as a [supplementary material](#). Mock tables are presented in Supplementary material 1.

Discussion

The primary aim of this paper is to minimise the risks of bias associated with selective outcome reporting and erroneous data-driven results. We therefore present a pre-defined statistical analysis plan for the MBT-RCT trial.

Strengths

Our methodology has several strengths. First, our methodology is pre-defined, and our analyses will adhere to this statistical analysis plan. Second, we have limited problems with multiplicity because we only assess one primary outcome, and our conclusions will primarily be based on the results of the primary outcome [29]. Third, all analyses will be conducted according to the intention-to-treat principle, and, if necessary, we will use multiple imputations and best-worst/worst-best case scenarios to assess the potential impact of missing data [35]. Furthermore, we plan to systematically assess if the underlying statistical assumptions are fulfilled for all statistical analyses.

Limitations

A potential limitation of the MBT-RCT trial is that no systematic review of the effects of short-term compared to long-term psychotherapy for borderline personality disorder, or for psychiatric disorders in general, was available prior to planning of this trial. Hence, estimations of anticipated intervention effects, estimations of variances used in our sample size, and power estimations, etc. may be erroneous. We are currently performing such a review, which will be submitted for publication prior to completion of this trial [40]. Second, we expect a significant amount of missing data, due to the instability of the trial population. Even though we plan to handle missing data appropriately, no statistical method can guarantee the validity of trial results if the missingness is substantial. Third, even though the trial will be sufficiently powered to confirm or reject intervention effects on the primary and secondary outcomes, the relatively small number of randomised participants may result in a risk of baseline differences which also may bias the trial results especially on non-primary outcomes. However, we will carefully consider the low sample size when interpreting the trial results. Fourth, as participants are not blinded to the allocated treatment, results from all participant-reported outcomes are at risk of bias [10]. Fifth, therapists are likewise not blinded to the allocated treatment and may have an allegiance to one of the interventions. Sixth, we have planned several subgroup analyses. However, subgroup analyses are per definition underpowered, and will be considered hypothesis-generating only. We will carefully consider these limitations when interpreting the results.

Conclusion

We have developed this statistical analysis plan in line with the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines, which should increase the validity of the MBT-RCT trial by mitigation of analysis bias.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-021-05450-y>.

Additional file 1: Table S1. Baseline characteristics of the trial population. BPD: Borderline Personality Disorder; MINI: Mini International Neuropsychiatric Interview; No.: Number; SCID-5-PD: Structured Clinical Interview for DSM-5 Personality Disorders; SD: Standard Deviation. **Table S2.** Primary and secondary outcome results (ITT Population). GAF: Global Assessment of Functioning; ITT: Intention To Treat; SF-36: Short Form Health Survey – 36; SD: Standard Deviation; WSAS: Work and Social Adjustment Scale; ZAN-BPD: Zanarini Rating Scale for Borderline Personality Disorder.

Acknowledgements

We are grateful to the participants and to our colleagues at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre for their hard work and good cooperation.

Authors' contributions

SJ drafted this protocol update, receiving ongoing supervision from JCJ. All authors amended and approved the final manuscript.

Funding

The MBT-RCT trial is funded by research grants from TrygFonden and from the Mental Health Services Research Foundation, Capital Region of Denmark. Funding applications have undergone anonymous peer review. Neither the funding bodies nor the sponsor will be involved in the collection, analysis, or interpretation of the data, or in writing the manuscripts. The grants will be administered by the head of administration at Stolpegaard Psychotherapy Center. At the end of the trial, the budget will undergo external auditing independently of the trial sponsor and investigators.

Availability of data and materials

Trial investigators, the steering committee, and statisticians at the Copenhagen Trial Unit will have access to the data. After the end of the trial, approval for making the final dataset publicly available in a depersonalised format in the Danish Data Archive will be applied for through the Danish Data Protection Agency.

Declarations

Ethics approval and consent to participate

There are no ethical concerns regarding this trial. Both the experimental intervention and control intervention follow the treatment guidelines published by the National Institute for Health and Care Excellence and the Danish Health and Medicines Authority. Any participants can, at any time, withdraw their consent without any implications for future treatment at the clinic. Prior to commencing the trial, ethical approval was obtained from the regional research ethics committee (ID number H-18023136), and approval was obtained from the Danish Data Protection Agency (I-suite number 6553). During the trial period, data on serious adverse events will be collected and will be reported to the regional research ethics committee.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Stolpegårdsvej 20, 2820 Gentofte, Denmark. ²Department of Psychology, University of Copenhagen, Copenhagen, Denmark. ³Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. ⁴Anna Freud Centre, Kantor Centre of Excellence, London, UK. ⁵Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

Received: 15 January 2021 Accepted: 12 July 2021

Published online: 28 July 2021

References

- Bateman A, Fonagy P. Mentalization-based treatment for personality disorders: a practical guide. Oxford, England: Oxford University Press; 2016. <https://doi.org/10.1093/med:psych/9780199680375.001.0001>.
- Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020;5:CD012955. <https://doi.org/10.1002/14651858.CD012955.pub2>.
- Juil S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196. <https://doi.org/10.1186/s13063-019-3306-7>.
- World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191–4.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Guideline: Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice (ICH-GCP). 2015. Available at: <https://ichgcp.net/da> [Accessed July 9, 2021].
- Dwan K, Gamble C, Williamson PR, Altman DG. Reporting of clinical trials: a review of research funders' guidelines. *Trials*. 2008;9(1):66. <https://doi.org/10.1186/1745-6215-9-66>.
- Thomas L, Peterson ED. The value of statistical analysis plans in observational research: defining high-quality research from the start. *JAMA*. 2012;308(8):773–4. <https://doi.org/10.1001/jama.2012.9502>.
- Gabriel SE. Getting the methods right—the foundation of patient-centered outcomes research. *New Engl J Med*. 2012;367(9):787–90. <https://doi.org/10.1056/NEJMp1207437>.
- Zanarini MC. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Personal Dis*. 2003;17(3):233–42. <https://doi.org/10.1521/pedi.17.3.233.22147>.
- Juil S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med*. 2021;26(3):109. <https://doi.org/10.1136/bmjebm-2020-111407>.
- Karterud S, Bateman A. Manual for mentalization based psychoeducational group therapy (MBT-I). Oslo: Gyldendal Akademisk; 2011.
- Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personal Ment Health*. 2017;11(4):215–28. <https://doi.org/10.1002/pmh.1382>.
- Brazier JE, Harper R, Jones N, O'cathain A, Thomas K, Usherwood T, et al. Validating the SF-36 Health Survey Questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160–4. <https://doi.org/10.1136/bmj.305.6846.160>.
- Hall RC. Global assessment of functioning: a modified scale. *Psychosom*. 1995;36(3):267–75. [https://doi.org/10.1016/S0033-3182\(95\)71666-8](https://doi.org/10.1016/S0033-3182(95)71666-8).
- Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary report. *Psychopharmacol Bull*. 1973;9(1):13–28.
- Jakobsen JC, Ovesen C, Winkel P, Hilden J, Gluud C, Wetterslev J. Power estimations for non-primary outcomes in randomised clinical trials. *BMJ Open*. 2019;9(6):e027092. <https://doi.org/10.1136/bmjopen-2018-027092>.
- Gratz K, Tull M, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099–112. <https://doi.org/10.1017/S0033291713002134>.
- Priebe S, Bhatti N, Barnicot K, Bremner S, Gaglia A, Katsakou C, 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. <https://doi.org/10.1159/000338897>.
- Phillips R, Schneider J, Molosankwe I, Leese M, Foroushani PS, Grime P, et al. Randomized controlled trial of computerized cognitive behavioural therapy for depressive symptoms: effectiveness and costs of a workplace intervention. *Psychol Med*. 2014;44(4):741–52. <https://doi.org/10.1017/S0033291713001323>.
- Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Houck PR, Counihan PJ, et al. Telephone-delivered collaborative care for treating post-CABG depression: a randomized controlled trial. *JAMA*. 2009;302(19):2095–103.
- Jørgensen CR, Freund C, Boye R, Jordet H, Andersen D, Kjølbye M. Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. *Acta Psychiatr Scand*. 2013;127(4):305–17. <https://doi.org/10.1111/j.1600-0447.2012.01923.x>.
- Simonsen S, Heinskou T, Sørensen P, Folke S, Lau M. Personality disorders: patient characteristics and level of outpatient treatment service. *Nordic J Psychiatry*. 2017;71(5):325–31.
- Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355–64. <https://doi.org/10.1176/appi.ajp.2009.09040539>.
- StataCorp. Stata Statistical Software: release 16. College Station, TX: StataCorp LLC; 2019.
- Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Stat Med*. 2011;31:328–40.
- Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ*. 2012; 345(sep14 1):e5840. <https://doi.org/10.1136/bmj.e5840>.
- Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. *BMC Med Res Methodol*. 2003;3(1):26. <https://doi.org/10.1186/1471-2288-3-26>.
- Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*. 2010; 8(1):18. <https://doi.org/10.1186/1741-7015-8-18>.
- Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol*. 2014;14(34):120–133.
- Zhang J, Quan H, Ng J, Stepanavage ME. Some statistical methods for multiple endpoints in clinical trials. *Control Clin Trials*. 1997;18(3):204–21. [https://doi.org/10.1016/S0197-2456\(96\)00129-8](https://doi.org/10.1016/S0197-2456(96)00129-8).
- Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567(7748):305–7. <https://doi.org/10.1038/d41586-019-00857-9>.
- Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. *Clin Trials J Soc Clin Trials*. 2005;2(4):282–90. <https://doi.org/10.1191/1740774505cn0980a>.
- Demets DL, Lan KG. Interim analysis: the alpha spending function approach. *Stat Med*. 1994;13(13-14):1341–52. <https://doi.org/10.1002/sim.4780131308>.
- Guyatt GH, Briel M, Glasziou P, Bassler D, Montori VM. Problems of stopping trials early. *BMJ*. 2012;344(jun15 1):e3863. <https://doi.org/10.1136/bmj.e3863>.
- Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162. <https://doi.org/10.1186/s12874-017-0442-1>.
- Nielsen EE, Nørskov AK, Lange T, Thabane L, Wetterslev J, Beyersmann J, et al. Assessing assumptions for statistical analyses in randomised clinical trials. *BMJ Evid Based Med*. 2019;24:185–9.
- Nørskov AK, Lange T, Nielsen EE, Gluud C, Winkel P, Beyersmann J, et al. Assessment of assumptions of statistical analysis methods in randomised clinical trials: the what and how. *BMJ Evid Based Med*. 2020;26(3):121–6.
- Frydenberg SJM. An introduction to Stata for health researchers, Fourth Edition: Stata Press; 2014.
- Oliveira AG. Biostatistics decoded: Wiley; 2013.
- Juil S, Poulsen S, Lunn S, Sørensen P, Jakobsen JC, Simonsen S. Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Sys Rev*. 2019;8(1):169. <https://doi.org/10.1186/s13643-019-1099-0>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Table 1

Baseline characteristics of the trial population

Characteristic	Short-term MBT (n=83)	Long-term MBT (n=83)
Demographic characteristics		
Age, mean (SD), years	XX (X)	XX (X)
Female sex – no. (%)	XX (X)	XX (X)
Civil status – no. (%)		
Married	X (X)	X (X)
Living together	X (X)	X (X)
Single	X (X)	X (X)
Living situation – no. (%)		
Alone	X (X)	X (X)
With others	X (X)	X (X)
Education level after high school – no. (%)		
Currently studying	X (X)	X (X)
Short education	X (X)	X (X)
Medium education	X (X)	X (X)
Long education	X (X)	X (X)
No higher education	X (X)	X (X)
Employment status – no. (%)		
On social welfare	X (X)	X (X)
Part time employed	X (X)	X (X)
Employed	X (X)	X (X)
Unemployed/homemaker	X (X)	X (X)
Clinical characteristics		
Psychiatric comorbidity (MINI) – no. (%)		
Anxiety disorders	X (X)	X (X)
Major depressive disorder	X (X)	X (X)
Post-traumatic stress disorder	X (X)	X (X)
Proportion of participants with subthreshold BPD – no. (%)	X (X)	X (X)
No. of BPD criteria (SCID-5-PD), mean (SD)	X (X)	X (X)
Personality disorder comorbidity (SCID-5-PD) – no. (%)		
Avoidant personality disorder	X (X)	X (X)
Obsessive-compulsive personality disorder	X (X)	X (X)
Dependent personality disorder	X (X)	X (X)
Paranoid personality disorder	X (X)	X (X)
Proportion of participants with one or more suicide-attempts the past 8 months – no. (%)	X (X)	X (X)
Proportion of participants with one or more acts of severe self-harm the past 8 months – no. (%)	X (X)	X (X)
Proportion of participants on antidepressants	X (X)	X (X)
Proportion of participants on antipsychotics		X (X)
Proportion of participants on other psychoactive drugs	X (X)	X (X)
Mean (SD) number of days from randomization to assessment time-point		
8 months post-randomization	X (X)	X (X)
16 months post-randomization	X (X)	X (X)
24 months post-randomization	X (X)	X (X)

BPD; Borderline Personality Disorder; MINI: Mini International Neuropsychiatric Interview; No.: Number; SCID-5-PD: Structured Clinical Interview for DSM-5 Personality Disorders; SD: Standard Deviation

Table 2
Primary and secondary outcome results (ITT Population)

Outcome	Short-term MBT group (n= 83)		Long-term MBT group (n= 83)		Estimate (95% CI)	p-value
	No. analyzed	Result	No. analyzed	Result		
Primary						
Borderline symptoms, ZAN-BPD, mean (SD)						
- Baseline	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 8 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 16 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
Secondary						
Functional impairment, WSAS, mean (SD)						
- Baseline	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 8 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 16 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
Quality of life, SF-36 – mental component, mean (SD)						

- Baseline	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 8 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 16 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
Global functioning, GAF, mean (SD)						
- Baseline	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 8 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
- 16 months	X	XX (X)	X	XX (X)	Mean difference (XX to YY)	$p = ZZ$
Proportion of participants with severe self-harm the past 8 months						
- Baseline	X	XX/YY (ZZ%)	x	XX/YY (ZZ%)	Relative risk XX (YY to ZZ)	$p = ZZ$
- 8 months	X	XX/YY (ZZ%)	x	XX/YY (ZZ%)	Relative risk XX (YY to ZZ)	$p = ZZ$
- 16 months	X	XX/YY (ZZ%)	x	XX/YY (ZZ%)	Relative risk XX (YY to ZZ)	$p = ZZ$

GAF: Global Assessment of Functioning; ITT: Intention To Treat; SF-36: Short Form Health Survey – 36; SD: Standard Deviation; WSAS: Work and Social Adjustment Scale; ZAN-BPD: Zanarini Rating Scale for Borderline Personality Disorder

Paper 5



The Capacity to End: Termination of Mentalization-Based Therapy for Borderline Personality Disorder

Sophie Juul^{1,2} · Sebastian Simonsen¹ · Anthony Bateman³

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Terminating a therapeutic relationship can be a challenging phase with patients suffering from borderline personality disorder. Despite the critical importance of the termination phase, the proportion of psychotherapy literature devoted to the demands and challenges of this phase is small. This paper describes a mentalization-based approach to detect and intervene against such challenges. It is proposed that termination challenges, while operating through overlapping and interactive mechanisms, can be attributed to (1) patient factors, (2) therapist factors, and (3) therapeutic relationship factors. The paper has clinical implications and suggests that the aim of enhancing mentalizing capacities should include mentalizing the often complicated and mixed feelings associated with separation and loss of the therapeutic relationship. To facilitate this process, we propose the use of a “termination formulation”, in which patients’ outcomes and future goals are recapitulated in the termination phase of mentalization-based therapy.

Keywords Borderline personality disorder · Mentalization-based therapy · Treatment termination

Introduction

Borderline personality disorder (BPD) is a prevalent psychiatric condition characterized by a pervasive pattern of symptoms such as intense and unstable relationships, chronic feelings of emptiness, intense anger, fear of abandonment, intolerance for aloneness, and a lack of a stable sense of self (American Psychiatric Association 2013). These fundamental aspects of BPD, particularly the relational aspects, can be understood as stemming from impairments in their underlying attachment organization (Levy et al. 2005).

Attachment theory, developed by Bowlby (1972, 1973, 1998), proposed that internal representations of self and others provide prototypes for all later relationships. Such representations exist outside of awareness and are quite resistant

to change (Waters et al. 2000). Bowlby’s suggestion that early experience with the caregiver organizes later attachment relationships, has been used to explain the development of BPD psychopathology (Fonagy et al. 2000). Patients with BPD typically experience considerable distortions of attachment representations, resulting in a disrupted capacity to depict mental states in themselves and others (Levy et al. 2005). Further, feelings of abandonment anxiety can be easily triggered in attachment situations, which are likely to be transferred to the therapeutic relationship.

Consequently, despite the obvious need for effective treatments, patients with BPD can be hard to engage effectively in treatment as they bring insecure pre-occupied, avoidant, or disorganized attachment strategies into their interaction with services and treatment (Choi-Kain et al. 2009). A common concern for clinicians when treating patients with BPD is premature treatment termination (De Panfilis et al. 2012; Wnuk et al. 2013). In a systematic review of therapy non-completion in patients with personality disorders, a median non-completion rate was estimated at 37% (McMurran et al. 2010). In contrast, some BPD patients are also likely to be frequent and continuous users of mental health services (Soeteman et al. 2008), often expressing a need for more help, presenting the treating clinicians with a difficult judgement of whether to terminate or to extend

✉ Sophie Juul
sophie.juul@regionh.dk

¹ Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Stolpegårdsvej 20, 2820 Gentofte, Denmark

² Department of Psychology, University of Copenhagen, Copenhagen, Denmark

³ Anna Freud Centre, Kantor Centre of Excellence, London, Great Britain

the treatment. There is reason to believe that such termination challenges characterized by high rates of premature treatment termination, prolonged treatments, and continuous referrals to specialist treatment, can to some extent be theoretically understood within an attachment framework, e.g. as a way of coping with anticipated and foreseen separation or loss. Thus, implicitly or explicitly, enhanced mentalizing of themes associated with separation and loss in the treatment of BPD could help prevent reenactment of attachment anxiety and it can help the patient form a new and more stable representation of interpersonal relationships.

Mentalization-based therapy is not immune to termination challenges activated by attachment anxiety. However, it may offer ways to understand and deal with them. Mentalization-based therapy is a psychodynamic therapy based on attachment theory that has originally been developed for treating BPD (Bateman and Fonagy 2016). Mentalization refers to the capacity to understand one's own and others internal mental states such as feelings, wishes, desires, and beliefs. Frequent and extensive failures to this capacity, especially in the context of attachment relationships, is believed to play an important role in the psychopathology of BPD (Fonagy and Luyten 2009). Mentalization-based therapy currently has empirical support as an 18-months program for BPD. This treatment program has been implemented in clinical settings worldwide, both in day-hospital (Bateman and Fonagy 1999; Smits et al. 2019) and outpatient (Bateman and Fonagy 2009; Kvarstein et al. 2015) settings, and is also being tested in a 20-weeks short-term version (Juul et al. 2019).

There are three main phases to the trajectory of mentalization-based therapy: a beginning-, a middle-, and an ending-phase. Each phase has a distinct aim and harnesses specific therapeutic processes. According to the treatment manual, the ending phase should focus on mentalizing affective states associated with separation, and there should be a focus on how to maintain gains that have been made during treatment (Bateman and Fonagy 2016). However, there will be many variations to the clinical application of the ending phase between clinicians, with different types of patients, and in different time format (e.g. short-term or long-term, predefined or open-ended).

In this paper, we will elaborate on the mentalization-based framework for understanding termination challenges, and we will propose possible clinical solutions to these challenges. While termination challenges are not restricted to mentalization-based therapy, we believe that enhanced mentalizing of feelings associated with termination can be helpful in any form of psychotherapy for BPD.

Termination Challenges

For the ending of mentalization-based therapy, and other psychotherapies for BPD, to be effective in stimulating long term improvement, it is important to outline challenges that can threaten this effectiveness. Knowing that these challenges operate in overlapping and interactive patterns, we have divided them into the following three distinct domains: *patient factors*, *therapist factors*, and *therapeutic relationship factors*.

Patient Factors

According to the mentalization theory, all people, but particularly patients with BPD, present with disrupted mentalizing, or nonmentalizing modes. There are three typical nonmentalizing modes termed *psychic equivalence mode*, *teleological mode*, and *pretend mode* (Bateman and Fonagy 2016; Fonagy and Luyten 2009). These are all likely to influence the termination phase of treatment.

Psychic equivalence mode involves thoughts and feelings becoming “too real” to a point where it is extremely difficult for the patient to entertain possible alternative perspectives. There is often a suspension of doubt forcing the individual to believe that his or her own perspective is the only one possible (Bateman and Fonagy 2019). In the termination phase of treatment, a patient presenting with psychic equivalence mode will become hyperaroused and overwhelmed by feelings of abandonment. Transference reactions can be dominated by feelings of rejection, dependence or splitting, and the patient can perceive the therapist as negligent or inattentive to his or her needs. We can refer to this as “the emotions form reality”-problem. This can result in both aggressive or clinging behaviors, separately or simultaneously. The patient might manifest increased dependency on the therapist, reflected in attempts to postpone termination (e.g. by presenting new symptoms or crises). For the patient, postponing therapy may serve to protect against anxious or angry feelings related to the experience of rejection, abandonment, and separation (Joyce et al. 2007).

Teleological mode involves states of mind being recognized only if their outcomes are physically observable (Bateman and Fonagy 2019). For example, affection is only perceived to be true, if it is accompanied by physical contact. In the termination phase of treatment, a patient presenting with a teleological mode of functioning will believe, that he or she can only experience real improvement in therapy, if the therapist is present and continuously available for support. We can refer to it as “the no therapist, no help”-problem. The presence of the therapist

is associated with comfort, while absence of the therapist is associated with distress. From an attachment point of view, the patient will present with an inability to form a stable psychic representation of the therapist, once the therapist is not physically present—a phenomenon similar to lack of *object constancy*, as formulated in psychoanalytic theory (Fraiberg 1969; Hartmann 1952).

Pretend mode involves thoughts and feelings becoming too disconnected from reality. In more extreme cases, this may lead to feelings of derealization and dissociation (Bateman and Fonagy 2019). A patient dominated by pretend mode may be prone to hypermentalizing—a state in which he or she may speak about mental states but with very little or no connection to real affective experiences. In the termination phase of treatment, a patient dominated by pretend mode may prematurely detach or perhaps become overly compliant arguing that the therapy has been a success, even though mentalizing or functional capacities indicate otherwise. We may refer to this as “the as-if”-problem. The patient may show considerable cognitive understanding of why the therapy must come to an end, but there may be little understanding of the underlying affective experiences induced by the loss of the therapist. Their restricted engagement may protect them from their feared intense or overwhelming response to termination. This phenomenon may be similar to a dissociative experience and may be further explained by notion of *the false self* as proposed by Winnicott (1960).

Further, it is likely that the concept of *epistemic trust* may influence the termination phase as well. Epistemic trust refers to an individual’s willingness to consider new knowledge from another person as trustworthy, generalizable, and relevant to the self (Fonagy and Allison 2014). It is suggested that many forms of psychopathology may be underpinned by an inability to benefit from social communication due to *epistemic mistrust*. This results in these patients’ reluctance to modify their beliefs and expectations, even when facing social experiences that clearly indicate the value of doing so (Fonagy et al. 2019). In the termination phase of treatment, a patient characterized by epistemic mistrust or increased epistemic vigilance may not trust the therapist’s reinforcements of autonomy or validations of good mentalizing. Instead, the patient may experience these interventions as untrue or even as a sign of masked abandonment. In other words, patients may not trust their own ability to generate a representation of the therapist representing them, if the therapist is not physically present.

Therapist Factors

Therapists’ own conflicts associated with the ending of therapy and the therapeutic relationship may also play an important role in the termination phase of treatment. Different

therapist emotional responses (or countertransference) are likely to become activated when facing treatment termination. The relationship between BPD pathology and therapist countertransference has been studied both theoretically and empirically in the literature on personality pathology (Betan et al. 2005; Colli et al. 2014). Typical countertransference reactions include feelings of helplessness/inadequacy, overinvolvement, overprotection, or feelings of being overwhelmed and disorganized.

In the termination phase, a patient may react in a *psychic equivalent mode* of functioning e.g. by becoming overwhelmed by feelings of abandonment, and this may in turn evoke feelings of guilt, helplessness, overprotection or disorganization in the therapist. Thus, termination challenges arise, if the therapist acts on their own feelings and thus fails to intervene against the psychic equivalent mode. For example, rather than regulating arousal and facilitating independence or self-efficacy, the therapist reinforces the psychic equivalence by explicitly agreeing that more therapy is needed or validates the notion of abandonment. In this case, there is a risk that the intolerable feelings are worsened. Instead, a contrary move to a cognitive focus and diversion would be more appropriate to facilitate reflection on personal agency and life after treatment (Bateman and Fonagy 2016). The therapist could say: “I hear you saying that you feel anxious and sad about ending therapy. But you have also told me that you wish to be more independent. Can we entertain the last part for a little while? How would independence look like for you? In which situations would you know that you have become more independent?”.

If the patient also presents with a *teleological mode*, arguing that he or she can only recover in the presence of a therapist, the therapist may feel overprotective and (teleologically) act on this feeling by deciding to prolong the therapy or to refer the patient to another therapist or treatment modality. In these cases, there is a risk that the challenging ending merely becomes postponed instead of mentalized. We are not arguing, that transferring to another therapist or treatment modality is never appropriate. Patients who resist termination through self-destructive acting out around the time of the termination phase may be offered additional care, but the therapist must reflect on his or her own ability to stimulate a mentalizing understanding of this behavior with the patient. In such cases therapists should have a supervision forum, which can be used to process feelings of e.g. guilt or remorse, if they must terminate the therapy with a patient who is still suffering emotionally and behaviorally. The same applies to forced terminations that can occur in clinical practice, for example when a training therapist must move to a new treatment setting or an experienced therapist must relocate for personal or professional reasons.

If the patient was dominated by a *pretend mode* of functioning throughout the therapy, and this was not properly

addressed, it can cause challenges to the termination phase of treatment. Pretend mode that is unaddressed results in endless therapy that merely “looks like” therapy. Therapists may then feel responsible if symptom severity is unchanged at the end of treatment and may thus decide to prolong the treatment because of continued treatment-demanding psychopathology (Sharp et al. 2013). With a patient dominated by pretend mode, another possibility is that the therapist avoids shifting from cognitive to affective mentalizing of feelings associated with ending. This lack of contrary move could cause both the patient and therapist to prematurely detach and agree that the therapy was a success and must come to an end. In this case, the therapist should stop and focus on the affective experience here-and-now, for example by saying: “You tell me that ending therapy is fine with you. How do you feel right now telling me that?”. If still stuck in pretend mode, the patient may reply: “I don’t really know what I am feeling”, and this may be a cue that the affective experience of termination should be further explored.

Therapists often experience patients with severe personality disorders as “hard to reach” as evidenced by premature termination and poor outcomes (Fonagy et al. 2017a, b). However, somewhat paradoxically, such patients may also be the hardest to let go. In such cases therapists may themselves lack *epistemic trust* in their patients’ capacity to end. Therapists may experience countertransference reactions that interfere with their ability to recognize the patient’s strengths and resources (Joyce et al. 2007). In other words, termination challenges may arise when therapists are not sufficiently able to generate a mental representation of the patient which properly mirrors the patients’ own sense of agency. In such cases the therapist often holds a fixed and rigid representation of the patient as someone who cannot manage without their support. The therapists may then ask themselves: “Am I able to generate a clear sense of my patient’s clinical change?”, and “Is my patient allowed to change in my mind?”. With a lack of epistemic trust in the patients’ ability to reach a certain level of autonomy and believe the patient’s own reports, the therapist may run the risk of delivering therapy that maintains the patient in the belief that more therapy will always be needed and available.

Therapeutic Relationship Factors

Over the course of treatment, an emotional bond is usually developed between the patient and the therapist. Both parties may experience sadness and reservations about termination and may even be interested in continuing the relationship. In fact, the goal of the therapeutic relationship in mentalization-based therapy is to facilitate such a safe environment, in which both personal and interpersonal thoughts and feelings can be mentalized (Bateman and Fonagy 2016). However, willingness to manage sensitive topics and ask demanding

questions within the therapeutic relationship also plays an important role in the competently delivered mentalization-based therapy session. Conversely, low rated sessions are characterized by therapists who are not adequately confrontative, who avoid difficult content, and who may exaggerate efforts to accommodate or please the patient (Folmo et al. 2019; Möller et al. 2016). We speculate that when both the patient and the therapist fail to engage in a trustful yet confronting therapeutic environment, it may lead to termination challenges. We believe that a therapeutic focus on the relational aspect in session may increase mentalizing of the patient’s interpersonal patterns. On the contrary, therapy that is too supportive, in which both the patient and therapist avoid mentalizing difficult interpersonal themes, might suffer the problem that challenging interpersonal patterns will reveal themselves during the termination process. By then, it may be too late to properly address them.

Further, in the beginning phase of mentalization-based therapy, and in other treatments for BPD, a case formulation is developed in a collaboration between the therapist and the patient. The mentalization-based case formulation is a collaborative clinical agreement, designed to foster agency and a sense of control, which in turn helps foster alliance and relaxes epistemic mistrust (Bateman et al. 2019; Karterud and Kongerslev 2019). We speculate that an mentalization-based case formulation with treatment goals that are not appropriate to the available length of treatment may increase the likelihood of a challenging termination. Treatment goals can be formulated either too vaguely (i.e. to improve mentalization) or too narrowly (i.e. stop self-harming after group therapy), both of which may result in termination challenges when gains and processes are reviewed and recapitulated in the ending phase of therapy.

Overcoming Termination Challenges Based on Mentalization-Based Therapy

In the following, we will offer some possible remedies for the challenges that arise when ending mentalization-based therapy and discuss their relevance to psychotherapy with BPD-patients in general. We focus on three main areas: mentalizing loss of the therapeutic relationship, management of treatment structure, and finally we propose to extend the case formulation with a termination formulation.

Mentalizing Loss of the Therapeutic Relationship

Some patients become overemotional, over-attached, or attribute incorrect attitudes to the therapist as a way of handling separation. Some patients may become confused, some become angry or devaluing, and others may continue to “deflate” the relationship. If not mentalized within the

context of the therapeutic relationship, these problematic strategies may influence what the patients carry forward into new relationships, e.g. “the good never lasts”, “all people eventually let you down”, “nobody will ever be able to understand me”. Thus, ending badly increases epistemic mistrust and increases the possibility that patients become inflexible or rigid in the face of social change (Fonagy et al. 2019).

According to a review of the literature by Joyce et al. (2007), two perspectives have dominated the literature on termination of the therapeutic relationship; the “termination-as-loss” metaphor and the notion of “termination as transformation”. The two perspectives are not mutually exclusive but will often occur simultaneously. In mentalization-based therapy, and other psychodynamic therapies for personality disorders, for example transference-focused therapy (Kernberg et al. 2008), the therapist will carefully assess the patient’s implicit or explicit metaphor and its consequences for the ending phase. The therapist will validate and normalize the patient’s emotional experiences of separation and loss as understandable based on their life narrative, but should also clearly communicate the possibility of other, more secure ways of saying goodbye i.e. acknowledging mixed feelings and recapitulating the relationship story.

Management of Treatment Structure

In order to enhance mentalizing of challenging terminations, we suggest that the finite nature of therapy should not only be specified from the beginning and addressed in the case formulation, but it should also be considered in the session structure throughout the course of therapy. In short-term mentalization-based groups this is done by therapists labeling the session number at the beginning of each session, e.g. by stating that this is the fourth out of 20 sessions. This effectively reminds patients about the boundaries of therapy and nudges them to bring up relevant material. Further, each session should leave sufficient time for the therapist and patient(s) to reflect on the treatment process in terms of mentalizing/understanding each other (Bateman et al. 2019). The group therapists should ask the group: “How did we do today? Did we understand each other? What have we managed to help with and what have we not; what do we need to think about over the week to discuss next time?”. We believe that this process both has general benefits in terms of building epistemic trust but may also have more specific benefits in preparing and building a capacity to terminate therapy. Often the therapist will start this process by asking the patient to sum up topics and aspects of the session emphasizing his or her perspective on the process and issues of importance. The therapist will then reflect on this as a way of helping the patient compare his or her perspectives with that of the therapists. In this way patients get a clearer

representation of what they themselves feel is important and helpful and this stronger sense of subjectivity (and otherness) we believe could be helpful both in terms of preparing separation and in supporting stronger agency and more secure attachment strategies. In addition, asking the patient to think about what has been left over during the week on their own builds a sense of personal agency.

Extending the Case Formulation with a Termination Formulation

The ending of therapy should be foreseen and prepared for in the case formulation at the beginning of therapy. The case formulation should include a relational passport (Karterud and Kongerslev 2019) specifying attachment patterns and foreseen reactions to the ending of treatment. The therapist should probe the patient for prior experiences with ending and loss and ask the patient for expectations about their reactions when the ending of treatment will be near. From a mentalization-based perspective we do not expect patients with BPD to be very conscious about separation reactions. Attachment and separation trauma tend to be implicitly represented in patients’ minds. However, the explicit addressing of the issue in the case formulation serves the purpose of shedding light on previously non-mentalized aspects of relationship patterns, for example when patients have been unable to read and take advantage of socially inclusive cues. In mentalization-based therapy, a key purpose is for the patient to shift from implicit to explicit mentalizing of mental states (Bateman and Fonagy 2016). The therapy should seek to modify the patients’ pathological activation of attachment processes by helping to avoid automatic processing of events. This automatic processing typically occurs when patients recruit negative memories of relationships and events to interpret new experiences (Winter et al. 2014). In the context of terminating therapy, patients may automatically expect the termination to confirm their negative memory bias. By entertaining a more explicit awareness of this bias in the case formulation, the therapist may potentially help the patient to a more deliberate and controlled way of being in a relationship and it may serve as a model for future separation.

When reviewing the case formulation, it is recommended that the patient and the therapist should reflect on strategies to maintain mentalizing capacities outside of the treatment system. This is important, given that patients with BPD often come from deprived social environments, and an increase in mentalizing following therapy may bring about psychological pain and vulnerability (Bo et al. 2017). This pain is exacerbated when patients continue to live in a family or social environment which promotes non-mentalizing responses to problems, for example using coercive or even violent control of others. Many patients might (rightfully)

see it as a best option to return to old strategies. In most cases, however, it is our experience that even within very deprived environments, it will be possible to identify one or two people with whom the patient can build a more stable and trusting relationship. It is important to recognize this early in therapy and perhaps include it in the case formulation to help support such relationships, while at the same time be transparent about the fact that the therapist will not always be available for support. Thus, the case formulation can be a particularly useful tool for enabling a smooth transition from therapy to real life.

In order to facilitate this transition, we propose to extend the case formulation with a simple form called the *termination formulation*. This form addresses three areas of termination and asks the patient to reflect on (a) what has been achieved and what needs further work; as mentioned, this process has been rehearsed at the end of sessions over time, (b) if there is anything the patient needs to address in the relationship with either the individual or group therapist or fellow group members before ending therapy, and (c) how the therapeutic progress can be retained or further developed after the therapy has ended. To foster agency and a sense of self-direction, the patient is given this form to take home and is asked to write down his or her own thoughts and feelings leaving three to four sessions to process the content and manage issues that come up. The termination formulation can be a useful tool in any form of psychotherapy tradition for BPD, as it entails pantheoretical themes of termination (Norcross et al. 2017).

Sometimes a case formulation is either not made, or it is made in ways that do not foster good outcomes (Bateman 2011; Simonsen et al. 2011). In such cases, implementing this simple termination formulation can be a particularly useful strategy to help identify main outcomes and support the maintenance of these outcomes once therapy has ended, even when these have not been clearly formulated at the beginning. If new issues or themes come up when preparing the termination formulation, or if countertransference issues may interfere with the development of the termination formulation in a problematic way, supervision becomes of the upmost importance. In mentalization-based supervision, supervisors are encouraged to help therapists clarify and extend their “roadmap” of realistic treatment trajectories and individualized outcomes for their patients. Furthermore, problems with overwhelming or confusing countertransference reactions are common when treating patients with BPD (Betan et al. 2005; Colli et al. 2014). Therefore, therapists performing mentalization-based therapy are encouraged to bring up patients, who are facing their ending phase of treatment, in supervision. We believe that mentalizing one’s own feelings towards termination is an important step towards being able to facilitate a beneficial ending phase for the patient.

Conclusion

Challenging terminations are likely to be an inherent part of all forms of personality disorder treatments regardless of treatment modality. In this paper, we have proposed an approach based on mentalization-based therapy to detect and intervene against such challenges. We have proposed that termination challenges can be attributed to (1) patient factors, (2) therapist factors, and (3) therapeutic relationship factors. We have suggested that explicit mentalizing of the often complicated and mixed feelings associated with separation and loss of the therapeutic relationship is an important part of the termination phase. To facilitate this process, we have proposed the use of a *termination formulation* in extension to the case formulation, in which patient’s outcomes, interpersonal issues in therapy, and future goals are recapitulated in the termination phase. At this point there is no empirical evidence that better mentalizing of termination issues causes better overall outcomes in the treatment of BPD. However, we are confident that providing therapists and patients with a more explicit framework for termination addresses a previously neglected issue and can enhance both therapists’ and patients’ capacity to end.

Funding Not applicable.

Data Availability Data sharing is not applicable to this article as no new data were created or analysed in this study.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington, DC: American Psychiatric Publishing.
- Bateman, A. (2011). Commentary on ‘Minding the difficult patient’: Mentalizing and the use of formulation in patients with borderline personality disorder comorbid with antisocial personality disorder. *Personality and Mental Health*, 5(1), 85–90. <https://doi.org/10.1002/pmh.156>.
- Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *American Journal of Psychiatry*, 156(10), 1563–1569. <https://doi.org/10.1176/ajp.156.10.1563>.
- Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry*, 166(12), 1355–1364. <https://doi.org/10.1176/appi.ajp.2009.09040539>.

- Bateman, A., & Fonagy, P. (2016). *Mentalization-based treatment for personality disorders: A practical guide*. Oxford, United Kingdom: Oxford University Press.
- Bateman, A., & Fonagy, P. (2019). *Handbook of mentalizing in mental health practice* (2nd ed.). Washington, DC: American Psychiatric Association Publishing.
- Bateman, A., Kongerslev, M., & Bo Hansen, S. (2019). Mentalization-based treatment group therapy (MBT-G) for adults and adolescents. In A. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practices*. Washington, DC: American Psychiatric Association Publishing.
- Betan, E., Heim, A. K., Conklin, C. Z., & Westen, D. (2005). Countertransference phenomena personality pathology in clinical practice: An empirical investigation. *American Journal of Psychiatry*, 162, 890–898.
- Bo, S., Sharp, C., Fonagy, P., & Kongerslev, M. (2017). Hypermentalizing, attachment, and epistemic trust in adolescent BPD: Clinical illustrations. *Personality Disorders*, 8(2), 172–182. <https://doi.org/10.1037/per0000161>.
- Bowlby, J. (1972). *Attachment and loss: Volume I: Attachment*. London: Penguin Books.
- Bowlby, J. (1973). *Attachment and loss: Volume II: Separation, anxiety and anger* (pp. 1–429). London: The Hogarth Press and the Institute of Psycho-Analysis.
- Bowlby, J. (1998). *Attachment and loss: Volume III: Loss—sadness and depression*. London: Pimlico.
- Choi-Kain, L. W., Fitzmaurice, G. M., Zanarini, M. C., Laverdière, O., & Gunderson, J. G. (2009). The relationship between self-reported attachment styles, interpersonal dysfunction, and borderline personality disorder. *The Journal of Nervous and Mental Disease*, 197(11), 816–821.
- Colli, A., Annalisa Tanzilli, A., Dimaggio, G., & Lingardi, V. (2014). Patient personality and therapist response: An empirical investigation. *American Journal of Psychiatry*, 171, 102–108.
- De Panfilis, C., Marchesi, C., Cabrino, C., Monici, A., Politi, V., Rossi, M., et al. (2012). Patient factors predicting early drop-out from psychiatric outpatient care for borderline personality disorder. *Psychiatry Research*, 200(2–3), 422–429.
- Folmo, E. J., Karterud, S. W., Kongerslev, M. T., Kvarstein, E. H., & Stänicke, E. (2019). Battles of the comfort zone: Modelling therapeutic strategy, alliance, and epistemic trust—A qualitative study of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy*. <https://doi.org/10.1007/s10879-018-09414-3>.
- Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy*, 51(3), 372.
- Fonagy, P., Allison, E., & Campbell, C. (2019). Mentalizing, resilience, and epistemic trust. In A. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (2nd ed.). Washington, DC: American Psychiatric Association Publishing.
- Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology*, 21(4), 1355–1381. <https://doi.org/10.1017/S0954579409990198>.
- Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2017a). What we have changed our minds about: Part 1. Borderline personality disorder as a limitation of resilience. *Borderline Personality Disorder and Emotion Dysregulation*, 4, 11. <https://doi.org/10.1186/s40479-017-0061-9>.
- Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2017b). What we have changed our minds about: Part 2. Borderline personality disorder, epistemic trust and the developmental significance of social communication. *Borderline Personality Disorder and Emotion Dysregulation*, 4, 9. <https://doi.org/10.1186/s40479-017-0062-8>.
- Fonagy, P., Target, M., & Gergely, G. (2000). Attachment and borderline personality disorder: A theory and some evidence. *Psychiatric Clinics*, 23(1), 103–122.
- Fraiberg, S. (1969). Libidinal object constancy and mental representation. *The Psychoanalytic Study of the Child*, 24(1), 9–47.
- Hartmann, H. (1952). The mutual influences in the development of ego and id. *The Psychoanalytic Study of the Child*, 7(1), 9–30.
- Joyce, A. S., Piper, W. E., Ogrodniczuk, J. S., & Klien, R. H. (2007). *Termination in psychotherapy: A psychodynamic model of processes and outcomes*. Washington, DC: American Psychological Association.
- Juul, S., Lunn, S., Poulsen, S., Sorensen, P., Salimi, M., Jakobsen, J. C., et al. (2019). Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: A protocol for a randomized clinical trial. *Trials*, 20(1), 196. <https://doi.org/10.1186/s13063-019-3306-7>.
- Karterud, S., & Kongerslev, M. T. (2019). Case formulations in mentalization-based treatment (MBT) for patients with borderline personality disorder. In U. Kramer (Ed.), *Case formulation for personality disorders* (pp. 41–60). London: Elsevier.
- Kernberg, O. F., Yeomans, F. E., Clarkin, J. F., & Levy, K. N. (2008). Transference focused psychotherapy: Overview and update. *The International Journal of Psychoanalysis*, 89(3), 601–620.
- Kvarstein, E. H., Pedersen, G., Urnes, O., Hummelen, B., Wilberg, T., & Karterud, S. (2015). Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—does it make a difference? *Psychology and Psychotherapy*, 88(1), 71–86. <https://doi.org/10.1111/papt.12036>.
- Levy, K. N., Meehan, K. B., Weber, M., Reynoso, J., & Clarkin, J. F. (2005). Attachment and borderline personality disorder: Implications for psychotherapy. *Psychopathology*, 38(2), 64–74.
- McMurrin, M., Huband, N., & Overton, E. (2010). Non-completion of personality disorder treatments: A systematic review of correlates, consequences, and interventions. *Clinical Psychology Review*, 30(3), 277–287. <https://doi.org/10.1016/j.cpr.2009.12.002>.
- Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2016). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research*. <https://doi.org/10.1080/10503307.2016.1158433>.
- Norcross, J. C., Zimmerman, B. E., Greenberg, R. P., & Swift, J. K. (2017). Do all therapists do that when saying goodbye? A study of commonalities in termination behaviors. *Psychotherapy*, 54(1), 66.
- Sharp, C., Ha, C., Carbone, C., Kim, S., Perry, K., Williams, L., et al. (2013). Hypermentalizing in adolescent inpatients: Treatment effects and association with borderline traits. *Journal of Personality Disorders*, 27(1), 3–18.
- Simonsen, S., Nørgaard, N. L., Larsen, K. A., & Bjørnholm, K. I. (2011). Mentalization-inspired case formulation: Minding the difficult patient. *Personality and Mental Health*, 5(1), 73–79. <https://doi.org/10.1002/pmh.154>.
- Smits, M. L., Feenstra, D. J., Eeren, H. V., Bales, D. L., Laurensen, E. M., Blankers, M., ... Verheul, R. (2019). Day hospital versus intensive out-patient mentalisation-based treatment for borderline personality disorder: Multicentre randomised clinical trial. *The British Journal of Psychiatry*, 1–6.
- Soeteman, D. I., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. V. (2008). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry*, 69, 259–265.
- Waters, E., Merrick, S., Treboux, D., Crowell, J., & Albersheim, L. (2000). Attachment security in infancy and early adulthood: A twenty-year longitudinal study. *Child Development*, 71(3), 684–689.

- Winnicott, D. W. (1960). Ego distortion in terms of true and false self. *The maturational process and the facilitating environment: Studies in the theory of emotional development* (pp. 140–157). New York: International Universities Press.
- Winter, D., Elzinga, B., & Schmahl, C. (2014). Emotions and memory in borderline personality disorder. *Psychopathology*, 47(2), 71–85.
- Wnuk, S., McMain, S., Links, P. S., Habinski, L., Murray, J., & Guimond, T. (2013). Factors related to dropout from treatment in two outpatient treatments for borderline personality disorder. *Journal of Personality Disorders*, 27(6), 716–726.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper 6

Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder: A Randomized Clinical Trial (MBT-RCT)

Sophie Juul^{a,b} Janus Christian Jakobsen^{a,c} Emilie Hestbaek^{b,d}
Caroline Kamp Jørgensen^{a,c} Markus Harboe Olsen^{a,e} Marie Rishede^f
Frederik Weischer Frandsen^{b,f} Sune Bo^d Susanne Lunn^d Stig Poulsen^d
Per Sørensen^b Anthony Bateman^g Sebastian Simonsen^{b,d}

^aCopenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ^bStolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Copenhagen, Denmark; ^cDepartment of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; ^dDepartment of Psychology, University of Copenhagen, Copenhagen, Denmark; ^eDepartment of Neuroanaesthesiology, The Neuroscience Centre, The Neuroscience Centre, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ^fDepartment of Psychiatry Mental Health Services East, Region Zealand, Roskilde, Denmark; ^gDepartment of Clinical, Educational and Health Psychology, University College London, London, UK

Keywords

Mentalization-based therapy · Borderline personality disorder · Short-term psychotherapy · Randomized clinical trial

Abstract

Introduction: Borderline personality disorder (BPD) is a severe and prevalent psychiatric disorder. Mentalization-based therapy (MBT) is an evidence-based intervention for BPD, and several countries offer treatment programs for BPD lasting for years, which is resource demanding. No previous trial has compared short-term with long-term MBT. **Objective:** The aim of the study was to assess the efficacy and safety of short-term versus long-term MBT for outpatients with BPD. **Methods:** Adult outpatients (≥ 18 years) with subthreshold or diagnosed BPD were randomly assigned (1:1) to short-term MBT (5 months) or long-term MBT (14 months). The primary

outcome was BPD symptoms assessed with the Zanarini Rating Scale for Borderline Personality Disorder. Secondary outcomes were functional impairment, quality of life, global functioning, and severe self-harm. All outcomes were primarily assessed at 16 months after randomization. This trial was prospectively registered at ClinicalTrials.gov, NCT03677037. **Results:** Between October 4, 2018, and December 3, 2020, we randomly assigned 166 participants to short-term MBT ($n = 84$) or long-term MBT ($n = 82$). Regression analyses showed no evidence of a difference when assessing BPD symptoms (MD 0.99; 95% CI: -1.06 to 3.03 ; $p = 0.341$), level of functioning (MD 1.44; 95% CI: -1.43 to 4.32 ; $p = 0.321$), quality of life (MD -0.91 ; 95% CI: -4.62 to 2.79 ; $p = 0.626$), global functioning (MD -2.25 ; 95% CI: -6.70 to 2.20 ; $p = 0.318$), or severe self-harm (RR 1.37; 95% CI: 0.70 – 2.84 ; $p = 0.335$). More participants in the long-term MBT group had a serious adverse event compared with short-term MBT (RR 1.63; 95% CI: 0.94 – 3.07 ; $p = 0.088$), primarily driven by

a difference in psychiatric hospitalizations (RR 2.03; 95% CI: 0.99–5.09; $p = 0.056$). **Conclusion:** Long-term MBT did not lead to lower levels of BPD symptoms, nor did it influence any of the secondary outcomes compared with short-term MBT.

© 2023 The Author(s).
Published by S. Karger AG, Basel

Introduction

Borderline personality disorder (BPD) is a mental disorder characterized by a pervasive pattern of symptoms, such as affective dysregulation, identity diffusion, interpersonal problems, self-harming behavior, and suicide-related mortality [1]. According to epidemiological studies, 1.6% of the general population suffers from BPD [2]. In clinical populations, it is the most common personality disorder [2], with a prevalence of between 9 and 22% of all psychiatric outpatients [3]. BPD causes a significant socioeconomic burden, not only because of direct health care costs but also because of indirect costs like social and occupational dysfunction [4]. These findings emphasize the need for efficacious and cost-effective treatments for this highly prevalent disorder. Currently, psychological interventions are considered the primary treatment of choice for BPD [5].

Traditionally, long-term comprehensive psychotherapy programs have been considered necessary for patients with BPD [6]. These include, but are not limited to, long-term psychodynamic therapy (for complex mental health disorders including BPD) lasting minimum a year [7], dialectical behavior therapy (DBT) up to 1 year, transference-focused psychotherapy (TFP) up to 3 years, schema-focused therapy (SFT) up to 3 years, and mentalization-based therapy (MBT), which was originally manualized and tested for 18 months but is sometimes delivered in longer durations in clinical practice, e.g., up to 3 years of outpatient treatment [8].

MBT is a psychological intervention rooted in psychodynamic and attachment theory, developed specifically to treat BPD [9]. MBT was originally designed as a long-term program, with 18 months of psychotherapy consisting of four components: (1) psychoeducation, (2) case formulation, (3) group therapy, and (4) individual therapy. Mentalization refers to the ability to understand one's own and others' mental states. The theoretical assumption is that BPD patients are more vulnerable to lose the capacity to mentalize when experiencing emotional distress. Increasing mentalization skills is assumed to minimize BPD symptoms such as emotional dysregulation, impulsivity, and suicidal ideation.

A systematic review with meta-analysis of all psychological therapies for BPD was published in 2020 [8]. The authors concluded that long-term MBT was more effective than treatment as usual in reducing self-harm, suicidality, and depression [8]. The original 18-month MBT program is prescribed by the MBT manual [9], but it is unclear if this is the most optimal treatment duration. We recently published a systematic review with meta-analysis assessing the beneficial and harmful effects of shorter-term versus longer-term psychotherapy for all adult mental health disorders [10], and we found no previous trial comparing the effects of different durations of MBT. The aim of the present trial was to assess the beneficial and harmful effects of short-term versus long-term MBT for outpatients with subthreshold or diagnosed BPD.

Materials and Methods

This trial is reported in accordance with the Consolidated Standards of Reporting Trials (CONSORT) 2010 statement [11] (online suppl. material pp. 1–2; for all online suppl. material, see <https://doi.org/10.1159/000534289>) and the methodological recommendations for trials of psychological interventions by Guidi et al. [12]. The trial was prospectively registered at ClinicalTrials.gov (NCT03677037, registered on September 19, 2018) before the first participant was randomized. The methodology was predefined and described in detail previously [13, 14] (online suppl. material, pp. 3–20).

Study Design and Participants

In this investigator-initiated, parallel group, single-centered, randomized clinical superiority trial, we consecutively screened patients from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre in the Mental Health Services of the Capital Region of Denmark. Eligible participants were adults (≥ 18 years) with subthreshold or diagnosed BPD according to the diagnostic criteria provided in the Diagnostic and Statistical Manual for Mental Disorders, 5th edition (DSM-5) [1] who provided written informed consent. Exclusion criteria were reduced intellectual functioning (IQ < 75), a full diagnosis of schizotypal personality disorder or antisocial personality disorder according to DSM-5, a comorbid mental health disorder requiring specialist treatment, current (past 2 months) substance dependence assessed with the Mini International Neuropsychiatric Interview (MINI) [15], concurrent psychotherapeutic treatment, inability to understand Danish, and lack of informed consent.

The trial was overseen by a steering committee and by an independent data and safety monitoring committee who reviewed the data and performed one predefined, blinded interim analysis after half of the target sample size had been randomized. The trial was approved by the Danish Data Protection Agency (I-suite: 6553; August 30, 2018) and the Regional Ethics Committee of Health Research of the Capital Region of Denmark (H-18023136; August 27, 2018).

Randomization and Masking

Copenhagen Trial Unit, a Danish center for clinical intervention research, was responsible for the central randomization. Participants were randomly assigned (1:1) to short-term MBT or long-term MBT by an independent investigator not otherwise involved in the trial according to a computer-generated allocation sequence with permuted blocks of various sizes. Randomization was stratified by sex (male/female) and low or high (0-11/12-36) BPD symptom severity at baseline assessed with the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) [16].

Outcome assessors, the data safety and monitoring committee, data managers, statisticians, and the steering committee were blind to treatment allocation [13, 14]. Trial participants and therapists were not blind to treatment allocation due to the difficulties of implementing efficient blinding procedures in trials assessing psychological interventions [17]. Blinding was broken on December 18, 2022, after the statistical report was finalized and after the steering group had agreed on two blinded abstracts.

Procedures

Patients were initially screened for eligibility during their routine intake interviews at the Outpatient Clinic for Personality Disorders. A psychiatrist or consultant psychiatrist performed a psychiatric assessment, assessed patient eligibility, provided the patient with a trial-specific information booklet, and obtained written informed consent. Eligible and consenting patients were then seen by a researcher who confirmed eligibility with the MINI interview [15] and the Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD) [18], collected demographic information, and performed baseline assessments before randomization.

Three external MBT experts, who were blind to treatment allocation, assessed adherence to the MBT manual in both short-term and long-term MBT using the Mentalization-Based Treatment Adherence and Competence Scale [19]. All raters were certified by one of the developers of the scale (A.B.). A total of 270 videos were obtained. From these videos, 59 videos (22%) from different treatment stages were randomly chosen for adherence ratings.

Short-term MBT was organized as a 20-week program of therapy in closed groups, commencing with five sessions of introduction to MBT followed by 15 weekly sessions of group MBT. Group therapy was accompanied by individual sessions every second week. The participants were treated by two group therapists, one of them also being the individual therapist, i.e., conjoined therapy. The MBT therapists followed the manual developed by Bateman and Fonagy [9].

Long-term MBT was organized as a 14-month program. All participants randomized to long-term MBT initially entered a 6-week introduction to MBT. When the introduction group finished, participants were allocated to one of eight slow-open MBT treatment groups. Treatment was then organized as 12 months of weekly group therapy sessions combined with individual therapy every second week. The long-term MBT group consisted of a total of 13.5 months of active treatment, but since there was some waiting time from the MBT intro group to the beginning of the MBT treatment group, we pragmatically defined the long-term MBT intervention as 14 months of duration. The participants were treated by two group therapists and a third individual therapist, i.e., combined therapy. Both group and individual therapies were

manualized by Bateman and Fonagy [9]. More information about the trial interventions can be found in the online supplementary material, p. 7.

Outcomes

The primary outcome was severity of borderline symptomatology assessed with the ZAN-BPD interview by a blinded assessor [16]. Secondary outcomes were functional impairment assessed with the Work and Social Adjustment Scale (WSAS) [20], quality of life assessed with the Short-Form Health Survey (SF-36) mental component [21], global functioning assessed with the Global Assessment of Functioning scale (GAF) [22], and severe self-harm defined as the proportion of participants with one or more deliberate act of self-harm resulting in visible tissue damage (not including less severe self-harm, e.g., head-banging, superficial cutting, or hitting oneself without tissue damage).

All outcomes were assessed at 8 months, 16 months, and 24 months (will be presented in a separate publication) after randomization. We predefined the 16-month time point as the assessment time point of primary interest to ensure that the participants were assessed at similar time points after randomization regardless of the length of their intervention.

Safety was assessed as serious adverse events. Two blinded outcome assessors independently went through patient medical records to assess serious adverse events using the International Conference on Harmonization of technical requirements for registration of pharmaceuticals for human use – Good Clinical Practice (ICH-GCP) definition of a serious adverse event (any untoward medical occurrence that resulted in death, was life-threatening, required hospitalization or prolonging of existing hospitalization and resulted in persistent or significant disability or jeopardized the participant) [23].

Statistical Analysis

We calculated that 166 participants were needed to provide 80% power at a 5% significance level for the primary outcome ZAN-BPD. We defined 3.5 points as a minimal important difference for ZAN-BPD.

The analyses followed the statistical analysis plan, which was published before unblinding of the data [14]. All continuous outcomes were analyzed using linear regression and dichotomous outcomes were analyzed using logistic regression. The analyses were based on an intention-to-treat population and were primarily adjusted for the baseline value of the outcome of interest and the stratification variables used in the randomization.

We used a five-step procedure [24] to assess if the thresholds for statistical and clinical significance were crossed, and we handled missing data according to the procedure suggested by Jakobsen et al. [25]. A more detailed explanation of our statistical methodology is available in our published and predefined statistical analysis plan [14].

Results

Between October 4, 2018, and December 3, 2020, 207 patients were initially screened for eligibility to the trial. We randomly assigned 166 participants to short-term MBT ($n = 84$) or long-term MBT ($n = 82$) (Fig. 1).

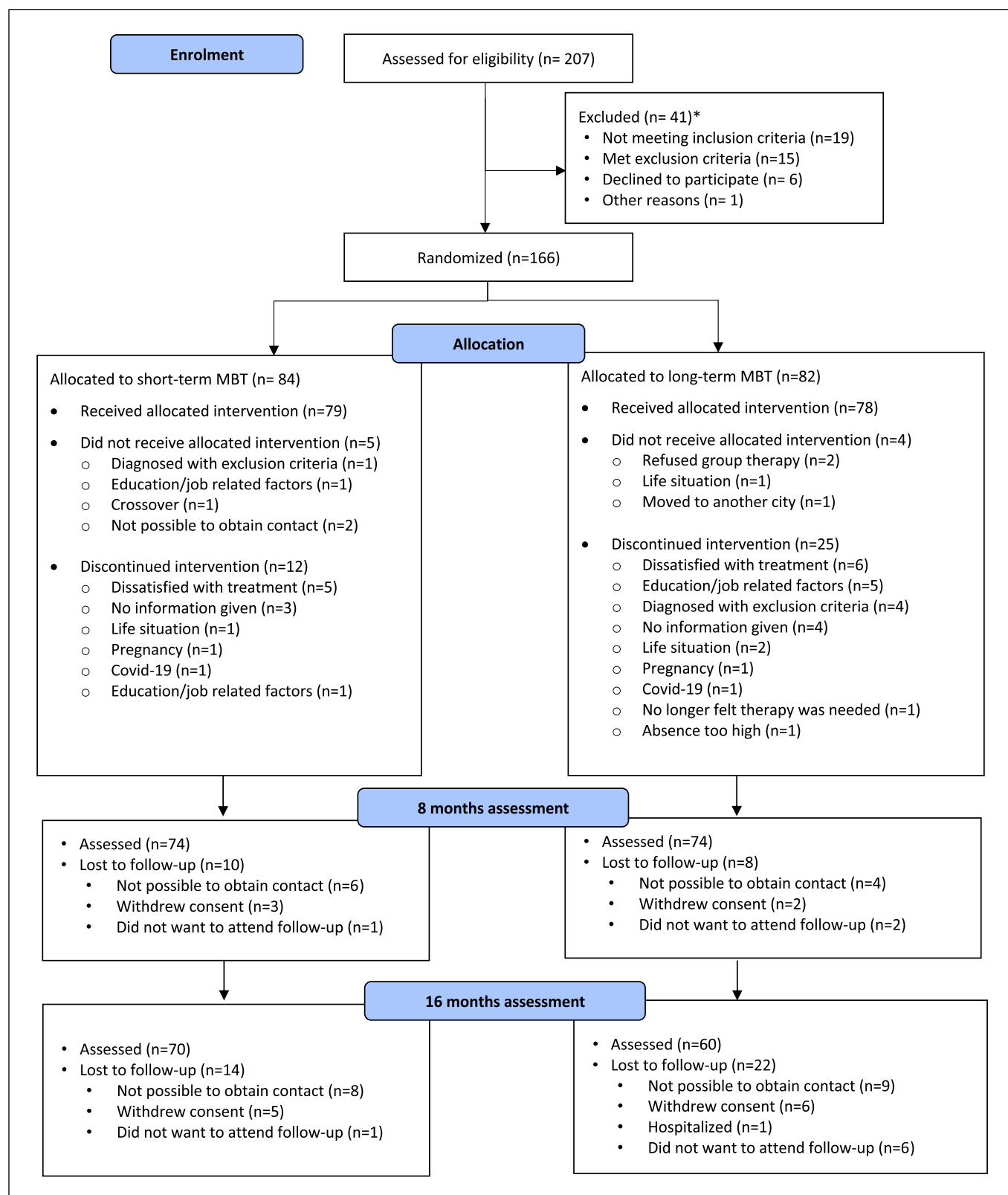


Fig. 1. Consolidated Standards of Reporting Trials (CONSORT) flowchart.

Table 1. Baseline sociodemographic and clinical characteristics

Characteristics	Short-term MBT (<i>n</i> = 84)	Long-term MBT (<i>n</i> = 82)
Sociodemographic characteristics		
Age, mean (SD), years	31.7 (10.1)	32.2 (10.1)
Female sex, <i>n</i> (%)	75 (89.1)	76 (92.7)
Civil status, <i>n</i> (%)		
Single	45 (53.6)	41 (50)
In a relationship (living together)	20 (23.8)	13 (15.9)
In a relationship (living separately)	11 (13.1)	19 (23.2)
Married	8 (9.5)	9 (10.9)
Living situation, <i>n</i> (%)		
With others	48 (57.1)	53 (64.6)
Alone	36 (42.9)	28 (34.2)
Other	0 (0)	1 (1.2)
Education level after high school, <i>n</i> (%)		
Short education	15 (17.9)	11 (13.4)
Medium education	20 (23.8)	17 (20.7)
Long education	6 (7.1)	5 (6.1)
No higher education	33 (39.3)	38 (46.3)
Other education	10 (11.9)	11 (13.4)
Employment status, <i>n</i> (%)		
On social welfare (/unemployed)	38 (45.2)	47 (57.3)
Currently studying	21 (25)	19 (23.2)
Employed	18 (21.4)	14 (17.1)
Unemployed/homemaker	2 (2.4)	0 (0)
Retired	1 (1.2)	0 (0)
Self-employed	1 (1.2)	1 (1.2)
Other	3 (3.6)	1 (1.2)
Clinical characteristics		
Psychiatric comorbidity (MINI), <i>n</i> (%)		
Anxiety disorders	42 (50)	53 (64.6)
Major depressive disorder	30 (35.7)	28 (34.1)
Post-traumatic stress disorder	17 (20.2)	24 (29.3)
Eating disorders	12 (14.3)	12 (14.6)
Obsessive compulsive disorder	10 (11.9)	7 (8.5)
Proportion of participants with subthreshold BPD, <i>n</i> (%)	20 (23.8)	13 (15.9)
No. of BPD criteria in total (SCID-5-PD), mean (SD)	6.2 (1.4)	6.6 (1.1)
No. of PD criteria in total (SCID-5-PD), mean (SD)	18.1 (4.9)	17.9 (5.5)
Personality disorder comorbidity (SCID-5-PD), <i>n</i> (%)		
Avoidant personality disorder	25 (29.8)	21 (25.6)
Obsessive-compulsive personality disorder	14 (16.7)	27 (32.9)
Paranoid personality disorder	15 (17.9)	20 (24.4)
Dependent personality disorder	8 (9.5)	7 (8.5)
Proportion of participants with one or more suicide-attempts the past 8 months, <i>n</i> (%)	5 (5.9)	6 (7.3)
Proportion of participants with one or more acts of severe self-harm the past 8 months, <i>n</i> (%)	32 (38.1)	33 (40.2)
Proportion of participants on antidepressants, <i>n</i> (%)	33 (39.3)	35 (42.7)
Proportion of participants on antipsychotics, <i>n</i> (%)	10 (11.9)	11 (13.3)
Proportion of participants on other psychoactive drugs, <i>n</i> (%)	21 (25.0)	19 (23.2)
Mean (SD) number of days from randomization to assessment time point		
8 months post-randomization	213 (12.4)	219 (13.8)
16 months post-randomization	389 (15.2)	354 (16.4)

BPD, borderline personality disorder; MINI, Mini International Neuropsychiatric Interview; No., number; SCID-5-PD, Structured Clinical Interview for DSM-5 Personality Disorders; SD, standard deviation.

Table 2. Primary and secondary outcome results (ITT population)

Outcome	Short-term MBT (<i>n</i> = 84)		Long-term MBT (<i>n</i> = 82)		Estimate (95% CI)	<i>p</i> value
	No. analyzed	result	No. analyzed	result		
Primary						
Borderline symptoms, ZAN-BPD, mean (SD)						
Baseline	84	13.9 (3.72)	82	13.8 (3.75)		
8 months	74	12.0 (5.33)	74	12.8 (5.01)	MD 0.95 (−0.57 to 2.46)	0.22
16 months	70	11.3 (6.22)	60	12.3 (5.91)	MD 0.99 (−1.06 to 3.03)	0.341
Secondary						
Functional impairment, WSAS, mean (SD)						
Baseline	84	22.9 (7.54)	82	21.9 (7.46)		
8 months	71	20.4 (8.59)	69	19.9 (8.36)	MD 0.50 (−1.95 to 2.95)	0.687
16 months	63	16.7 (8.91)	56	17.2 (9.36)	MD 1.44 (−1.43 to 4.32)	0.321
Quality of life, SF-36 – mental component, mean (SD)						
Baseline	84	50.3 (10.4)	82	49.2 (10.1)		
8 months	71	50.4 (10.8)	69	53.1 (11.6)	MD 2.97 (−0.78 to 6.72)	0.120
16 months	63	53.9 (9.94)	56	52.9 (10.2)	MD −0.91 (−4.62 to 2.79)	0.626
Global functioning, GAF, mean (SD)						
Baseline	84	46.5 (8.14)	82	46.3 (6.36)		
8 months	74	49.6 (13.5)	74	47.7 (10.5)	MD −1.28 (−5.05 to 2.49)	0.502
16 months	70	52.2 (15.0)	61	49.7 (11.7)	MD −2.25 (−6.70 to 2.20)	0.318
Proportion of participants with severe self-harm the past 8 months						
Baseline	84	32/84 (38.1%)	82	33/82 (40.2%)		
8 months	74	19/74 (25.7%)	74	27/74 (36.5%)	RR 1.43 (0.87–2.45)	0.153
16 months	68	13/68 (19.1%)	61	16/61 (26.2%)	RR 1.37 (0.70–2.84)	0.335

GAF, Global Assessment of Functioning; ITT, intention to treat; MD, mean difference; OR, odds ratio; SF-36, Short Form Health Survey – 36; SD, standard deviation; WSAS, Work and Social Adjustment Scale; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

In the short-term MBT group, five participants did not receive the allocated intervention, and 12 participants discontinued the intervention prematurely for various reasons. One participant randomized to short-term MBT received two consecutive short-term MBT group interventions. One participant randomized to short-term MBT crossed over to long-term MBT. According to our protocol, they were both included in the primary analysis per intention-to-treat principle. In the long-term MBT group, four participants did not receive the allocated intervention, and 25 discontinued the intervention prematurely for various reasons (Fig. 1).

Baseline sociodemographic and clinical characteristics can be found in Table 1. The psychopharmacological co-interventions were delivered similarly in both groups across the trial period (online suppl. material, p. 22). Overall, the two groups were balanced across baseline characteristics.

Participants allocated to short-term MBT attended a mean (SD) of 12 (6.2) individual therapy sessions and 14 (6.9) group sessions over a period of 5 months. Participants allocated to long-term MBT attended a mean of 17 (11) individual therapy sessions and 24 (14) group

sessions over a period of 14 months. Attendance data were missing for 21 participants. These data are based on the intention-to-treat population. Attendance data for the protocol population can be found in the online supplementary material, p. 22.

Independent and certified coders rated a random selection of 35 individual sessions and 24 group sessions with the Mentalization-Based Treatment Adherence and Competence Scale [19]. The mean (SD) score was 4.46 (0.94) (*n* = 30 videos) in the short-term MBT and 4.79 (1.1) in the long-term MBT (*n* = 29 videos) (≥ 3.5 was considered adherent) with no evidence of a difference between short-term and long-term MBT (*t* test: *p* = 0.22). Overall adherence to MBT was good for both individual therapy (4.3, 0.9) and group therapy (4.7, 0.8) in the short-term group and individual therapy (4.6, 1.1) and group therapy (5.0, 1.1) in the long-term group. Table 2 summarizes primary and secondary outcome data in the intention-to-treat population.

Linear regression showed no evidence of a difference between short-term and long-term MBT at 16 months when assessing borderline symptoms (ZAN-BPD MD

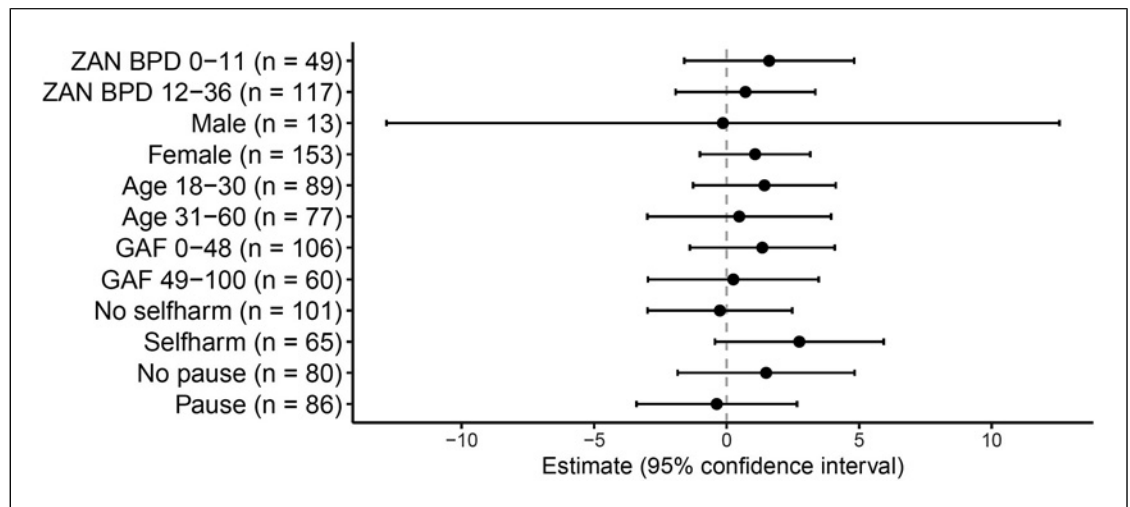


Fig. 2. Forest plot of subgroup analyses.

0.99; 95% CI: -1.06 to 3.03; $p = 0.341$). 14 (16.7%) of 84 participants in the short-term MBT group were lost to follow-up compared with 22 (26.8%) of 82 participants in the long-term MBT group. Sensitivity analysis showed that missing data had the potential to alter the results (online suppl. material, pp. 23–107). None of the pre-planned subgroup analyses (baseline severity of borderline symptoms; sex; age; baseline global functioning; baseline proportion of participants with severe self-harm; proportion of participants who had their group therapy temporarily paused due to COVID-19) showed evidence of a difference (Fig. 2). The outcome assessed at 8 months after randomization showed similar results (online suppl. material, pp. 23–107). Inter-rater reliability for ZAN-BPD was high (intraclass correlation: 0.92) based on the ratings of three independent assessors of a random subsample of 40 participants.

Linear regression showed no evidence of a difference between short-term and long-term MBT at 16 months when assessing the level of functioning (WSAS MD 1.44; 95% CI: -1.43 to 4.32; $p = 0.321$). The outcome assessed at 8 months after randomization showed similar results (online suppl. material, pp. 23–107).

Linear regression showed no evidence of a difference between short-term and long-term MBT at 16 months when assessing quality of life (SF-36 MD -0.91; 95% CI: -4.62 to 2.79; $p = 0.626$). The outcome assessed at 8 months after randomization showed similar results (online suppl. material, pp. 23–107).

Linear regression showed no evidence of a difference between short-term and long-term MBT at 16 months when assessing global functioning (GAF MD -2.25; 95%

CI: -6.70 to 2.20; $p = 0.318$). The outcome assessed at 8 months after randomization showed similar results (online suppl. material, pp. 23–107).

At 16 months, 13 of 68 (19.1%) participants performed self-harm in the short-term MBT group compared with 16 of 61 (26.2%) participants in the long-term MBT group (RR 1.37; 95% CI: 0.70–2.84; $p = 0.335$). The outcome assessed at 8 months after randomization showed similar results (online suppl. material, pp. 23–107).

At 16 months after randomization, 15 of 84 (17.9%) participants experienced a serious adverse event in the short-term MBT group compared with 24 of 82 (29.3%) of the participants in the long-term MBT group (RR 1.63; 95% CI 0.94–3.07; $p = 0.088$) (online suppl. material pp. 23–107). Logistic regression showed no evidence of a difference but indicated a harmful effect of long-term MBT primarily driven by a difference in psychiatric hospitalizations: 9 of 84 (10.7%) participants had a psychiatric hospitalization in the short-term MBT group compared with 18 out of 82 (22%) in the long-term MBT group (RR 2.03; 95% CI: 0.99–5.09; $p = 0.056$) (online suppl. material, p. 108). A table of individual serious adverse events can be found in the online supplementary material, p. 110.

Discussion

This is the first randomized clinical trial assessing short-term versus long-term MBT for adult outpatients with BPD. Our analyses showed that at 16 months after

randomization, no evidence of a difference was observed either when assessing BPD symptomatology, level of functioning, quality of life, global functioning, and self-harm. Serious adverse events occurred more often in the long-term MBT group primarily driven by more psychiatric hospitalizations, but this result was not statistically significant. The therapists in the trial showed good adherence to the MBT manual, and were found to be delivering MBT as intended.

Our trial has important practical implications for the treatment of BPD, as the high prevalence of the disorder combined with high costs of existing treatments causes a significant barrier to accessing evidence-based psychotherapies for BPD in many parts of the world [26]. Given that access to evidence-based psychotherapy for BPD is a problem for many patients due to scarce resources in the mental health systems, shortening the overall length of psychotherapy in outpatient mental health settings may increase access to evidence-based care. Our trial challenges the prevailing opinion that people with personality disorders require long-term psychotherapy and supports the use of briefer psychotherapy programs for BPD. Based on our results, administrators and clinicians in the mental health system may give serious consideration to short-term MBT as a desirable option for outpatients with BPD.

In our trial, more patients in the long-term MBT group experienced a serious adverse event primarily driven by a difference in psychiatric hospitalizations. As the follow-up period was identical for the compared groups (i.e., 16 months after randomization), our analyses might indicate a “true” difference in the risk of serious adverse events, because the participants in both groups were observed in comparable follow-up time periods. While this difference was not statistically significant, it is still worth commenting on, as the confidence interval (0.99–5.08) showed that benefit could almost be excluded and an increased risk of psychiatric hospitalizations in the long-term MBT group was possible. Furthermore, thresholds for statistical significance should be less conservative when considering increased risks of harmful effects of experimental interventions [27]. The difference in serious adverse events might be a chance finding (“random error”); MBT may overwhelm BPD patients and lead to psychiatric hospitalizations; and/or being in MBT may increase odds that patients contact acute psychiatric care in times of crisis.

This trial has several strengths. First, we used a rigorous trial methodology to reduce the risks of bias, and

our methodology was predefined in our published protocol and detailed statistical analysis plan prior to unblinding the results [13, 14]. Second, we assessed both beneficial and harmful effects and included objective assessments of serious adverse events, which is rare in psychotherapy trials for BPD [8]. Third, our trial has a high degree of external validity, as it was conducted in a real-world outpatient setting, and the excellent compliance with MBT suggests that this approach could be acceptable for implementation in routine clinical use with appropriate training and supervision. Fourth, our analyses showed limited signs of heterogeneity as all subgroup analyses showed no evidence of a difference, and our results were consistent across most outcomes. Fifth, all participants were assessed at the same time points regardless of their allocated intervention taking into account “regression toward the mean” potentially influencing the present results. Sixth, the psychopharmacological treatment was comparable between groups at all time points, reducing the risk of differences in psychopharmacological treatment influencing the present results.

This trial also has limitations. First, consistent with other trials assessing participants with BPD [8], we experienced a large proportion of missing outcome data. Large proportions of missing data should always warrant cautious interpretations of results [25]. Second, we included participants with subthreshold or diagnosed BPD. Hence, our trial population was heterogeneous and, on average, showed lower baseline BPD symptoms, including lower proportions of self-harm and higher levels of global functioning, as compared to previous trials assessing the effects of MBT for BPD [28, 29]. Our analyses did not indicate that the effects of MBT were influenced by baseline severity, but subthreshold BPD is not a diagnosis defined in either DSM-5 [1] or the International Classification of Diseases – eleventh version (ICD-11) [30] which is a limitation of the present trial. Our results may only be generalized to the population we included in the trial, i.e., participants with either subthreshold or diagnosed BPD. This should be considered, when interpreting our results. We chose this approach for pragmatic reasons, as many outpatients in everyday clinical practice do not meet the criteria for a full BPD diagnosis but are still functionally impaired, and we wished to include these patients to increase the external validity of the trial. Similar inclusion criteria have been used in other trials assessing psychotherapy for BPD [31, 32]. Third, the long-term MBT group in this trial was 14 months of MBT consisting of psychoeducation followed by weekly group therapy combined with individual

therapy every second week. Our long-term MBT intervention is not identical to the original, manualized long-term MBT program, which is 18 months of treatment with both weekly group and individual therapies [9]. Hence, the evidence base for the long-term MBT group in this trial compared with treatment as usual is unclear. Fourth, clinicians and participants were not blind to the treatment allocation, which could potentially bias the results [17]. Fifth, we predefined the minimal clinically important difference (MCID) for the primary outcome as a 3.5-point superiority margin. However, this estimation was based on clinical experience with the population and an approximation of the expected standard deviation (SD/2) from previous trials [13]. Hence, the estimation of the MCID in this trial is uncertain and might be smaller or larger than 3.5. This has to be considered when interpreting the results. Furthermore, some of the items in the ZAN-BPD scale might reflect more severe problems than others, and the MCID might depend on the baseline severity, further leading to problems interpreting our present results. Finally, due to the COVID-19 pandemic, we experienced protocol deviations during the trial period, as group therapy in both trial interventions was temporarily paused for 2 months in 2020 and again in 2021. According to our subgroup analyses, these pauses did not seem to influence the outcomes, and they were identical in both groups.

Different short-term psychotherapies exist for patients with BPD, including brief dialectical behavior therapy (DBT) for 20 weeks [33], emotion regulation group therapy (ERGT) for 14 weeks [32], and systems training for emotional predictability and problem solving (STEPPS) for 20 weeks [34]. However, only one previous randomized clinical trial has assessed the effects of different durations of psychotherapy for BPD. The results of our trial are similar to a recent noninferiority trial assessing the effects of 6 versus 12 months of DBT for BPD [35], but this trial assessed the effects of a different intervention and included a narrower participant group, assessing only acute participants (eligibility criteria were one or more suicide attempts or acts of self-harm within the past 2 months). The authors concluded that the noninferiority hypothesis was supported for the primary outcome, total self-harm, and for several secondary outcomes.

In conclusion, long-term MBT did not lead to lower levels of BPD symptoms, nor did it influence any of the secondary outcomes compared with short-term MBT. In light of scarce treatment resources for BPD, briefer psychotherapy (5 months instead of 14 months) may be considered a desirable option for outpatients with BPD.

Acknowledgments

We thank the trial participants and the trial clinicians from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark; the independent members of the Data Safety and Monitoring Committee (Prof. Ole Jakob Storebø and Prof. Mickey Kongerslev); our student assistants (Mathilde Hasselby-Andersen, Anders Sonne Munch, Amanda Ark Søndergaard, Laura Alsing Juul, Rebecca Engel Thomas, Rebecca Kjær Andersen, and Olivia Frost); and the independent coders of MBT adherence (Nana Lund Nørgaard and Anne Blom Corlin).

Statement of Ethics

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Regional Ethics Committee of Health Research of the Capital Region of Denmark (approval number H-18023136; August 27, 2018) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All participants provided written informed consent.

Conflict of Interest Statement

A.B. declares receiving book royalties from Oxford University Press and American Psychiatric Publishing Inc. for publishing books on mentalizing. A.B. declares payment for training courses in MBT from Anna Freud Centre, London. No other disclosures were reported.

Funding Sources

The trial was funded by Trygfonden (S.J., Grant No. 123488) and by the Mental Health Services in the Capital Region of Denmark Research Fund (S.J. and S.S., Grant No. N/A). The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had access to all study data and had final responsibility for the decision to submit it for publication.

Author Contributions

S.J. and S.S. conceptualized the trial, secured the funding, designed the short-term MBT intervention in collaboration with A.B. and clinicians at the outpatient clinic, and were the principal investigator (S.J.) and sponsor (S.S.) of the trial. PS provided organizational support for the trial. F.W.F. assisted in the clinical implementation of the trial. The trial was overseen by a steering committee consisting of S.J., S.P., P.S., J.C.J., S.L., and S.S. S.J. recruited participants and performed baseline assessments assisted by E.H., C.K.J., and M.R. Blinded outcome

assessments were performed by E.H., C.K.J., M.R., and several student assistants. A.B. and S.B. trained and supervised the staff in both short-term and long-term MBT. J.C.J. supervised the design of the trial protocol and statistical analysis plan. M.H.O. and J.C.J. independently performed blinded statistical analyses. S.J., J.C.J., and M.H.O. had full access to all the blinded trial data and took responsibility for the integrity of the data and the accuracy of the data analysis. S.B. assisted in the coding of adherence to MBT. S.J., J.C.J., and S.S. drafted the paper. All authors commented on and approved the final version of the paper.

Data Availability Statement

Data from this trial are not publicly available due to ethical grounds. Access to deidentified data from this trial may be obtained by contacting the corresponding author. All requests for data access must be approved by the Regional Ethics Committee of Health Research of the Capital Region of Denmark and the institution with whom the primary investigator is affiliated. Data access will be governed by a data-sharing agreement executed between the Capital Region of Denmark and the institution with whom the primary investigator is affiliated.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub; 2013. Report No: 0890425574.
- Togersen J. Prevalence, sociodemographics, and functional impairment. In: Oldham J, Skodol AE, Bender DS, editors. *Textbook of personality disorders*. 2nd ed. Washington (DC): American Psychiatric Publishing; 2014. p. 109–31.
- Zimmerman M, Rothschild L, Chelminski I. The prevalence of DSM-IV personality disorders in psychiatric outpatients. *Am J Psychiatry*. 2005;162(10):1911–8.
- Sansone RA, Sansone LA. Employment in borderline personality disorder. *Innov Clin Neurosci*. 2012;9(9):25–9.
- Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385(9969):735–43.
- National Institute for Health and Care Excellence (NICE). Borderline personality disorder: recognition and management. Clinical guideline [CG78]. 2009. Available from: <https://www.nice.org.uk/guidance/cg78> [accessed September 19, 2022].
- Leichsenring F, Rabung S. Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis. *Br J Psychiatry*. 2011;199(1):15–22.
- Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020;5(5):CD012955.
- Bateman A, Fonagy P. *Mentalization-based treatment for personality disorders: a practical guide*. Oxford, England: Oxford University Press; 2016.
- Juul S, Jakobsen JC, Jørgensen CK, Poulsen S, Sørensen P, Simonsen S. The difference between shorter- versus longer-term psychotherapy for adult mental health disorders: a systematic review with meta-analysis. *BMC Psychiatry*. 2023;23(1):438.
- Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Obstet Gynecol*. 2010;115(5):1063–70.
- Guidi J, Brakemeier E-L, Bockting CLH, Cosci F, Cuijpers P, Jarrett RB, et al. Methodological recommendations for trials of psychological interventions. *Psychother Psychosom*. 2018;87(5):276–84.
- Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019; 20(1):196.
- Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, 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; 22(1):497.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59(Suppl 20):22–33; quiz 34–57.
- Zanarini MC, Vujanovic AA, Parachini EA, Boulanger JL, Frankenburg FR, Hennen J. Zanarini rating scale for borderline personality disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Pers Disord*. 2003;17(3):233–42.
- Juul S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med*. 2021; 26(3):109.
- First MB, Williams JB, Karg RS, Spitzer RL. *User's guide for the SCID-5-CV structured clinical interview for DSM-5® disorders: clinical version*. American Psychiatric Publishing, Inc.; 2016.
- Bateman A, Bales D. *MBT. HJAQMf. A quality manual for MBT*. 2020:Available from: <https://www.annafreud.org/media/14278/adherence-scale-august-2020.pdf> [accessed January 10, 2023].
- Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personal Ment Health*. 2017;11(4):215–28.
- Brazier JE, Harper R, Jones N, O'cathain A, Thomas K, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160–4.
- Hall RC. Global assessment of functioning: a modified scale. *Psychosom*. 1995;36(3): 267–75.
- International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use). *ICH harmonised guideline: integrated addendum to ICH E6(R1): guideline for good clinical practice (ICH-GCP)*. 2015.
- Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance: a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol*. 2014;14(34):34.
- Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials: a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.
- Iliakis EA, Sonley AK, Ilagan GS, Choi-Kain LW. Treatment of borderline personality disorder: is supply adequate to meet public health needs? *Psychiatr Serv*. 2019;70(9): 772–81.
- European Medicines Agency. *Guideline on multiplicity issues in clinical trials*. EMA/CHMP/44762/2017; 2016.
- Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12): 1355–64.

- 29 Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563–9.
- 30 World Health Organization (WHO). *International classification of Diseases: Eleventh revision (ICD-11)*; 2019/2021.
- 31 Jørgensen MS, Storebø OJ, Bo S, Poulsen S, Gondan M, Beck E, et al. Mentalization-based treatment in groups for adolescents with Borderline Personality Disorder: 3- and 12-month follow-up of a randomized controlled trial. *Eur Child Adolesc Psychiatry*. 2021; 30(5):699–710.
- 32 Gratz KL, Tull MT, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099–112.
- 33 McMain SF, Guimond T, Barnhart R, Habinski L, Streiner DL. A randomized trial of brief dialectical behaviour therapy skills training in suicidal patients suffering from borderline disorder. *Acta Psychiatr Scand*. 2017;135(2):138–48.
- 34 Blum N, St John D, Pfohl B, Stuart S, McCormick B, Allen J, et al. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *Am J Psychiatry*. 2008;165(4):468–78.
- 35 McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022;91:382–97.



Supplementary material

for

Short-term versus long-term mentalization-based therapy for borderline personality disorder: a randomized clinical trial (MBT-RCT)

Table of contents

CONSORT 2020 Checklist	1
Trial protocol	3
Statistical analysis plan (SAP)	13
Differences between the protocol/SAP and the manuscript	21
Psychopharmacological co-interventions during the trial period	22
Attendance data	23
Final statistical report for the MBT-RCT trial	25
Post-hoc statistical analyses for the MBT-RCT trial	60
Individual serious adverse events	111



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	6
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	8
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	12
Participants	4a	Eligibility criteria for participants	8
	4b	Settings and locations where the data were collected	8
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	10-11
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	11-12
	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines	8
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	9
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	9
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	9
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	9
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	9

		assessing outcomes) and how	
Statistical methods	11b	If relevant, description of the similarity of interventions	10-11
	12a	Statistical methods used to compare groups for primary and secondary outcomes	12-13
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12-13
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	13-14, Figure 1
	13b	For each group, losses and exclusions after randomisation, together with reasons	Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up	13
	14b	Why the trial ended or was stopped	NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Table 1, pp15-17
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	15-17
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	15-17
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	15-17
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	16-17
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	19-20
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	19-20
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	17-20
Other information			
Registration	23	Registration number and name of trial registry	8
Protocol	24	Where the full trial protocol can be accessed, if available	12
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	21

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

STUDY PROTOCOL

Open Access



Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial

Sophie Juul^{1,2*} , Susanne Lunn², Stig Poulsen², Per Sørensen¹, Mehrak Salimi¹, Janus Christian Jakobsen³, Anthony Bateman⁴ and Sebastian Simonsen¹

Abstract

Background: Psychotherapy for borderline personality disorder is often lengthy and resource-intensive. However, the current length of outpatient treatments is arbitrary and based on trials that never tested if the treatment intensity could be reduced. As a result, there is insufficient evidence to inform the decision between short-term and long-term psychotherapy for borderline personality disorder. Mentalization-based therapy is one treatment option for borderline personality disorder and consists traditionally of an 18-month treatment program.

Methods/design: This trial is an investigator-initiated single-center randomized clinical superiority trial of short-term (20 weeks) compared to long-term (14 months) mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder. Participants will be recruited from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark. Participants will be included if they meet a minimum of four DSM-V criteria for borderline personality disorder. Participants will be assessed before randomization, and at 8, 16, and 24 months after randomization. The primary outcome is severity of borderline symptomatology assessed with the Zanarini Rating Scale for borderline personality disorder. Secondary outcomes include self-harm incidents, functional impairment (Work and Social Adjustment Scale, Global Assessment of Functioning) and quality of life (Short-Form Health Survey 36). Severity of psychiatric symptoms (Symptom Checklist 90-R) will be included as an exploratory outcome. Measures of personality functioning, attachment, borderline symptoms, group alliance, and mentalization skills will be included to explore potential predictors and mechanisms of change.

Discussion: This trial will provide evidence of the beneficial and harmful effects of short-term compared to long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder.

Trial registration: ClinicalTrials.gov, [NCT03677037](https://clinicaltrials.gov/ct2/show/study/NCT03677037). Registered on September 19, 2018.

Keywords: Mentalization-based therapy, Borderline personality disorder, Randomized clinical trial, Treatment intensity

* Correspondence: sophie.juul@regionh.dk

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Gentofte, Capital Region of Denmark, Denmark

²Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

Borderline personality disorder is a psychiatric condition characterized by a pervasive pattern of symptoms such as interpersonal conflicts, identity diffusion, impulsivity, and emotional dysregulation [1]. According to epidemiological studies, 1.6% of the general population suffer from borderline personality disorder [2]. In clinical populations, it is the most common personality disorder [2], with a prevalence of between 9% and 22% of all psychiatric outpatients [3–5]. Borderline personality disorder is associated with high levels of psychiatric comorbidity, particularly depression, anxiety disorders, eating disorders, substance abuse [6–8], and other personality disorders [9]. Together, these findings emphasize the need for the development of efficacious and cost-effective treatments for this severe and highly prevalent disorder.

While pharmacological treatment may reduce some borderline-related symptoms, there is still no convincing evidence that it is suitable for treating all diagnostic criteria [10]. Although further evidence is still warranted, psychotherapy continues to be the primary treatment of choice for borderline personality disorder [11]. During the last 10–15 years, studies have established the efficacy of different forms of intensive, specialized long-term psychotherapy modalities. These have recently been evaluated in a systematic review and meta-analysis exploring the efficacy of psychotherapies for borderline personality disorder, in which it was concluded that dialectical behavior therapy and psychodynamic therapies (transference focused therapy and mentalization-based therapy) significantly improved borderline-relevant outcomes [12]. However, no single treatment modality has been established as the primary treatment of choice.

Mentalization-based therapy (MBT) is a psychodynamic therapy rooted in attachment and cognitive theory [13], which was developed specifically for treating borderline personality disorder [14]. Mentalization refers to the capacity to understand one's own and others' mental states. The theoretical assumption is that patients with borderline personality disorder are more vulnerable to lose this capacity when experiencing emotional distress. The MBT manual offers therapeutic techniques to identify these shifts and to bring the patient back into a mentalizing mode [14, 15]. The therapy program consists of four basic components: (1) psycho-education, (2) case formulation, (3) group therapy, and (4) individual therapy. All of these aim to enhance the patient's capacity to mentalize. Increasing mentalization skills is assumed to minimize borderline-related symptoms such as emotional dysregulation, impulsivity, and suicidal ideation. However, information about the mechanisms that produce a change in MBT, or in psychotherapy in general, is still limited [16, 17].

MBT for adult borderline personality disorder has been tested in cohort studies [18, 19] and one randomized but uncontrolled trial [20]. Two forms of MBT have been tested in randomized controlled trials: day hospital MBT [21, 22] and intensive outpatient MBT [13], each lasting a maximum of 18 months. For a systematic review of the current evidence base of MBT for borderline personality disorder, see Vogt and Norman [16].

Bateman and Fonagy [13] assessed the effects of intensive outpatient MBT in a randomized clinical trial, in which 134 participants with a confirmed borderline personality disorder diagnosis were randomized either to 18 months of outpatient MBT, combining weekly group and individual sessions with different therapists, or to structured clinical management. In this trial, MBT was superior to structured clinical management in terms of its effects on suicide attempts, severe incidents of self-harm, and on self-reported measures. Treatment effects were sustained at the 5-year follow-up [23]. Nevertheless, only 134 participants were randomized, which questions whether the trial was powered to assess the chosen outcomes, and only 41 were assessed after 5 years. Further, the trial investigators were also the developers of MBT. Thus, the small sample size and the substantial problems with incomplete outcome data, especially at the long-term follow-up, are threats to the validity of the study.

However, while intensive outpatient MBT currently has empirical support as an 18-month program for borderline personality disorder, evidence that this is the optimal length of the intervention is not available. Consequently, MBT is now offered for different lengths of time (both shorter and longer) in outpatient settings around the world [14]. Various other short-term psychotherapies for borderline personality disorder have already been developed and tested in randomized clinical trials, e.g., emotion regulation group therapy [24], systems training for emotional predictability and problem-solving [25, 26], and brief dialectical behavior therapy skills training [27]. However, all the trials have either compared a short-term experimental group to a short-term control group or tested the short-term treatment as an adjunctive to treatment as usual. Thus, these trials do not provide guidance on evidence-based decisions regarding the optimal length of treatment for borderline patients. In addition, no empirical evidence is available to identify which subtypes of patients would benefit from short-term treatment and which would require more intensive treatment [28].

We performed a preliminary literature search (PubMed and Cochrane Library) for trials comparing different lengths of psychotherapy for borderline personality disorder. No such trials were found. When we expanded our search terms to all types of psychiatric disorders, only few trials were identified [29, 30]. We are

currently working on a protocol for a more comprehensive systematic review, including a full assessment of risk of bias and a trial sequential analysis of short-term compared with long-term psychotherapy for all psychiatric disorders. The systematic review will be submitted for publication before data collection is completed in this trial.

Methods/design

Objective

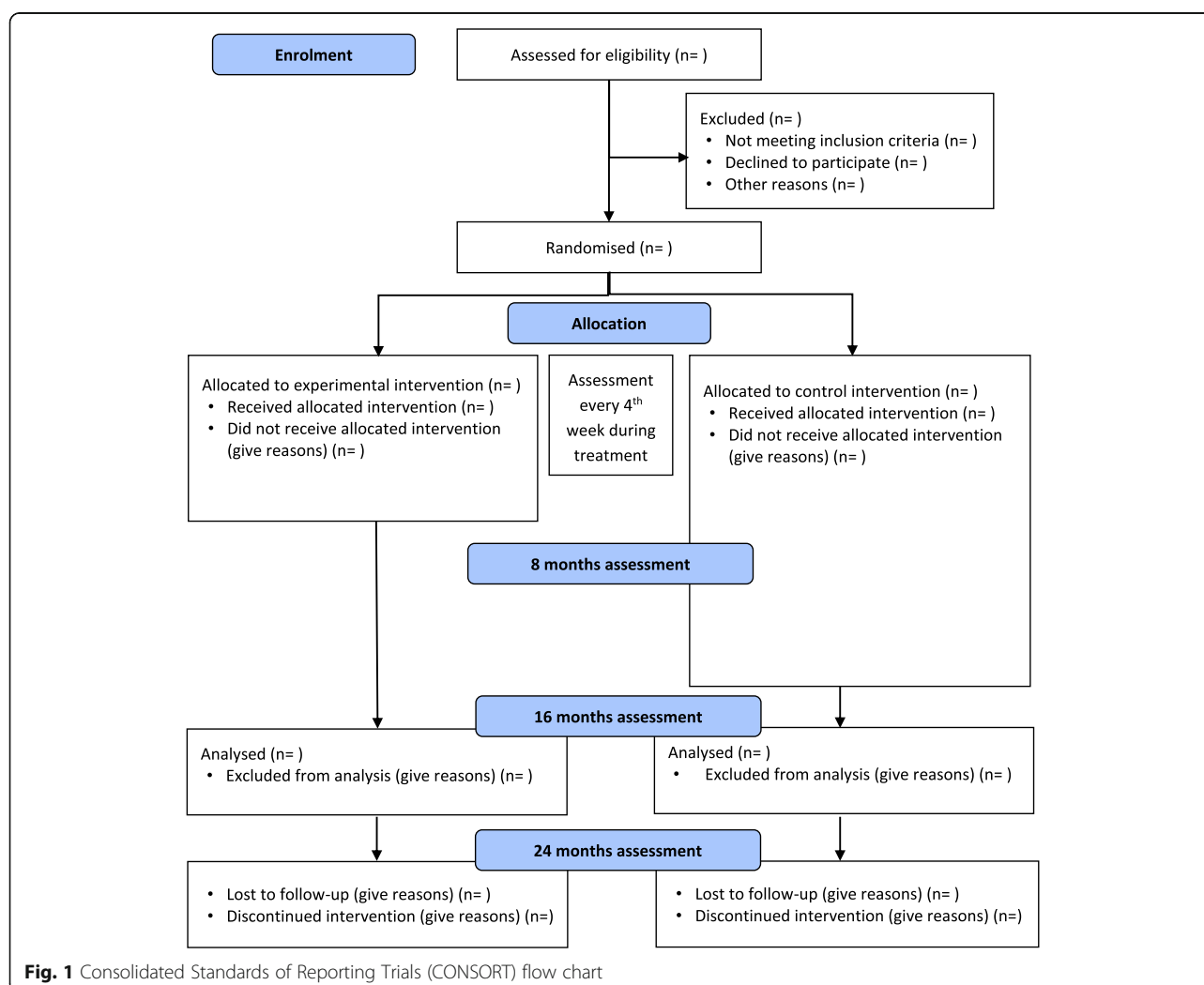
The primary objective of this trial will be to evaluate the beneficial and harmful effects of short-term (20 weeks) MBT compared with long-term (14 months) MBT for adult outpatients with subthreshold or diagnosed borderline personality disorder. We will evaluate the treatments on the primary outcome (borderline symptomatology), secondary outcomes (self-harm incidents, quality of life, and functional impairment, and exploratory outcomes (psychiatric symptoms). Measures of personality functioning, attachment, borderline psychopathology, group

cohesion, and mentalization skills will be included as predictor and mediator variables.

Design

We have designed an investigator-initiated parallel-group single-centre randomized clinical superiority trial of short-term versus long-term MBT for outpatients with subthreshold or diagnosed borderline personality disorder. The Consolidated Standards of Reporting Trials (CONSORT) flow chart for the trial is shown in Fig. 1. [31, 32]. The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) participant timeline is given in Fig. 2, and the SPIRIT checklist is given in Additional file 1 [33].

We will consider for participation all patients referred to the trial site. Patients will be included in the trial, if they comply with the eligibility criteria listed in Table 1. There are inclusion and exclusion criteria as part of the procedure for clinical intake at the trial site, and criteria specific to this trial. For a detailed overview of typical



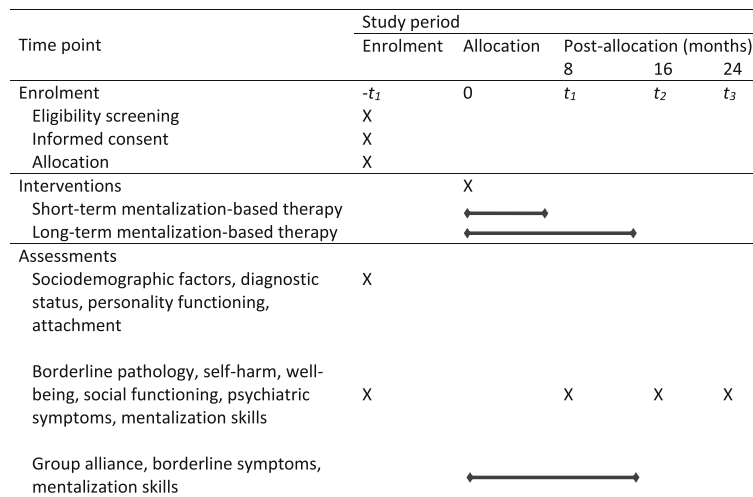


Fig. 2 Participant timeline for the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)

patient characteristics at the trial site, see Simonsen et al. [34].

We will include participants with at least subthreshold borderline personality disorder. According to the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V) [1], the threshold for a full diagnosis is five out of nine diagnostic criteria. However, there is increasing evidence that even four confirmed diagnostic criteria are not qualitatively different from diagnosed borderline personality disorder in terms of impairment, and that the diagnostic threshold should be more inclusive than established by the DSM system to reflect the dimensionality of the construct [35, 36]. For this reason, previous trials have included participants with at least a subthreshold diagnosis [24, 37], and we will do the same in this trial.

Trial site and personnel

The trial site is the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark (from now on referred to as “the clinic”). The clinic specializes in MBT for borderline personality disorder. Patients are

referred from the Capital Region of Denmark via a central visitation unit, where they are initially screened for eligibility before referral to the clinic. Once referred to the clinic, psychiatrists and attending physicians will perform the initial selection and screening of a participant to the trial and collect informed consent. The principal investigator, sponsor-investigator, or a trained research assistant will then conduct the baseline assessments of the participant. All post-baseline assessments will be carried out by trial investigators who are blind to treatment allocation.

Trial therapists provide therapy to both the short-term and long-term treatment groups. Before commencing the trial, all trial therapists at the clinic will have received training in the short-term MBT program by trial investigators and national and international MBT specialists. The training covers relevant topics like case formulations, termination of psychotherapy, and case-specific supervision. The training will continue throughout the trial period. Therapist treatment fidelity will be rated by an independent certified rater. This will allow us to investigate whether the delivered interventions adhere to the MBT manual.

Table 1 Eligibility criteria

	Criteria exclusive to the outpatient clinic	Criteria exclusive to the trial
Inclusion criteria	Aged 18–60 Personality disorder(s) considered to be the primary diagnosis/diagnoses	A minimum of four confirmed DSM-V diagnostic criteria for borderline personality disorder Written informed consent
Exclusion criteria	Possibility of a learning disability (IQ < 75) A full diagnosis of schizotypal personality disorder or antisocial personality disorder Presence of a comorbid psychiatric disorder that requires specialist treatment Current (past 2 months) substance dependence including alcohol Concurrent psychotherapeutic treatment outside the clinic	Unable to understand Danish Lack of informed consent

DSM-V Diagnostic and Statistical Manual of Mental Disorders, 5th edition

Randomization

Copenhagen Trial Unit, a Danish center for clinical intervention research, will be responsible for the central randomization. Trial investigators will call designated staff at Copenhagen Trial Unit using a central telephone to randomize eligible participants to either the experimental group or the control group with a 1:1 allocation according to a computer-generated allocation sequence with permuted blocks of various sizes generated by Copenhagen Trial Unit and unknown to the investigators, secretaries, and clinical staff at the trial site. This is done to eliminate any predictability in the random sequence. The randomization is stratified by (1) sex and (2) high/low scores on the primary outcome measure, the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) [38] at baseline.

Interventions

The short-term MBT program delivered in this trial is overall similar to the existing long-term program, but differs structurally in the following ways: (1) the short-term program is lower in treatment intensity (both duration and exposure), (2) the same therapists provide both group and individual sessions in the short-term program (conjoined psychotherapy), whereas the group therapy and individual therapy are provided by different therapists in the long-term program (combined psychotherapy), and (3) the short-term program is structured in closed groups, in which all participants start and finish the program together, whereas the long-term program is structured as slow-open groups, in which a new participant can enter a group when another finishes. Both interventions in this trial adhere to the treatment guidelines provided by the National Institute for Health and Care Excellence [39].

Experimental intervention

The short-term MBT program is designed as a 20-week psychotherapy program consisting of five sessions of introductory MBT (MBT-I) followed by 15 sessions of group MBT (MBT-G) accompanied by conjoined individual sessions every second week and two psycho-educative meetings with participants and their relatives. Seven to nine participants and two therapists will be included in each short-term MBT group. The groups are closed to enhance cohesion between group participants. A total of 11 short-term MBT groups will be included in this trial.

Originally, MBT-I was a 12-session introductory psycho-educative program covering relevant topics like personality disorders, attachment, and mentalization [40]. The original manual has been modified for our 5-week intervention. A copy of our modified manual is available upon request. After the completion of MBT-I,

the same group of participants will move on to MBT-G, which consists of 15 sessions of mentalization-based psychotherapy in groups, as manualized by Bateman and Fonagy [14]. In our short-term MBT program, group sessions will be accompanied by individual psychotherapy every second week with one of the two group therapists. As part of the individual therapy, a case formulation will be prepared and subsequently shared by the participants in the group. The overall purpose of the individual sessions is for the therapist and participant to develop a consensus of the participant's main difficulties and to establish psychotherapeutic focus points in the group therapy. Furthermore, participants and relatives will be invited to two psycho-educative meetings hosted by the therapists at the beginning of the treatment program to enhance the mentalization work at home. The participants in the short-term MBT program will furthermore be offered three individual follow-up sessions after the end of treatment.

Control intervention

Long-term MBT is organized as a 14-month program and has been implemented at the clinic for the past 7 years. All participants randomized to long-term MBT will initially enter a 6-week MBT-I program manualized by Karterud and Bateman [40] and modified for our 6-week intervention in collaboration with the authors. New MBT-I groups commence every time new participants are recruited and randomized to long-term MBT. A maximum of 12 participants can enter an MBT-I group. When MBT-I finishes, participants will be allocated to one of eight slow-open MBT treatment groups. MBT-G is then organized as 12 months of weekly group therapy sessions, also manualized by Bateman and Fonagy [14]. In the long-term MBT program, group sessions will be accompanied by combined individual MBT sessions every 2 weeks throughout the program. As part of the individual therapy, a case formulation will be developed and subsequently shared with the group by the participant. Furthermore, participants and relatives will also be invited to two psycho-educative meetings hosted by the therapists to enhance the mentalization work at home. When a participant drops out or completes MBT-G, a new participant can start in the group. This procedure continues until the target sample size has been randomized to the long-term MBT program. The participants in the long-term MBT program will also be offered three individual follow-up sessions after the end of treatment.

Concomitant interventions

Participants who are receiving psychotropic treatment will be allowed to continue their medical treatment while participating in the trial. The medical protocol will

follow national as well as international medical recommendations for the treatment of borderline personality disorder and comorbid disorders [39, 41]. Psychiatrists or attending physicians at the clinic will assess the need for additional psychotropic treatment and are asked to adhere to the guidelines. All participants, regardless of treatment condition, will be asked about their current medication by trial personnel during trial interviews to allow us to measure any potential differences in the use of psychotropic medication between the groups.

Baseline assessment at trial intake

Baseline assessments will be carried out prior to randomization by the principal investigator, the sponsor-investigator, and a trained research assistant, all of whom are also clinical psychologists. *General psychopathology* will be assessed with the Mini International Neuropsychiatric Interview (MINI) [42]. *Personality disorders* will be assessed with the Structured Clinical Interview for DSM-V Personality Disorders (SCID-5-PD), formerly known as SCID-II [43]. The SCID-5-PD is considered the gold standard for clinician-administered semi-structured interviews designed to assess personality disorders according to DSM-V criteria [44].

Outcomes

Primary outcome

The primary outcome is the severity of borderline symptomatology assessed with the ZAN-BPD [38], which is a clinician-administered scale for the assessment of change in borderline psychopathology over time. Each of the nine borderline personality disorder criteria are rated on a 0 to 4 anchored scale reflecting the severity of symptoms. The rating is intended to reflect both the frequency and the severity of borderline psychopathology. The interview provides a total score of borderline psychopathology ranging from 0 to 36. ZAN-BPD will be assessed by investigators blind to treatment allocation at baseline, and at the 8-, 16-, and 24-month follow-ups. We will video-record interviews to allow an assessment of inter-rater reliability based on the intraclass correlation coefficient. The results will be evaluated using the guidelines provided by Cicchetti [45].

Secondary outcomes

Functional impairment will be assessed with the Work and Social Adjustment Scale (WSAS) [46, 47]. This self-report scale will be assessed at baseline, and at the 8-, 16-, and 24-month follow-ups. *Quality of life* will be assessed with the Short-Form Health Survey (SF-36) [48], which consists of a mental and a physical component. We will use the mental component as a secondary outcome and the physical component as an exploratory outcome. This self-report scale will

be given at baseline, and at the 8-, 16-, and 24-month follow-ups. *Global functioning* will be measured with the Global Assessment of Functioning (GAF) split version [49]. GAF will be assessed by investigators blind to treatment allocation at baseline, and at the 8-, 16-, and 24-month follow-ups. Inter-rater reliability will be calculated using the previously mentioned guidelines. *Severe self-harm* (dichotomous data) will be measured as the proportion of participants with severe self-harm defined as deliberate acts of self-harm resulting in visible tissue damage. Self-harm will be assessed by investigators blind to treatment allocation using the Suicide and Self-harm Inventory (SSHI) (citation) at baseline, and at the 8-, 16-, and 24-month follow-ups.

Exploratory outcomes

Symptom distress will be measured with the Global Severity Index (GSI) of the Symptom Checklist 90-R (SCL-90-R) [50]. SCL-90-R will be given at baseline, and at the 8-, 16-, and 24-month follow-ups.

Potential predictors and mediators

Questionnaires are given at baseline and every fourth week throughout the intervention period for both intervention groups to allow us to explore predictors and mechanisms of change. In a separate statistical analysis plan, which will be submitted for publication before data collection is completed in this trial, we will describe how these exploratory analyses will be performed. The following predictors and mediators will be investigated.

Personality functioning will be assessed with the Levels of Personality Functioning Scale, Brief Form (LPFS-BF) [51], which is a newly developed brief 12-item self-report questionnaire assessing levels of personality functioning according to the DSM-V alternative model for personality disorders. *Attachment* will be assessed with the brief self-report Relationship Questionnaire (RQ), which gives continuous and categorical ratings of the four attachment styles [52]. *Borderline symptomatology* will be assessed using the Zanarini Rating Scale for Borderline Personality Disorder, Self-Report Version (ZAN-BPD-SRV) [53]. *Group alliance* will be assessed using the 12-item version of the Group Questionnaire (GQ) [54], which is a brief self-report measure of the three core components of group alliance: alliance to the other participants, alliance to the therapists, and group cohesion as a whole. *Mentalization skills* will be assessed with the 15-item Mentalization Questionnaire (MZQ) [55].

For an overview of all measures and the corresponding time of assessment, see Table 2.

Table 2 Assessments administered at baseline and each follow-up point throughout the trial

Assessment points	Self-report measures	Expert ratings
Baseline	LPFS-BF, RQ, SF-36, WSAS, SCL-90-R, MZQ	MINI, SCID-5-PD, ZAN-BPD, GAF, SSHI
Every 4 weeks	GQ, MZQ, ZAN-BPD-SRV	–
Follow-ups at 8, 16, and 24 months	SF-36, WSAS, SCL-90-R, MZQ	ZAN-BPD, SSHI, GAF

GAF Global Assessment of Functioning, GQ Group Questionnaire, LPFS-BF Level of Personality Functioning Scale, Brief Form, MINI Mini International Neuropsychiatric Interview, MZQ Mentalization Questionnaire, RQ Relationship Questionnaire, SCID-5-PD Structured Clinical Interview for DSM-V Personality Disorders, SCL-90-R Symptom Checklist 90-R, SF-36 Short-Form Health Survey 36, SSHI Suicide and Self-Harm Inventory, WSAS Work and Social Adjustment Scale, ZAN-BPD Zanzarini Rating Scale for Borderline Personality Disorder, ZAN-BPD-SRV Zanzarini Rating Scale for Borderline Personality Disorder, Self-Report Version

Blinding

Trial participants and therapists will not be blind to treatment allocation. This is due to the difficulties of implementing an efficient blinding procedure in psychotherapy trials. Baseline assessments will be done before allocation of participants, and the outcome assessments will be performed by blinded assessors. Participants will be instructed to withhold information of their allocation group when assessed. The statistical analyses will be conducted by blinded external statisticians from Copenhagen Trial Unit with the intervention groups coded as A and B. The steering committee will write and agree on two abstracts while the blinding is intact; one assuming the experimental intervention group is A and the control intervention group is B, and the other assuming the opposite. After this, the randomization code will be broken [56, 57].

Participant discontinuation and withdrawal

Participants can withdraw from the trial at any time without giving a reason and without consequences for future treatment at the clinic. To secure data for the trial, a trial investigator will contact the participant and ask what aspects of the trial the participant wishes to withdraw from: (1) the trial intervention or control group, (2) the assessment interviews, or (3) use of already collected data in analyses. If the participant specifies that they wish to withdraw fully and thereby withdraw from all the points above, their data will be deleted and not used in any analysis. The trial investigator will encourage the participant to continue attending the follow-up assessments.

Data management

The data in this trial will be collected using electronic case report forms developed in the data collection system REDCap. The system has been approved by the Danish Data Protection Agency and fulfills the requirements for data security. Data from the interviews will be entered directly into REDCap on a tablet, and all self-report measures will be collected from REDCap. For a detailed overview of outcome measures and data collection time points, see Table 2. The only source data are participants' signed consent forms.

Student assistants employed in the research department but not otherwise involved in the trial will make sure that all self-report measures are sent to participants at the right times, and that the data are complete for all participants enrolled in the trial. The REDCap database has an integrated audit trail to document any access to and changes of the data. The validated data will be exported to SAS for further statistical analyses by statisticians from Copenhagen Trial Unit.

Statistical plan and data analysis

Sample size

The sample size was determined by the predicted change in the primary outcome measure, ZAN-BPD. A 3.5-point superiority margin is considered to be the minimal important difference. Consistent with previous trials that have used ZAN-BPD as an outcome measure for a patient group like ours [24, 58], we expect a standard deviation of 8. With power set at 80% and alpha set at 5% two-tailed, a sample size of 83 participants is needed in each treatment group, corresponding to a total of 166 participants.

Statistical methods

All continuous outcomes will be assessed using linear regression and dichotomous outcomes will be analyzed using logistic regression. The analyses will be based on an intention-to-treat population and will primarily be adjusted for the baseline value of the outcome of interest and the stratification variables used in the randomization. We will secondly adjust all analyses for the following design variables: age (18–30 and 30–60) and functional impairment as assessed with the overall baseline GAF score (0–48 and 49–100) [34].

We will use a five-step procedure [59] to assess if the thresholds for statistical and clinical significance are crossed and we will handle missing data according to the procedure suggested by Jakobsen et al. [60]. A detailed statistical analysis plan will be published before the analyses commence, in which we will provide a detailed description of all primary, secondary, and exploratory analyses. All analyses will be performed blinded with the two intervention groups concealed as A and B.

Interim analyses

An external data safety monitoring committee will perform interim analyses when 50% of the data have been collected according to the good clinical practice guidelines of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use [61]. It will decide whether the trial should stop or carry on. Early stopping criteria will follow the recommendations of Jakobsen et al. [59].

Discussion

This trial will provide evidence of the beneficial and harmful effects of short-term compared to long-term MBT for outpatients with subthreshold or diagnosed borderline personality disorder. To the best of our knowledge, short-term MBT has never been tested before. Gaining more information on how different lengths of treatment work for specific subtypes of patients may help to minimize the potential burden from long-term psychotherapy for some, while at the same time it may identify subtypes of patients for whom short-term psychotherapy is contraindicated. This knowledge may enhance the cost-effectiveness of treatment options for borderline personality disorder. Further, this trial may provide information on the potential predictors and mediators of treatment response.

The present trial has several strengths. First, it has a high degree of external validity because of the relatively inclusive eligibility criteria. Second, the methodology is based on CONSORT, and was predefined and described in detail before randomization began, including, e.g., blinding of all possible parties and implementation of a central randomization system both for generating an allocation sequence and for concealing allocation. Third, the implementation of systematic treatment fidelity ratings allows us to investigate treatment fidelity in both groups.

Our trial also has limitations. First, no systematic review of the effects of short-term compared to long-term psychotherapy for psychiatric disorders is currently available. As mentioned earlier, we are currently performing such a review. Second, the long-term MBT intervention, which is 14 months of treatment in this trial, diverges in intensity (both duration and exposure) from the original 18-month program [13]. This is due to the fixed length of the treatment packages, which have been implemented in the Danish mental health care system. Third, we cannot account for any potential confounding variables because of the structural differences between the groups: the short-term MBT program is closed and conjoined, whereas the long-term program is slow-open and combined.

Dissemination policy

The Danish population will be informed of the trial as well as its final results through national media. The

results of the trial will be presented at all outpatient clinics treating borderline personality disorder in the Mental Health Services, Capital Region of Denmark, by the principal investigator or sponsor-investigator. The final and interim results will be presented at national and international conferences. Further, associations for patients and relatives will be informed about the results of the trial and its future implications. The trial results will be written up by the steering committee and will be published in international peer-reviewed journals. The government of Denmark will be informed of the results before a press release is issued but will have no influence on the reporting of results.

Trial status

The current protocol is version 1, dated 9 October 2018. The first participant was enrolled on 24 September 2018. Recruitment is expected to be completed by September 1, 2020

Additional file

Additional file 1: SPIRIT checklist. (DOCX 61 kb)

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials; DSM-V: Diagnostic and Statistical Manual of Mental Disorders, 5th edition; GAF: Global Assessment of Functioning; GQ: Group Questionnaire; GSI: Global Severity Index; LPFS-BF: Levels of Personality Functioning Scale, Brief Form; MBT: Mentalization-based therapy; MBT-G: Mentalization-based group therapy; MBT-I: Introduction to Mentalization-Based Therapy; MINI: Mini International Neuropsychiatric Interview; MZQ: Mentalization Questionnaire; RQ: Relationship Questionnaire; SCID-5-PD: Structured Clinical Interview for DSM-V Personality Disorders; SCL-90-R: Symptom Checklist 90-R; SF-36: Short-Form Health Survey 36; SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials; SSHI: Suicide and Self-Harm Inventory; WSAS: Work and Social Adjustment Scale; ZAN-BPD: Zangarini Rating Scale for Borderline Personality Disorder; ZAN-BPD-SRV: Zangarini Rating Scale for Borderline Personality Disorder, Self-Report Version

Acknowledgements

We are grateful to the participants and to our colleagues at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre for their hard work and good cooperation.

Funding

The trial is funded by research grants from TrygFonden and from the Mental Health Services Research Foundation, Capital Region of Denmark. Funding applications have undergone anonymous peer review. Neither the funding bodies nor the sponsor will be involved in the collection, analysis, or interpretation of the data, or in writing the manuscripts. The grants will be administered by the head of administration at Stolpegaard Psychotherapy Center. At the end of the trial, the budget will undergo external auditing independently of the trial sponsor and investigators.

Availability of data and materials

The trial investigators, the steering committee, and statisticians at the Copenhagen Trial Unit will have access to the data. After the end of the trial, approval for making the final dataset publicly available in a depersonalized format in the Danish Data Archive will be applied for through the Danish Data Protection Agency.

Trial sponsor

The sponsor of the trial is the Mental Health Services in the Capital Region of Denmark, Kristineberg 3, 2100 Copenhagen Ø.

Protocol amendments

No substantial deviations from the protocol will be implemented without the prior review and approval of the regulatory authorities, except where such may be necessary to eliminate an immediate hazard to the trial participants. In such case, the deviation will be reported to the regulatory authorities as soon as possible. Any deviation, however minor, will be documented and made available in the trial master file in a designated table developed for this purpose.

Authors' contributions

SS and SJ conceived the initial idea of a randomized controlled trial comparing short-term and long-term MBT for borderline personality disorder, along with MS, the acting chief psychologist at the clinic. PS approved of the idea and provided organizational support. SS and SJ developed the trial design with contributions and supervision from JCJ, SL, SP, PS, and AB. SJ wrote the protocol, which was carefully discussed with SS, JCJ, SL, SP, PS, and AB. JCJ designed the plan for the statistical analyses and wrote the sections on sample size and statistical approach. All authors read and approved the final manuscript.

Ethics approval and consent to participate

There are no ethical concerns regarding this trial. Both the experimental intervention and control intervention follow the treatment guidelines published by the National Institute for Health and Care Excellence and the Danish Health and Medicines Authority. Any participants can, at any time, withdraw their consent without any implications for future treatment at the clinic. Prior to commencing the trial, ethical approval was obtained from the regional research ethics committee (ID number H-18023136) and approval was obtained from the Danish Data Protection Agency (I-suite number 6553). During the trial period, data on adverse events will be collected at four time points and will be reported to the regional research ethics committee.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Gentofte, Capital Region of Denmark, Denmark. ²Department of Psychology, University of Copenhagen, Copenhagen, Denmark. ³Copenhagen Trial Unit, Center for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. ⁴St. Ann's Hospital, London, England.

Received: 11 October 2018 Accepted: 19 March 2019

Published online: 05 April 2019

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Arlington: American Psychiatric Pub; 2013.
- Togersen J. In: Oldham J, Skodol AE, Bender DS, editors. Prevalence, Sociodemographics, and Functional Impairment, in Textbook of personality disorders. Washington DC: American Psychiatric Publishing; 2014. p. 109–31.
- Fossati A, et al. Patterns of covariation of DSM-IV personality disorders in a mixed psychiatric sample. *Compr Psychiatry*. 2000;41(3):206–15.
- Zimmerman M, Rothschild L, Chelminski I. The prevalence of DSM-IV personality disorders in psychiatric outpatients. *Am J Psychiatry*. 2005;162(10):1911–8.
- Korzekwa MI, et al. Estimating the prevalence of borderline personality disorder in psychiatric outpatients using a two-phase procedure. *Compr Psychiatry*. 2008;49(4):380–6.
- Conklin CZ, Westen D. Borderline Personality Disorder in Clinical Practice. *Am J Psychiatry*. 2005;162:867–75.
- Zanarini MC, et al. Axis I comorbidity of borderline personality disorder. *Am J Psychiatry*. 1998;155:1733–9.
- Trull TJ, et al. Borderline personality disorder and substance use disorders A review and integration. *Clin Psychol Rev*. 2000;20(2):235–53.
- Barrachina J, et al. Axis II comorbidity in borderline personality disorder is influenced by sex, age, and clinical severity. *Compr Psychiatry*. 2011;52(6):725–30.
- Lieb K, et al. Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. *Br J Psychiatry*. 2010;196(1):4–12.
- Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385(9969):735–43.
- Cristea IA, et al. Efficacy of Psychotherapies for Borderline Personality Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. 2017;74(4):319–28.
- Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355–64.
- Bateman A, Fonagy P. Mentalization-based treatment for personality disorders: a practical guide. Oxford: Oxford University Press; 2016.
- Karterud S. Mentalization-Based Group Therapy (MBT-G): A theoretical, clinical, and research manual. Oxford: OUP Oxford; 2015.
- Vogt KS, Norman P. Is mentalization-based therapy effective in treating the symptoms of borderline personality disorder? A systematic review. *Psychol Psychother*. 2018. <https://doi.org/10.1111/papt.12194>
- Kazdin AE. Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol*. 2007;3:1–27.
- Kvarstein EH, et al. Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—does it make a difference? *Psychol Psychother*. 2015; 88(1):71–86.
- Bales D, et al. Treatment outcome of 18-month, day hospital mentalization-based treatment (MBT) in patients with severe borderline personality disorder in the Netherlands. *J Personal Disord*. 2012;26(4):568–82.
- Jorgensen CR, et al. Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. *Acta Psychiatr Scand*. 2013;127(4):305–17.
- Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563–9.
- Laurensen EM, Luyten P, Kikkert MJ, Westra D, Peen J, Soons MB, Dekker JJ. Day hospital mentalization-based treatment v. specialist treatment as usual in patients with borderline personality disorder: randomized controlled trial. *Psychol Med*. 2018;48(15):2522–29.
- Bateman A, Fonagy P. 8-year follow-up of patients treated for borderline personality disorder: mentalization-based treatment versus treatment as usual. *Am J Psychiatry*. 2008;165(5):631–8.
- Gratz KL, Tull MT, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099–112.
- Blum N, St. John D, Pfohl B, Stuart S, McCormick B, Allen J, Arndt S, Black DW. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for Outpatients With Borderline Personality Disorder: A Randomized Controlled Trial and 1-Year Follow-Up. *Am J Psychiatry*. 2008;165:468–78.
- Bos EH, van Wel EB, Appelo MT, Verbraak MJ. A randomized controlled trial of a Dutch version of systems training for emotional predictability and problem solving for borderline personality disorder. *J Nerv Ment Dis*. 2010; 198(4):299–304.
- McMain SF, et al. A randomized trial of brief dialectical behaviour therapy skills training in suicidal patients suffering from borderline disorder. *Acta Psychiatr Scand*. 2017;135(2):138–48.
- McMain SF. Advances in the Treatment of Borderline Personality Disorder: An Introduction to the Special Issue. *J Clin Psychol*. 2015;71(8):741–6.
- Knekt P, et al. Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*. 2008; 38(5):689–703.
- Lorentzen S, et al. Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial. *Br J Psychiatry*. 2013;203(3):280–7.
- Boutron I, et al. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med*. 2008;148(4):295–309.

32. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med Res Methodol*. 2001;1:2.
33. Chan A, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Ann Intern Med*. 2013;158:200–7.
34. Simonsen S, Heinskou T, Sørensen P, Folke S, Lau ME. Personality disorders: patient characteristics and level of outpatient treatment service. *Nord J Psychiatry*. 2017;71(5):325–31.
35. Clifton A, Pilkonis PA. Evidence for a single latent class of Diagnostic and Statistical Manual of Mental Disorders borderline personality pathology. *Compr Psychiatry*. 2007;48(1):70–8.
36. Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence, and effect of personality disorder. *Lancet*. 2015;385(9969):717–26.
37. Beck E, et al. Mentalization-based treatment in groups for adolescents with borderline personality disorder (BPD) or subthreshold BPD versus treatment as usual (M-GAB): study protocol for a randomized controlled trial. *Trials*. 2016;17(1):314.
38. Zanarini. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Personal Disord*. 2003;17(3):233–42.
39. National Institute for Health and Care Excellence (NICE). Borderline Personality Disorder: the NICE Guideline on Treatment and Management. *Br Psychol Soc R Coll Psychiatrists*. 2009.
40. Karterud S, Bateman A. Manual for mentaliseringsbasert psykoedukativ gruppeterapi (MBT-I). Oslo: Gyldendal akademisk; 2011.
41. Sundhedsstyrelsen, National klinisk retningslinje for behandling af emotionel ustabil personlighedsstruktur, borderline type. 2015.
42. Lecrubier Y, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry*. 1997;12:224–31.
43. First MB, et al. The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part II: Multi-site test-retest reliability study. *J Personal Disord*. 1995;9(2):92–104.
44. Lobbstaël J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clin Psychol Psychother*. 2011;18(1):75–9.
45. Cicchetti DV. Guidelines, Criteria, an Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychol Assess*. 1994;6(4):284–90.
46. Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: Psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personal Ment Health*. 2017;11(4):215–28.
47. Miles HH, Barrabee EL, Finesinger JE. Evaluation of psychotherapy: with a follow up study of 62 cases of anxiety neurosis. *Psychosom Med*. 1951 Mar-Apr;13(2):83–105.
48. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Med Care*. 1992;30(6):473–83.
49. Jones SH, et al. Reliability and Validity of the Global Assessment of Functioning (GAF). *Br J Psychiatry*. 1995;166:654–9.
50. Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: a study in construct validation. *J Clin Psychol*. 1977;33(4):981–9.
51. Hutsebaut J, Feensta D, Kamphuis J. Development and Preliminary Psychometric Evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM-5 Level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). *Personal Disord Theory Res Treat*. 2016;7(2):192–4.
52. Bartholomew K, Horowitz LM. Attachment Styles Among Young Adults: A Test of a Four-Category Model. *J Pers Soc Psychol*. 1991;61(2):226–44.
53. Zanarini, et al. Development of the self-report version of the Zanarini Rating Scale for Borderline Personality Disorder. *Personal Ment Health*. 2015;9(4):243–9.
54. Krogel J, et al. The Group Questionnaire: a clinical and empirically derived measure of group relationship. *Psychother Res*. 2013;23(3):344–54.
55. Hausberg M, Schulz H, Pieglar T, Happach C, Klöpper M, Brütt A, et al. Is a self-rated instrument appropriate to assess mentalization in patients with mental disorders? Development and first validation of the Mentalization Questionnaire (MZQ). *Psychother Res*. 2012;22(6):699–709.
56. Gotzsche PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials*. 1996;17(4):285–90 discussion 290–3.
57. Jarvinen TL, et al. Blinded interpretation of study results can feasibly and effectively diminish interpretation bias. *J Clin Epidemiol*. 2014;67(7):769–72.
58. Priebe S, 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.
59. Jakobsen JC, et al. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol*. 2014;14:34.
60. Jakobsen JC, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.
61. Englev E, Petersen KP. ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives. *Dan Med J*. 2003;165(16):1659–62.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



UPDATE

Open Access



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)

Sophie Juul^{1,2,3*} , Sebastian Simonsen¹, Stig Poulsen², Susanne Lunn², Per Sørensen¹, Anthony Bateman⁴ and Janus Christian Jakobsen^{3,5}

Abstract

Background: Psychotherapy for borderline personality disorder is often extensive and resource-intensive. Mentalisation-based therapy is a psychodynamically oriented treatment option for borderline personality disorder, which includes a case formulation, psychoeducation, and group and individual therapy. The evidence on short-term compared with long-term mentalisation-based therapy is currently unknown.

Methods/design: The Short-Term MBT Project (MBT-RCT) is a single-centre, parallel-group, investigator-initiated, randomised clinical superiority trial in which short-term (20 weeks) will be compared with long-term (14 months) mentalisation-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder. Outcome assessors, data managers, the data safety and monitoring committee, statisticians, and decision-makers will be blinded to treatment allocation. Participants will be assessed before randomisation and at 8, 16, and 24 months after randomisation. The primary outcome will be the severity of borderline symptomatology assessed with the Zanarini Rating Scale for Borderline Personality Disorder. Secondary outcomes will be functional impairment (Work and Social Adjustment Scale), quality of life (Short-Form Health Survey 36—mental component), global functioning (Global Assessment of Functioning), and proportion of participants with severe self-harm. In this paper, we present a detailed statistical analysis plan including a comprehensive explanation of the planned statistical analyses, methods to handle missing data, and assessments of the underlying statistical assumptions. Final statistical analyses will be conducted independently by two statisticians following the present plan.

* Correspondence: sophie.juul@regionh.dk

¹Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Stolpegårdsvej 20, 2820 Gentofte, Denmark

²Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Discussion: We have developed this statistical analysis plan before unblinding of the trial results in line with the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines, which should increase the validity of the MBT-RCT trial by mitigation of analysis bias.

Trial registration: ClinicalTrials.gov [NCT03677037](https://clinicaltrials.gov/ct2/show/study/NCT03677037). Registered on 19 September 2018

Background

Mentalisation-based therapy (MBT) is a long-term, psychodynamically oriented psychotherapy developed specifically to treat patients with borderline personality disorder [1]. Even though more randomised clinical trials at low risk of bias are still needed, long-term MBT is considered one of the most evidence-based interventions currently available for patients with borderline personality disorder [2]. However, the optimal duration of MBT for borderline personality disorder is currently unclear.

The Short-Term MBT Project (MBT-RCT) is a single-centre, parallel-group, investigator-initiated, randomised clinical superiority trial with the objectives to assess the beneficial and harmful effects of short-term (20 weeks) compared with long-term (14 months) MBT for outpatients with subthreshold or diagnosed borderline personality disorder [3]. The Helsinki Declaration [4] and the International Conference on Harmonization of Good Clinical Practice (ICH-GCP) [5] Guidelines recommend that clinical trials should be analysed according to a pre-specified plan to prevent selective outcome reporting bias and data-driven analysis results [6–8].

In this publication, we describe the pre-planned statistical analyses of the primary and secondary outcomes in the MBT-RCT trial. The main publication of the trial results will adhere to this statistical analysis plan as approved by the steering group.

Methods

The design of the MBT-RCT trial has been described in detail previously [3]. The trial population will be adults (18 years of age or older) with subthreshold or diagnosed borderline personality disorder assessed with the Structural Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD). Participants will be eligible for enrolment

if they meet all of the following inclusion criteria and none of the exclusion criteria as presented in Table 1.

The MBT-RCT trial is registered at ClinicalTrials.gov (identifier: NCT03677037), is carried out in compliance with the Declaration of Helsinki [4], and is approved by the Danish Data Protection Agency (approval number: 6553) and by the Regional Research Ethics Committee of the Capital Region of Denmark (approval number: H-18023136).

Randomisation and blinding

Copenhagen Trial Unit, a Danish centre for clinical intervention research, will be responsible for the central randomisation. Randomisation will be performed with a 1:1 allocation according to a computer-generated allocation sequence with permuted blocks of various sizes. The allocation sequence will be concealed from all trial investigators. The randomisation is stratified by (1) sex and (2) high/low baseline scores on the primary outcome measure, the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) [9].

Outcome assessors, data managers, statisticians, the data safety and monitoring committee, and decision-makers will be blinded to treatment allocation [10]. Trial participants and therapists will not be blind to the treatment allocation. This is due to the difficulties of implementing an efficient blinding procedure in trials assessing psychological interventions [10].

Trial interventions

Experimental intervention

The short-term MBT programme is organised as a 20-week programme consisting of 20 weeks of group therapy in closed groups commencing with five sessions of psychoeducative introduction to MBT [11] followed by 15 sessions of group MBT group therapy accompanied

Table 1 Inclusion and exclusion criteria

	General criteria of the outpatient clinic	Criteria exclusive to the trial
Inclusion criteria	- Aged 18–60 - Personality disorder(s) considered to be the primary diagnosis	- A minimum of four confirmed DSM-5 diagnostic criteria for borderline personality disorder - Written informed consent
Exclusion criteria	- Learning disability (IQ < 75) - A full diagnosis of schizotypal personality disorder or antisocial personality disorder - Presence of a comorbid psychiatric disorder that requires specialist treatment - Current (past 2 months) substance dependence including alcohol - Concurrent psychotherapeutic treatment outside the clinic	- Unable to understand Danish - Lack of informed consent

by conjoined individual sessions every second week. All participants are furthermore invited to two psychoeducative meetings with other participants and their relatives. The participants will be treated by two group therapists, one of them also being the individual therapist. Both group and individual therapies are manualised by Bateman and Fonagy [1].

Control intervention

Long-term MBT is organised as a 14-month programme and has been implemented at the clinic for the past 10 years. All participants randomised to long-term MBT will initially enter a 6-week psychoeducative introduction to MBT [11]. New psychoeducative MBT groups commence every time new participants are recruited and randomised to long-term MBT. When the psychoeducative group finishes, participants will be allocated to one of eight slow-open MBT treatment groups. Treatment is then organised as 12 months of weekly group therapy sessions combined with individual therapy every second week. All participants are furthermore invited to two psychoeducative meetings with other participants and their relatives. The participants will be treated by two group therapists and a third individual therapist. Both group and individual therapies are manualised by Bateman and Fonagy [1].

Baseline characteristics

The baseline characteristics will be assessed from inclusion in the trial. A mock table of the complete predefined baseline table can be found in Supplementary Material 1. The baseline characteristics will be as follows:

1. Demographic characteristics:
 - (a) Age
 - (b) Sex
 - (c) Civil status
 - (d) Living situation (alone/with others)
 - (e) Education level
 - (f) Employment status
2. Clinical characteristics:
 - (a) Psychiatric comorbidity, e.g. a diagnosis of an anxiety disorder, major depressive disorder, post-traumatic stress disorder (reported if the frequency is above or equal to 10% in any of the intervention groups)
 - (b) Proportion of participants with subthreshold borderline personality disorder
 - (c) Mean number of borderline personality disorder diagnostic criteria
 - (d) Personality disorder comorbidity (reported if the frequency is above or equal to 10% in any of the intervention groups)

- (e) Proportion of participants with one or more suicide attempts 8 months prior to randomisation
 - (f) Proportion of participants with severe self-harm incidents defined as deliberate acts of self-harm resulting in visible tissue damage 8 months prior to randomisation
 - (g) Proportion of participants on psychopharmacological medication (e.g. antidepressants, antipsychotics) at baseline
3. Mean (SD) number of days from randomisation to assessment time point for all time points (8, 16, and 24 months post-randomisation)

Outcomes

The outcomes were predefined as primary, secondary, and exploratory [3]. This publication describes the statistical analysis plan of the primary and secondary outcomes only.

Primary outcome

- Severity of borderline symptomatology assessed with the Zanarini Rating Scale for Borderline Personality Disorder [9]

Secondary outcomes

- Functional impairment (assessed with the Work and Social Adjustment Scale (WSAS) [12])
- Quality of life (assessed with the Short-Form Health Survey (SF-36) mental component) [13]
- Global functioning (assessed with the Global Assessment of Functioning scale (GAF)) [14]
- Severe self-harm (defined as the proportion of participants with one or more deliberate acts of self-harm resulting in visible tissue damage)

Exploratory outcomes

- Symptom distress (assessed with the Symptom Checklist 90 (SCL-90)) [15]
- Quality of life (assessed with the Short-Form Health Survey (SF-36) physical component [13])

Assessment time points

All outcomes will be assessed at baseline and at 8, 16, and 24 months after randomisation. Investigator-administered outcomes (severity of borderline symptomatology, severe self-harm, and global functioning) will be assessed by blinded assessors at all time points. We will use the 16-month time point as the primary time point of interest, as it is the time point closest to the end of treatment in the long-term MBT group. In an

exploratory analysis, we will consider reporting the results of the comparison between the end of treatment in both groups (i.e. data from the 8 months time point in the short-term group compared with data from the 16 months time point in the long-term group). Data from the 24-month time point, as well as results of the exploratory outcomes, will be analysed using the same principles as described in this statistical analysis plan and published in a separate publication.

Safety

We will report the proportion of participants with one or more serious adverse events in both groups. We will use the International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use—Good Clinical Practice (ICH-GCP) definition of a serious adverse event, which is any untoward medical occurrence that resulted in death, was life-threatening, required hospitalisation or prolonging of existing hospitalisation, and resulted in persistent or significant disability or jeopardised the participant [5]. Two investigators will independently go through the participants' medical journals and assess possible serious adverse events at the 16 and 24 months time point of assessment according to the ICH-GCP definition.

Sample size and power estimations

The sample size estimation was based on the primary outcome, and our primary conclusions will be based on the results of the primary outcome. The outcomes in our outcome hierarchy were ranked according to clinical relevance, and we estimated the power of each non-primary outcome to ensure that we had sufficient power to confirm or reject minimally important intervention effects [16].

Sample size estimation

The sample size was determined by the predicted change in the primary outcome measure, ZAN-BPD. We considered a 3.5-point superiority margin to be the minimal important difference. Consistent with previous trials that have used ZAN-BPD as an outcome measure for a group of participants similar to ours [17, 18] we expect a standard deviation of 8. With power set at 80% and alpha set at 5% two-tailed, a sample size of 83 participants will be needed in each intervention group, corresponding to a total of 166 participants.

We did not adjust the sample size according to missing data. We plan to have as close to 0% missing data as possible. However, if missing data occur, we will use multiple imputations which will limit the loss of power (please see the 'Handling of missing data' section).

Power estimation for secondary outcomes

For the secondary outcomes, we have performed power calculations as presented in Tables 2 and 3 [16].

General analysis principles

Statistical analyses will be performed in Stata [24]. All analyses will be conducted according to the intention-to-treat principle (ITT). The intention-to-treat population will include all randomised participants, regardless of missing data, lost to follow-up, or adherence to the intervention. Thus, by performing an intention-to-treat analysis, we will assess the effects of being randomised to the interventions. We will consider performing a per-protocol analysis, if the number of participants who prematurely drops out of treatment exceeds 5% of the total trial population. By performing a per-protocol analysis, we will assess the effects of adhering to the intervention, which must be considered hypothesis-generating only.

It is generally recommended that regression analyses should be adjusted for the stratification variables used in the randomisation [25–27]. Thus, all analyses will primarily be adjusted for the stratification variables used in the randomisation (and the baseline value of the outcome of interest when assessing continuous outcomes). We will secondly adjust all analyses for the following adjustment variables: age (18–30/31–60), baseline global functioning as assessed with the GAF score (0–48/49–100), baseline proportion of participants with severe self-harm 8 months prior to randomisation (participants with one or more events/participants with no events), and proportion of participants who had their group therapy temporarily paused due to COVID-19 in March 2020 and January 2021 compared to the proportion of participants who did not have their group therapy temporarily paused.

We will perform the following subgroup analyses (test of interaction):

- Baseline severity of borderline symptomatology (ZAN-BPD scores 0–11/12–36)
- Sex (male/female)
- Age (18–30/31–60)
- Baseline global functioning (GAF scores 0–48/49–100)
- Baseline proportion of participants with severe self-harm incidents 8 months prior to randomisation (participants with one or more events/participants with no events)
- Proportion of participants who had their group therapy temporarily paused due to COVID-19 in March 2020 and January 2021 compared to participants who did not have their group therapy temporarily paused

Table 2 Power estimations for the secondary continuous outcomes

Outcome	Minimal clinically important difference	Expected standard deviation	Alpha	Sample size	Power	Reference
	δ	σ	α	n	$1-\beta$	
Functional impairment (WSAS)	4.5	10	5	166	82	Phillips et al. [19]
Quality of life (SF-36—mental component)	5	11	5	166	83	Rollman et al. [20]
Global functioning (GAF)	4	8.5	5	166	85	Jørgensen et al. [21]

GAF Global Assessment of Functioning, SF-36 Short-Form Health Survey 36, WSAS Work and Social Adjustment Scale

We will present the results of the subgroup analyses in forest plots.

Trial profile

The flow of trial participants will be displayed in a Consolidated Standards of Reporting Trials (CONSORT) diagram [28]. The number of screened patients who were assessed for eligibility, and the number included in the primary and secondary analyses, as well as all reasons for exclusions in the primary and secondary analyses, will be reported.

Statistical analyses

Analysis of continuous data

Continuous outcomes will be presented as means and standard deviations for each group together with 95% confidence intervals for the means of the groups and the mean differences between the groups. We will analyse the continuous outcomes using linear regression. All variables will be included as fixed effects.

Analysis of dichotomous data

Dichotomous outcomes will be presented as proportions of participants in each group with the event, together with risk ratios with 95% confidence intervals. We will analyse the dichotomous outcomes using logistic regression. All variables will be included as fixed effects. Odds ratios will be transformed to risk ratios estimating marginal effects using the NLCOM command in Stata [24].

Level of significance

The threshold for significance will be assessed according to a five-step procedure, suggested by Jakobsen and colleagues [29].

The first step will be to calculate and report confidence intervals and p -values for the primary and secondary outcomes. All confidence intervals will be 95% and

two-sided. We will use a p -value of less than 0.05 as the threshold for statistical significance for our primary outcome (see the ‘Sample size estimation’ section) since we plan to report on only one primary outcome. Since our primary conclusions will be based on one outcome result at one time point (16 months post-randomisation), we will limit problems associated with multiple testing due to multiple outcome comparisons [30, 31]. All remaining outcome results and assessment time points will be considered hypothesis-generating only.

The second step will be to calculate and report the Bayes factor [32] for primary and secondary outcomes. The Bayes factor is the ratio between the probability of the results given that the null hypothesis (H_0) is true divided by the probability of the results given that the alternative hypothesis (H_A) is true [32]. Calculating and reporting the Bayes factor will allow us to interpret the results of the primary outcome in relation to former trial results [17, 18].

The third step will be to use Lan-DeMets monitoring boundaries if the trial is stopped before the sample size is reached [33]. This is done to avoid a potential false rejection of the null hypothesis caused by an insufficient sample size [34].

The fourth step regarding adjustment of p -values based on multiple testing of the primary outcome is not applicable to our trial. We only have one single primary outcome, primarily assessed at one time point (16 months post-randomisation) [29].

The fifth step is the assessment of the clinical significance. The assessment of the clinical significance of our trial results will be based on the intervention effects we predefined in the sample size and power estimations.

Interim analyses

We have pre-planned one interim analysis, which will be conducted after half of the trial participants have been

Table 3 Power estimations of secondary dichotomous outcomes

Outcome	Expected proportion in the control group	Alpha significance level	Power	Reference
Severe self-harm incidents ^a	15%	5	14%	Simonsen et al. [22]; Bateman and Fonagy [23]

^aEven though the expected power is only 14%, we define severe self-harm incidents as a secondary outcome, because it is considered very important for this population

assessed at the 8 months post-randomisation time point. The timing and prevalence of any additional interim analyses will be decided exclusively by the members of the data monitoring and safety committee. The role of the data monitoring and safety committee will be to make recommendations to the steering group to either continue, change, hold, or terminate the trial. This recommendation will primarily be based on safety considerations. The data monitoring and safety committee will be provided with the following trial data: number of participants randomised, number of participants per intervention group, baseline ZAN-BPD scores for all participants, ZAN-BPD scores at the 8-month post-randomisation time point for participants in both intervention groups with available data at that time point, proportion of participants with one or more deliberate acts of self-harm at the 8-month post-randomisation time point, and serious adverse events. Based on evaluations of these data, the data monitoring and safety committee will decide whether they want further data from the principal investigator and when next to perform analyses on data.

Handling of missing data

Missing data will be handled according to the recommendations of Jakobsen and colleagues [35]. In short, if we experience missing data, we will consider to use multiple imputations and use best-worst/worst-best case scenarios to assess the potential impact of the missing data [35].

All randomised participants (the intention-to-treat population) will be included in the primary analysis of all outcomes. If it is not valid to ignore missing data (that is, if the missing data exceeds 5%), we will consider using multiple imputations and use best-worst/worst-best case scenarios to assess the potential impact of the missing data [35]. Best-worst and worst-best case scenarios assess the potential range of impact of the missing data for the trial results [35]. In the 'best-worst' case scenario, it is assumed that all participants lost to follow-up in the short-term group have had a beneficial outcome (e.g. had no self-harm incidents), and all those in the long-term group have had a harmful outcome (e.g. had one or more self-harm incidents). Conversely, in the 'worst-best' case scenario, it is assumed that all participants who were lost to follow-up in the short-term group have had a harmful outcome and that all those lost to follow-up in the long-term group have had a beneficial outcome [35]. When continuous outcomes are used, a 'beneficial outcome' will be defined as the group mean plus two SDs of the group mean (fixed imputation), and a 'harmful outcome' will be defined as the group mean minus two SDs of the group mean (fixed imputation) [35]. We do not expect any missing baseline

data, as participants will only be randomised once they have a complete baseline dataset.

Assessments of underlying statistical assumptions

We will assess the underlying statistical assumptions for all statistical analyses [36, 37]. We will test for major interactions between each covariate and the intervention variable for all regression analyses. We will, in turn, include each possible first-order interaction between included covariates and the intervention variable. For each combination, we will test if the interaction term is significant and we will assess the effect size. We will only consider concluding that there is evidence of an interaction if (1) the interaction is statistically significant following the Bonferroni-adjusted thresholds (0.05 divided by the number of possible interactions) and (2) if the interaction shows a clinically significant effect. If we conclude that the interaction is statistically significant, we will consider both presenting a separate analysis for each interaction as well as an overall analysis including the interaction term in the model [36].

Assessments of underlying statistical assumptions for linear regression

We will visually inspect quantile-quantile plots of the residuals [38, 39] to assess if the residuals are normally distributed, and we will use residuals plotted against covariates and fitted values [38, 39] to assess for homogeneity of variances. If the plots show deviations from the model assumptions, we will consider transforming the outcome, i.e. by using log transformation or square root and/or use robust standard errors [38, 39].

Assessments of underlying statistical assumptions for logistic regression

We will assess if the deviance divided by the degrees of freedom is significantly larger than 1 to assess for relevant overdispersion. If that is the case, we will consider using a maximum likelihood estimate of the dispersion parameter.

Statistical reports

Two independent statisticians will analyse blinded data on all outcomes with intervention groups concealed as, e.g., 'A' and 'B'. Two independent statistical reports will be delivered to the principal investigator (SJ) and will be shared with the steering group. If there are discrepancies between the two primary statistical reports, these will be identified and we will then consider which is the most correct result. A final statistical report will be prepared, and all two (or three, if anything is to be corrected) statistical reports will be published as a [supplementary material](#). Mock tables are presented in Supplementary material 1.

Discussion

The primary aim of this paper is to minimise the risks of bias associated with selective outcome reporting and erroneous data-driven results. We therefore present a pre-defined statistical analysis plan for the MBT-RCT trial.

Strengths

Our methodology has several strengths. First, our methodology is pre-defined, and our analyses will adhere to this statistical analysis plan. Second, we have limited problems with multiplicity because we only assess one primary outcome, and our conclusions will primarily be based on the results of the primary outcome [29]. Third, all analyses will be conducted according to the intention-to-treat principle, and, if necessary, we will use multiple imputations and best-worst/worst-best case scenarios to assess the potential impact of missing data [35]. Furthermore, we plan to systematically assess if the underlying statistical assumptions are fulfilled for all statistical analyses.

Limitations

A potential limitation of the MBT-RCT trial is that no systematic review of the effects of short-term compared to long-term psychotherapy for borderline personality disorder, or for psychiatric disorders in general, was available prior to planning of this trial. Hence, estimations of anticipated intervention effects, estimations of variances used in our sample size, and power estimations, etc. may be erroneous. We are currently performing such a review, which will be submitted for publication prior to completion of this trial [40]. Second, we expect a significant amount of missing data, due to the instability of the trial population. Even though we plan to handle missing data appropriately, no statistical method can guarantee the validity of trial results if the missingness is substantial. Third, even though the trial will be sufficiently powered to confirm or reject intervention effects on the primary and secondary outcomes, the relatively small number of randomised participants may result in a risk of baseline differences which also may bias the trial results especially on non-primary outcomes. However, we will carefully consider the low sample size when interpreting the trial results. Fourth, as participants are not blinded to the allocated treatment, results from all participant-reported outcomes are at risk of bias [10]. Fifth, therapists are likewise not blinded to the allocated treatment and may have an allegiance to one of the interventions. Sixth, we have planned several subgroup analyses. However, subgroup analyses are per definition underpowered, and will be considered hypothesis-generating only. We will carefully consider these limitations when interpreting the results.

Conclusion

We have developed this statistical analysis plan in line with the Declaration of Helsinki and the International Conference on Harmonization of Good Clinical Practice Guidelines, which should increase the validity of the MBT-RCT trial by mitigation of analysis bias.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-021-05450-y>.

Additional file 1: Table S1. Baseline characteristics of the trial population. BPD: Borderline Personality Disorder; MINI: Mini International Neuropsychiatric Interview; No.: Number; SCID-5-PD: Structured Clinical Interview for DSM-5 Personality Disorders; SD: Standard Deviation. **Table S2.** Primary and secondary outcome results (ITT Population). GAF: Global Assessment of Functioning; ITT: Intention To Treat; SF-36: Short Form Health Survey – 36; SD: Standard Deviation; WSAS: Work and Social Adjustment Scale; ZAN-BPD: Zanarini Rating Scale for Borderline Personality Disorder.

Acknowledgements

We are grateful to the participants and to our colleagues at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre for their hard work and good cooperation.

Authors' contributions

SJ drafted this protocol update, receiving ongoing supervision from JCJ. All authors amended and approved the final manuscript.

Funding

The MBT-RCT trial is funded by research grants from TrygFonden and from the Mental Health Services Research Foundation, Capital Region of Denmark. Funding applications have undergone anonymous peer review. Neither the funding bodies nor the sponsor will be involved in the collection, analysis, or interpretation of the data, or in writing the manuscripts. The grants will be administered by the head of administration at Stolpegaard Psychotherapy Center. At the end of the trial, the budget will undergo external auditing independently of the trial sponsor and investigators.

Availability of data and materials

Trial investigators, the steering committee, and statisticians at the Copenhagen Trial Unit will have access to the data. After the end of the trial, approval for making the final dataset publicly available in a depersonalised format in the Danish Data Archive will be applied for through the Danish Data Protection Agency.

Declarations

Ethics approval and consent to participate

There are no ethical concerns regarding this trial. Both the experimental intervention and control intervention follow the treatment guidelines published by the National Institute for Health and Care Excellence and the Danish Health and Medicines Authority. Any participants can, at any time, withdraw their consent without any implications for future treatment at the clinic. Prior to commencing the trial, ethical approval was obtained from the regional research ethics committee (ID number H-18023136), and approval was obtained from the Danish Data Protection Agency (I-suite number 6553). During the trial period, data on serious adverse events will be collected and will be reported to the regional research ethics committee.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Stolpegårdsvej 20, 2820 Gentofte, Denmark. ²Department of Psychology, University of Copenhagen, Copenhagen, Denmark. ³Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. ⁴Anna Freud Centre, Kantor Centre of Excellence, London, UK. ⁵Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

Received: 15 January 2021 Accepted: 12 July 2021

Published online: 28 July 2021

References

- Bateman A, Fonagy P. Mentalization-based treatment for personality disorders: a practical guide. Oxford, England: Oxford University Press; 2016. <https://doi.org/10.1093/med:psych/9780199680375.001.0001>.
- Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020;5:CD012955. <https://doi.org/10.1002/14651858.CD012955.pub2>.
- Juil S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196. <https://doi.org/10.1186/s13063-019-3306-7>.
- World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191–4.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Guideline: Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice (ICH-GCP). 2015. Available at: <https://ichgcp.net/da> [Accessed July 9, 2021].
- Dwan K, Gamble C, Williamson PR, Altman DG. Reporting of clinical trials: a review of research funders' guidelines. *Trials*. 2008;9(1):66. <https://doi.org/10.1186/1745-6215-9-66>.
- Thomas L, Peterson ED. The value of statistical analysis plans in observational research: defining high-quality research from the start. *JAMA*. 2012;308(8):773–4. <https://doi.org/10.1001/jama.2012.9502>.
- Gabriel SE. Getting the methods right—the foundation of patient-centered outcomes research. *New Engl J Med*. 2012;367(9):787–90. <https://doi.org/10.1056/NEJMp1207437>.
- Zanarini MC. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Personal Dis*. 2003;17(3):233–42. <https://doi.org/10.1521/pedi.17.3.233.22147>.
- Juil S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med*. 2021;26(3):109. <https://doi.org/10.1136/bmjebm-2020-111407>.
- Karterud S, Bateman A. Manual for mentalization based psychoeducational group therapy (MBT-I). Oslo: Gyldendal Akademisk; 2011.
- Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personal Ment Health*. 2017;11(4):215–28. <https://doi.org/10.1002/pmh.1382>.
- Brazier JE, Harper R, Jones N, O'cathain A, Thomas K, Usherwood T, et al. Validating the SF-36 Health Survey Questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160–4. <https://doi.org/10.1136/bmj.305.6846.160>.
- Hall RC. Global assessment of functioning: a modified scale. *Psychosom*. 1995;36(3):267–75. [https://doi.org/10.1016/S0033-3182\(95\)71666-8](https://doi.org/10.1016/S0033-3182(95)71666-8).
- Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary report. *Psychopharmacol Bull*. 1973;9(1):13–28.
- Jakobsen JC, Ovesen C, Winkel P, Hilden J, Gluud C, Wetterslev J. Power estimations for non-primary outcomes in randomised clinical trials. *BMJ Open*. 2019;9(6):e027092. <https://doi.org/10.1136/bmjopen-2018-027092>.
- Gratz K, Tull M, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099–112. <https://doi.org/10.1017/S0033291713002134>.
- Priebe S, Bhatti N, Barnicot K, Bremner S, Gaglia A, Katsakou C, 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. <https://doi.org/10.1159/000338897>.
- Phillips R, Schneider J, Molosankwe I, Leese M, Foroushani PS, Grime P, et al. Randomized controlled trial of computerized cognitive behavioural therapy for depressive symptoms: effectiveness and costs of a workplace intervention. *Psychol Med*. 2014;44(4):741–52. <https://doi.org/10.1017/S0033291713001323>.
- Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Houck PR, Counihan PJ, et al. Telephone-delivered collaborative care for treating post-CABG depression: a randomized controlled trial. *JAMA*. 2009;302(19):2095–103.
- Jørgensen CR, Freund C, Boye R, Jordet H, Andersen D, Kjølbye M. Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. *Acta Psychiatr Scand*. 2013;127(4):305–17. <https://doi.org/10.1111/j.1600-0447.2012.01923.x>.
- Simonsen S, Heinskou T, Sørensen P, Folke S, Lau M. Personality disorders: patient characteristics and level of outpatient treatment service. *Nordic J Psychiatry*. 2017;71(5):325–31.
- Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355–64. <https://doi.org/10.1176/appi.ajp.2009.09040539>.
- StataCorp. Stata Statistical Software: release 16. College Station, TX: StataCorp LLC; 2019.
- Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Stat Med*. 2011;31:328–40.
- Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ*. 2012;345(sep14 1):e5840. <https://doi.org/10.1136/bmj.e5840>.
- Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. *BMC Med Res Methodol*. 2003;3(1):26. <https://doi.org/10.1186/1471-2288-3-26>.
- Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*. 2010;8(1):18. <https://doi.org/10.1186/1741-7015-8-18>.
- Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol*. 2014;14(34):120–133.
- Zhang J, Quan H, Ng J, Stepanavage ME. Some statistical methods for multiple endpoints in clinical trials. *Control Clin Trials*. 1997;18(3):204–21. [https://doi.org/10.1016/S0197-2456\(96\)00129-8](https://doi.org/10.1016/S0197-2456(96)00129-8).
- Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567(7748):305–7. <https://doi.org/10.1038/d41586-019-00857-9>.
- Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. *Clin Trials J Soc Clin Trials*. 2005;2(4):282–90. <https://doi.org/10.1191/1740774505cn0980a>.
- Demets DL, Lan KG. Interim analysis: the alpha spending function approach. *Stat Med*. 1994;13(13-14):1341–52. <https://doi.org/10.1002/sim.4780131308>.
- Guyatt GH, Briel M, Glasziou P, Bassler D, Montori VM. Problems of stopping trials early. *BMJ*. 2012;344(jun15 1):e3863. <https://doi.org/10.1136/bmj.e3863>.
- Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162. <https://doi.org/10.1186/s12874-017-0442-1>.
- Nielsen EE, Nørskov AK, Lange T, Thabane L, Wetterslev J, Beyersmann J, et al. Assessing assumptions for statistical analyses in randomised clinical trials. *BMJ Evid Based Med*. 2019;24:185–9.
- Nørskov AK, Lange T, Nielsen EE, Gluud C, Winkel P, Beyersmann J, et al. Assessment of assumptions of statistical analysis methods in randomised clinical trials: the what and how. *BMJ Evid Based Med*. 2020;26(3):121–6.
- Frydenberg SJM. An introduction to Stata for health researchers, Fourth Edition: Stata Press; 2014.
- Oliveira AG. Biostatistics decoded: Wiley; 2013.
- Juil S, Poulsen S, Lunn S, Sørensen P, Jakobsen JC, Simonsen S. Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Syst Rev*. 2019;8(1):169. <https://doi.org/10.1186/s13643-019-1099-0>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Differences between the protocol/statistical analysis plan and the manuscript

Covid-19:

Due to the COVID-19 pandemic, group therapy sessions were temporarily paused from March 13th to May 11th, 2020, and from January 11th to March 15th, 2021, in both groups. Individual therapy continued every second week during this time according to plan, mainly face to face, but online consultations were a possibility. These alterations did not result in any change of the length of treatment in either group. E.g. if a participant was randomized to short-term MBT, the original fixed date that was set for termination were kept, so that the participant still received 20 weeks of psychotherapy. The same principle applied in the long-term group. The psychoeducative meetings with the patient and relatives were canceled during this period.

Abstracts

When the abstracts were approved by the steering committee, we did not have data for the individual serious adverse events. Hence the abstracts were slightly revised after blinding was broken to include information about serious adverse events. The analyses have been reported in the post-hoc statistical report in this supplementary file.

	Short-term MBT			Long-term MBT		
	Baseline	8 months	16 months	Baseline	8 months	16 months
	N = 84	N = 73	N = 71	N = 82	N = 74	N = 58
Proportion of participants on antidepressants	33 (39.3%)	26 (35.6%)	25 (35.2%)	35 (42.7%)	29 (39.2%)	19 (32.8%)
Proportion of participants on antipsychotics	10 (11.9%)	16 (21.9%)	15 (21.1%)	11 (13.3%)	11 (14.9%)	10 (17.2%)
Proportion of participants on other psychoactive drugs	21 (25.0%)	23 (31.5%)	21 (29.6%)	19 (23.2%)	25 (33.8%)	21 (36.2%)

ATTENDANCE DATA

Mean (SD) number of attended individual- and group therapy sessions in the intention-to-treat and per protocol population

Intention-to-treat population					
	n	Mean, individual therapy	SD, individual therapy	Mean, group therapy	SD, group therapy
Short-term MBT	71	12	6,2	13,9	6,9
Long-term MBT	74	16,6	10,7	23,8	13,6

Per protocol population					
	n	Mean, individual therapy	SD, individual therapy	Mean, group therapy	SD, group therapy
Short-term MBT	52	14	4,4	16,3	5,5
Long-term MBT	47	20,3	8,9	31	9,7

MBT RCT

Markus Harboe Olsen & Janus Christian Jakobsen

Markus Harboe Olsen's (MHO) first analysis was carried out 22:26, 6 November 2022, and included correction of the baseline occurrence for the self-harm analysis – this was corrected in the second and attached report from 21:05, 13 November 2022. Janus Christian Jakobsen's (JCJ) first analysis was carried out at 07:48, 7 November 2022 – analysis of global functioning was not included. This was corrected in the second and attached report from 21:11, 10 November 2022.

The results from the analyses carried out in R 4.2.1 (by MHO) and in Stata 17 (by JCJ) are not identical. However, the overall conclusions and all primary analyses are identical with subsequent identical results.

Overall, when assessing the assumptions for the linear regression analyses, the analyses did not fit perfectly and both MHO and JCJ explored the results using non-parametric tests which showed similar patterns.

The complete case analysis for the primary outcome was by both MHO and JCJ chosen as the primary analysis. MHO carried out different sensitivity – all mentioned in as a possibility in the statistical analysis plan – which showed similar results.

Interaction analysis did not show any significant interactions, and only MHO carried out the subsequent subgroup analysis. It was not necessary – however, the results can be used to present the subgroup analysis in a forest plot following the statistical analysis plan.

After presentation to the investigators, additional exploratory preplanned analyses were carried out using month 8 as an outcome. A third and final report created at 21:33, 28 November 2022 is attached as the final report.

Appendices:

1. The final report in R – 21:33, 28 November 2022
2. The report in R – 21:05, 13 November 2022
3. The report in Stata – 21:11, 10 November 2022

Statistical report for the MBT-RCT

Markus Harboe Olsen

28 november 2022 21:33

Table of contents

Primary outcome	3
Severity of borderline symptomatology (ZAN-BPD, 16-weeks)	3
Complete case analysis (primary analysis)	3
Multiple imputation (exploratory analysis)	5
Best-worst / worst-best (exploratory analysis)	6
8 vs. 16 weeks / 16 vs. 8 weeks (exploratory analysis)	7
Stratification variables (exploratory analysis)	8
Secondary outcomes	12
Functional impairment (WSAS, 16-weeks)	12
Quality of life (SF-36 mental component, 16-weeks)	14
Global functioning (GAF, 16-weeks)	16
Severe self-harm (16-weeks)	18
Serious adverse events (16-weeks)	18
Subgroup analysis	20
Baseline ZAN-BPD scores (0-11/12-36)	20
Sex (male/female)	21
Age (18-30/31-60)	22
Baseline GAF scores (0-48/49-100)	23
Baseline selfharm	24
Paused treatment due to COVID-19	25
Forest plot	25
Exploratory 8-week analyses	26
Severity of borderline symptomatology (ZAN-BPD, 8-weeks)	26
Functional impairment (WSAS, 8-weeks)	29

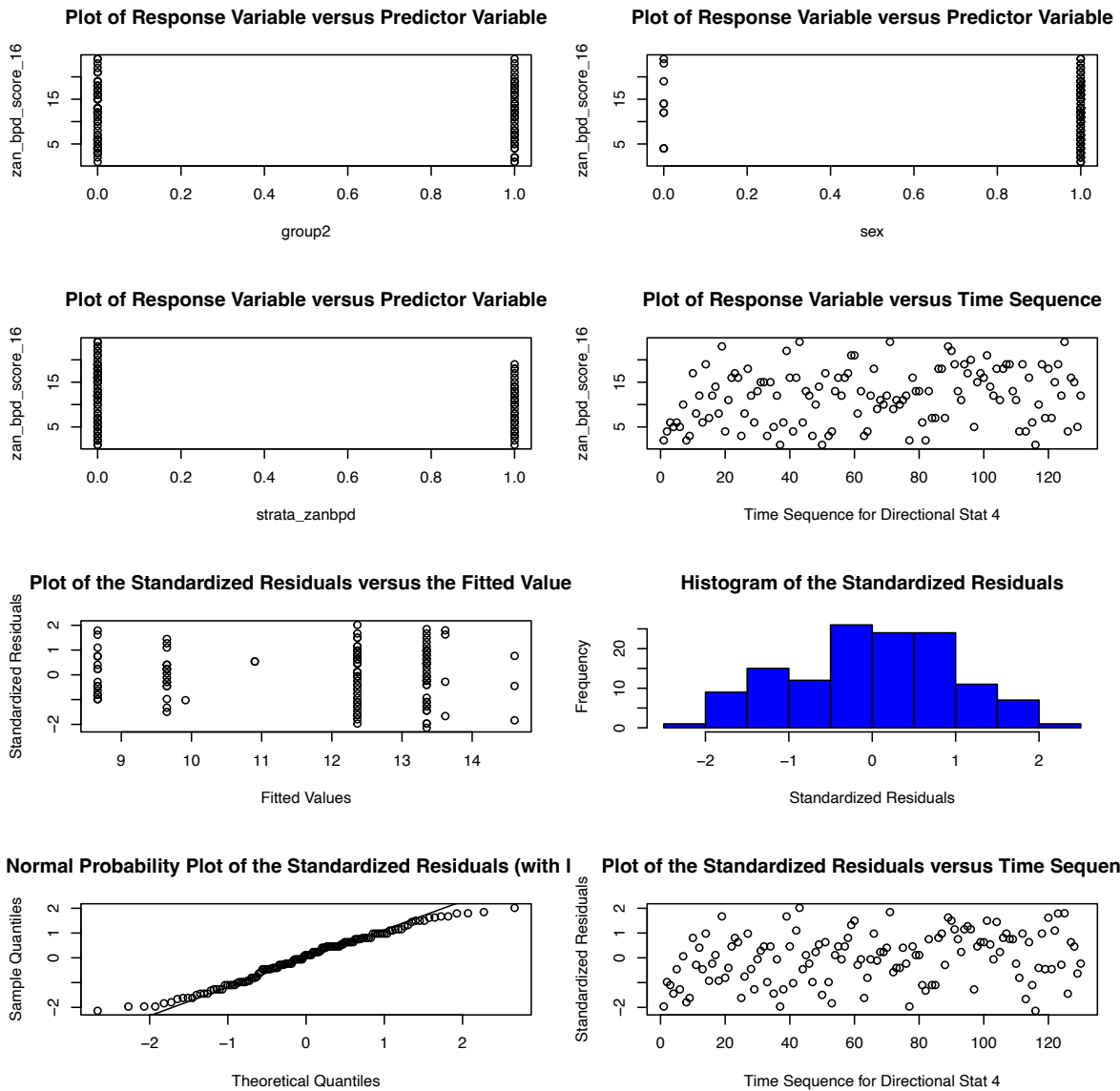
Quality of life (SF-36 mental component, 8-weeks)	31
Global functioning (GAF, 8-weeks)	33
Severe self-harm (8-weeks)	35

Primary outcome

Severity of borderline symptomatology (ZAN-BPD, 16-weeks)

Complete case analysis (primary analysis)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	4.0541883	0.3987219	Assumptions acceptable.
Skewness	0.4777362	0.4894495	Assumptions acceptable.
Kurtosis	3.4208026	0.0643797	Assumptions acceptable.
Link Function	0.1556467	0.6931965	Assumptions acceptable.
Heteroscedasticity	0.0000027	0.9986988	Assumptions acceptable.

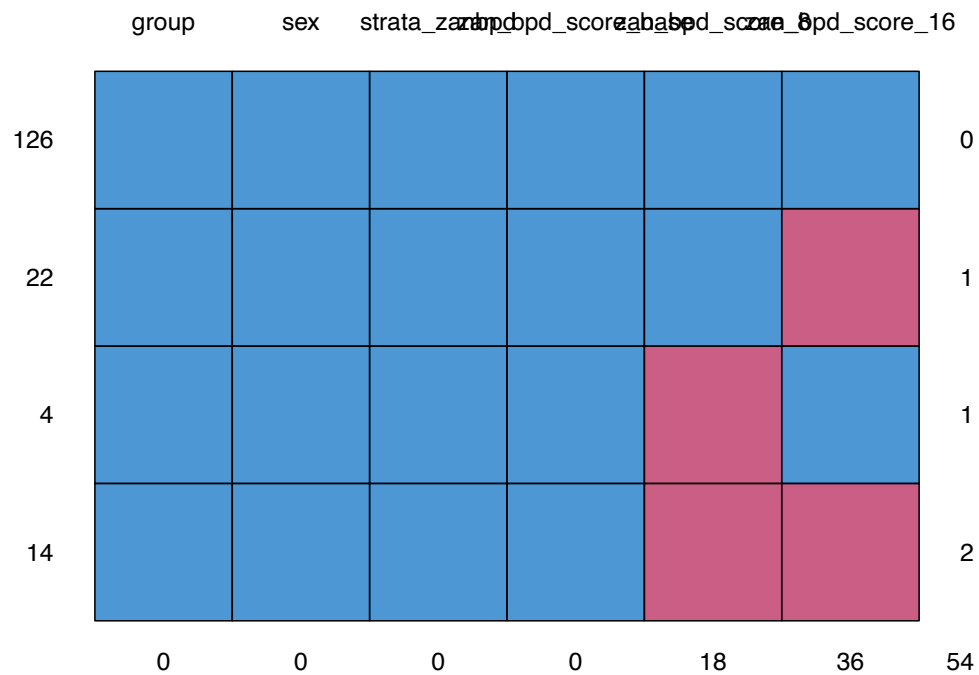
Summary data

	1	2
	(N=84)	(N=82)
zan_bpd_score_base		
Mean (SD)	13.9 (3.72)	13.8 (3.75)
zan_bpd_score_8		
Mean (SD)	12.0 (5.33)	12.8 (5.01)
Missing	10 (11.9%)	8 (9.8%)
zan_bpd_score_16		
Mean (SD)	11.3 (6.22)	12.3 (5.91)
Missing	14 (16.7%)	22 (26.8%)

Analysis

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.62	9.75 – 17.48	<0.001
group [2]	0.99	-1.06 – 3.03	0.341
sex	-1.25	-5.07 – 2.57	0.518
strata zanbpd	-3.70	-5.91 – -1.50	0.001
Observations	130		
R ² / R ² adjusted	0.089 / 0.067		

Multiple imputation (exploratory analysis)



Complete data summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.00	6.00	12.00	11.78	16.00	24.00	36

Imputed data summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	6.00	12.00	11.57	16.00	24.00

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	12.28	8.89 – 15.68	< 0.001
group [2]	1.53	-0.26 – 3.33	0.094
sex	-0.45	-3.80 – 2.89	0.789
strata zanbpd	-3.58	-5.55 – -1.60	< 0.001
Observations	166		
R ² / R ² adjusted	0.087 / 0.070		

Best-worst / worst-best (exploratory analysis)

Best 1 / Worst 2

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	11.64	8.28 – 15.00	< 0.001
group [2]	3.62	1.85 – 5.40	< 0.001
sex	-0.54	-3.86 – 2.77	0.746
strata zanbpd	-2.93	-4.88 – -0.98	0.003
Observations	166		
R ² / R ² adjusted	0.131 / 0.115		

Best 2 / Worst 1

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	15.12	11.78 – 18.46	< 0.001
group [2]	-1.60	-3.37 – 0.17	0.075
sex	-2.03	-5.32 – 1.26	0.226
strata zanbpd	-3.10	-5.04 – -1.16	0.002
Observations	166		
R ² / R ² adjusted	0.083 / 0.066		

8 vs. 16 weeks / 16 vs. 8 weeks (exploratory analysis)

1 is 8 weeks / 2 is 16 weeks

zan_816			
Predictors	Estimates	CI	p
(Intercept)	13.75	10.42 – 17.08	< 0.001
group [2]	0.50	-1.33 – 2.34	0.588
sex	-0.79	-4.12 – 2.53	0.637
strata zanbpd	-3.93	-5.94 – -1.92	< 0.001
Observations	134		
R ² / R ² adjusted	0.106 / 0.085		

2 is 8 weeks / 1 is 16 weeks

zan_816			
Predictors	Estimates	CI	p
(Intercept)	11.95	8.64 – 15.27	< 0.001
group [2]	1.42	-0.30 – 3.14	0.104
sex	0.88	-2.36 – 4.12	0.592
strata zanbpd	-4.76	-6.64 – -2.88	< 0.001
Observations	144		
R ² / R ² adjusted	0.168 / 0.151		

Stratification variables (exploratory analysis)

Age (18-30/31-60)

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.92	9.87 – 17.96	< 0.001
group [2]	1.04	-1.02 – 3.10	0.320
sex	-1.36	-5.21 – 2.49	0.486
strata zanbpd	-3.69	-5.91 – -1.48	0.001
covar ageTRUE	-0.55	-2.64 – 1.55	0.607
Observations	130		
R ² / R ² adjusted	0.091 / 0.062		

GAF (0-48/49-100)

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	14.13	10.15 – 18.10	< 0.001
group [2]	0.93	-1.12 – 2.97	0.372
sex	-1.38	-5.20 – 2.45	0.478
strata zanbpd	-3.33	-5.64 – -1.02	0.005
covar gafTRUE	-1.18	-3.36 – 0.99	0.283
Observations	130		
R ² / R ² adjusted	0.097 / 0.068		

Selfharm ($\geq 1/0$)

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.21	9.24 – 17.17	<0.001
group [2]	0.97	-1.07 – 3.01	0.350
sex	-1.26	-5.09 – 2.56	0.515
strata zanbpd	-3.62	-5.83 – -1.41	0.002
self harm baseline [1]	0.99	-1.09 – 3.07	0.348
Observations	130		
R ² / R ² adjusted	0.095 / 0.066		

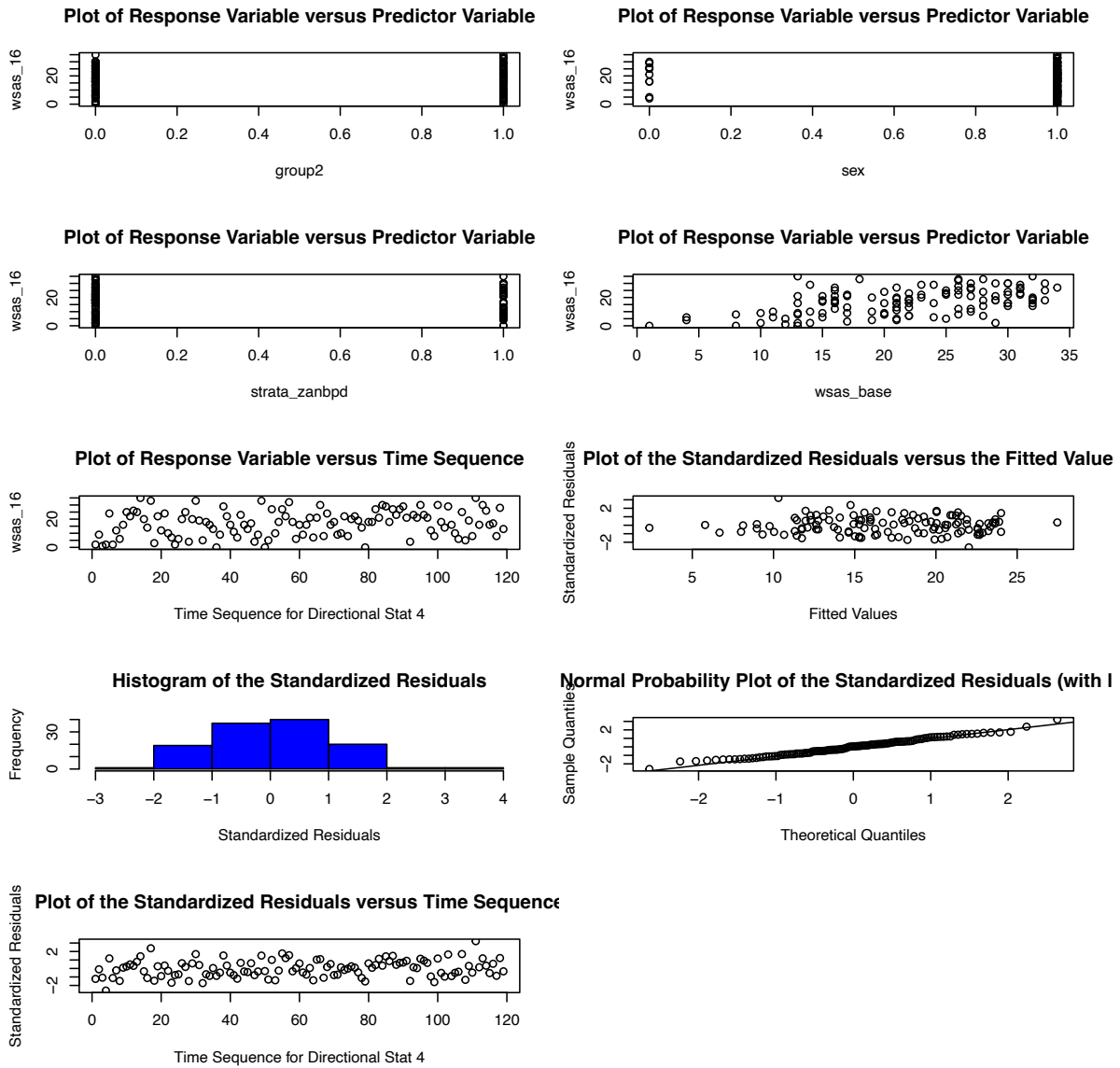
Paused treatment

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	14.29	10.34 – 18.24	< 0.001
group [2]	0.29	-1.95 – 2.53	0.798
sex	-0.75	-4.62 – 3.11	0.701
strata zanbpd	-3.67	-5.87 – -1.47	0.001
covid19 [Nej]	-1.68	-3.94 – 0.59	0.146
Observations	130		
R ² / R ² adjusted	0.104 / 0.076		

Secondary outcomes

Functional impairment (WSAS, 16-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	1.7158018	0.7878452	Assumptions acceptable.
Skewness	1.1646766	0.2804978	Assumptions acceptable.
Kurtosis	0.0103978	0.9187807	Assumptions acceptable.
Link Function	0.4868232	0.4853478	Assumptions acceptable.
Heteroscedasticity	0.0539042	0.8164038	Assumptions acceptable.

Summary data

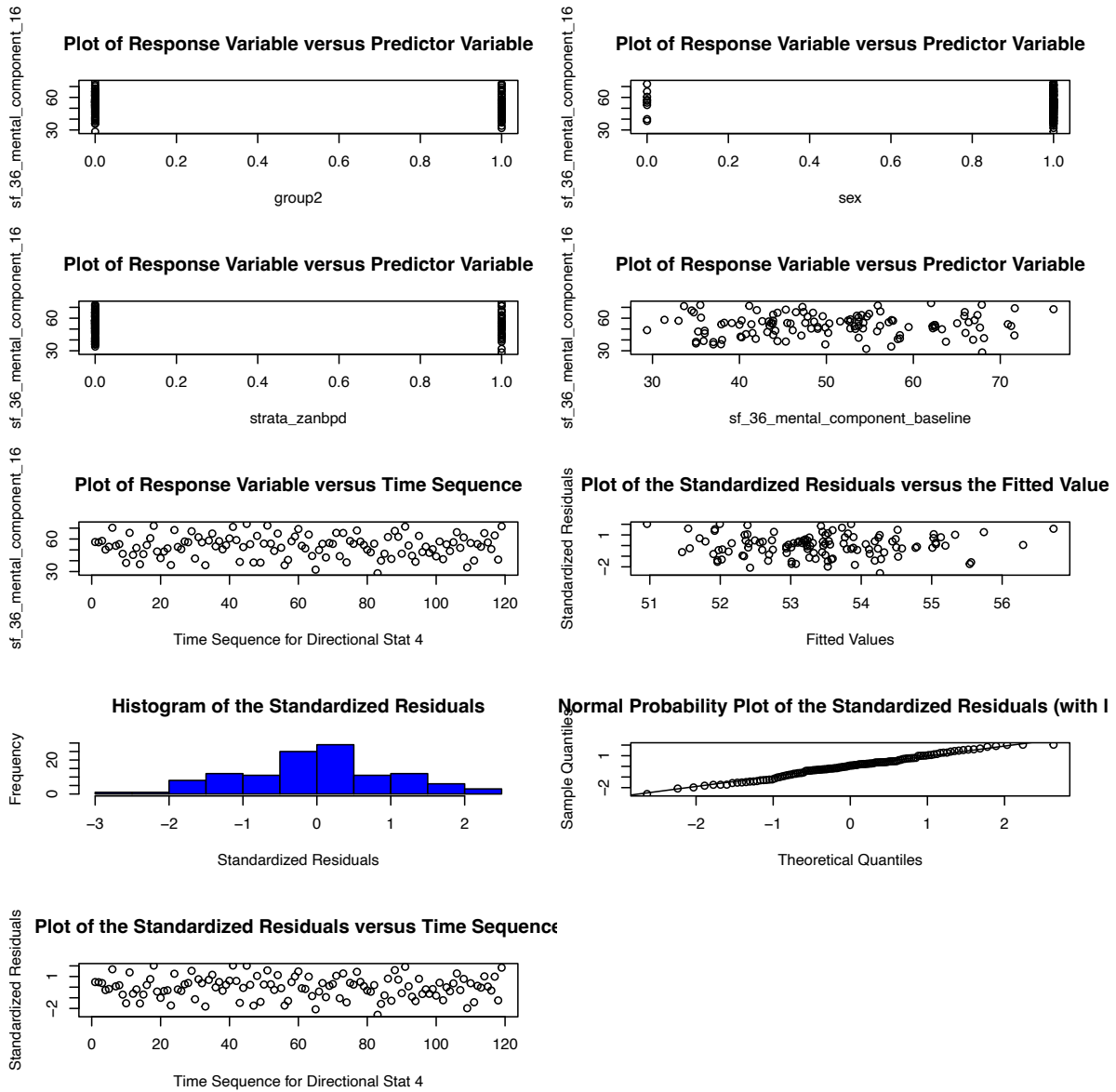
	1	2
	(N=84)	(N=82)
wsas_base		
Mean (SD)	22.9 (7.54)	21.9 (7.46)
wsas_8		
Mean (SD)	20.4 (8.59)	19.9 (8.36)
Missing	13 (15.5%)	13 (15.9%)
wsas_16		
Mean (SD)	16.7 (8.91)	17.2 (9.36)
Missing	21 (25.0%)	26 (31.7%)

Analysis

	wsas_16		
Predictors	Estimates	CI	p
(Intercept)	5.63	-1.33 – 12.58	0.112
group [2]	1.44	-1.43 – 4.32	0.321
sex	-4.24	-9.67 – 1.18	0.124
strata zanbpd	0.31	-2.82 – 3.45	0.844
wsas base	0.66	0.46 – 0.86	<0.001
Observations	119		
R ² / R ² adjusted	0.281 / 0.256		

Quality of life (SF-36 mental component, 16-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	6.4455556	0.1682533	Assumptions acceptable.
Skewness	0.2764293	0.5990512	Assumptions acceptable.
Kurtosis	0.8714568	0.3505523	Assumptions acceptable.
Link Function	5.2434878	0.0220291	Assumptions NOT satisfied!
Heteroscedasticity	0.0541817	0.8159404	Assumptions acceptable.

Summary data

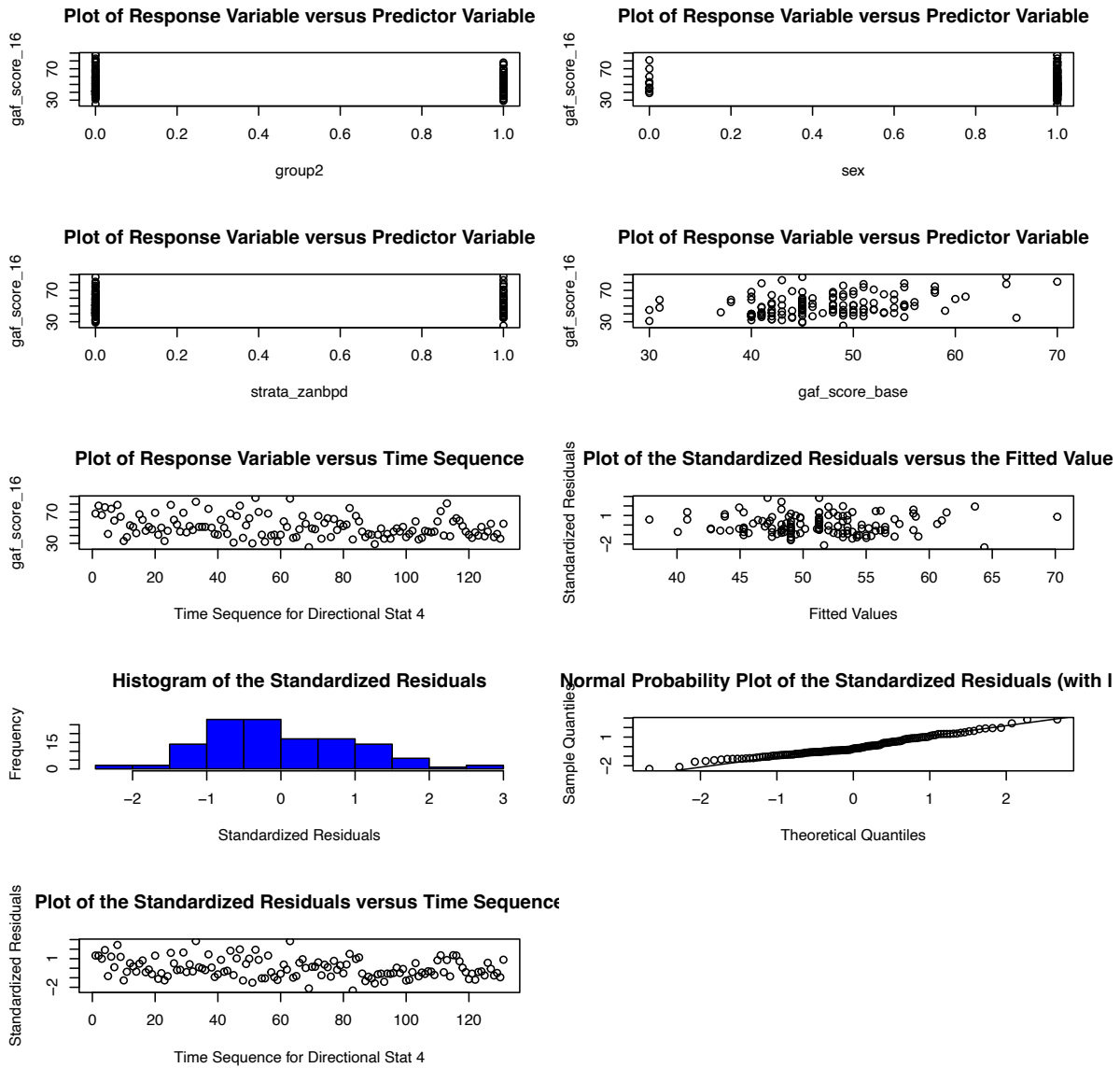
	1	2
	(N=84)	(N=82)
sf_36_mental_component_baseline		
Mean (SD)	50.3 (10.4)	49.2 (10.1)
sf_36_mental_component_8		
Mean (SD)	50.4 (10.8)	53.1 (11.6)
Missing	13 (15.5%)	13 (15.9%)
sf_36_mental_component_16		
Mean (SD)	53.9 (9.94)	52.9 (10.2)
Missing	21 (25.0%)	26 (31.7%)

Analysis

sf_36_mental_component_16			
Predictors	Estimates	CI	p
(Intercept)	52.91	41.94 – 63.88	<0.001
group [2]	-0.91	-4.62 – 2.79	0.626
sex	-2.48	-9.46 – 4.50	0.484
strata zanbpd	-0.90	-5.03 – 3.23	0.666
sf 36 mental component baseline	0.07	-0.11 – 0.25	0.439
Observations	119		
R ² / R ² adjusted	0.012 / -0.023		

Global functioning (GAF, 16-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	10.3506099	0.0349182	Assumptions NOT satisfied!
Skewness	5.8210651	0.0158354	Assumptions NOT satisfied!
Kurtosis	0.0169192	0.8965081	Assumptions acceptable.
Link Function	1.4496393	0.2285858	Assumptions acceptable.
Heteroscedasticity	3.0629862	0.0800943	Assumptions acceptable.

Summary data

	1	2
	(N=84)	(N=82)
gaf_score_base		
Mean (SD)	46.5 (8.14)	46.3 (6.36)
gaf_score_8		
Mean (SD)	49.6 (13.5)	47.7 (10.5)
Missing	10 (11.9%)	8 (9.8%)
gaf_score_16		
Mean (SD)	52.2 (15.0)	49.7 (11.7)
Missing	14 (16.7%)	21 (25.6%)

Analysis

gaf_score_16			
Predictors	Estimates	CI	p
(Intercept)	17.82	-0.77 – 36.42	0.060
group [2]	-2.25	-6.70 – 2.20	0.318
sex	-0.19	-8.59 – 8.21	0.964
strata zanbpd	-2.60	-7.69 – 2.49	0.314
gaf score base	0.75	0.40 – 1.09	<0.001
Observations	131		
R ² / R ² adjusted	0.139 / 0.111		

Wilcoxon rank sum test with continuity correction

```
data: df$gaf_score_16[df$group == 2] and df$gaf_score_16[df$group == 1]
W = 2014, p-value = 0.578
alternative hypothesis: true location shift is not equal to 0
```

Severe self-harm (16-weeks)

Summary data

	1	2
	(N=84)	(N=82)
self_harm_baseline		
0	52 (61.9%)	49 (59.8%)
1	32 (38.1%)	33 (40.2%)
self_harm_8		
0	55 (65.5%)	47 (57.3%)
1	19 (22.6%)	27 (32.9%)
Missing	10 (11.9%)	8 (9.8%)
self_harm_16		
0	55 (65.5%)	45 (54.9%)
1	13 (15.5%)	16 (19.5%)
Missing	16 (19.0%)	21 (25.6%)

Analysis

Relative risk derived by G-computation
RR: 1.3717 (sRR: 1.3754)
Confidence interval: 0.6983 - 2.8412
p-value: 0.3353

Serious adverse events (16-weeks)

Summary data

	1	2
	(N=84)	(N=82)
as.factor(sae_total)		
0	69 (82.1%)	58 (70.7%)
1	9 (10.7%)	15 (18.3%)
2	4 (4.8%)	3 (3.7%)
3	1 (1.2%)	3 (3.7%)
4	0 (0%)	1 (1.2%)
5	1 (1.2%)	1 (1.2%)
7	0 (0%)	1 (1.2%)

Analysis

Fisher's Exact Test for Count Data

```
data: table(df$sae_total > 0, df$group)
p-value = 0.1003
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.8625312 4.2797517
sample estimates:
odds ratio
 1.896076
```

Wilcoxon rank sum test with continuity correction

```
data: df$sae_total[df$group == "1"] and df$sae_total[df$group == "2"]
W = 3046.5, p-value = 0.08356
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -3.403897e-05 5.726355e-05
sample estimates:
difference in location
 -1.998471e-06
```

Subgroup analysis

Baseline ZAN-BPD scores (0-11/12-36)

P-value for interaction: 0.6904671

Baseline ZAN-BPD scores 0-11 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	9.60	3.40 – 15.79	0.003
group [2]	1.61	-1.60 – 4.82	0.317
sex	-1.23	-7.31 – 4.86	0.685
Observations	40		
R ² / R ² adjusted	0.034 / -0.018		

Baseline ZAN-BPD scores 12-36 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.70	8.86 – 18.53	<0.001
group [2]	0.71	-1.92 – 3.35	0.593
sex	-1.20	-6.10 – 3.70	0.627
Observations	90		
R ² / R ² adjusted	0.006 / -0.017		

:::

...

Sex (male/female)

P-value for interaction: 0.7697972

Sex 0 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	14.06	4.87 – 23.25	0.009
group [2]	-0.14	-12.84 – 12.56	0.980
strata zanbpd	-3.30	-17.15 – 10.55	0.591
Observations	10		
R ² / R ² adjusted	0.046 / -0.226		

Sex 1 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	12.33	10.76 – 13.90	<0.001
group [2]	1.08	-1.01 – 3.16	0.308
strata zanbpd	-3.71	-5.96 – -1.47	0.001
Observations	120		
R ² / R ² adjusted	0.091 / 0.076		

...

...

Age (18-30/31-60)

P-value for interaction: 0.6602134

Age 18-30 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.81	7.98 – 19.65	< 0.001
group [2]	1.43	-1.26 – 4.12	0.294
sex	-1.43	-7.35 – 4.50	0.632
strata zanbpd	-3.68	-6.57 – -0.80	0.013
Observations	76		
R ² / R ² adjusted	0.105 / 0.068		

Age 31-60 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.70	8.06 – 19.34	< 0.001
group [2]	0.48	-2.99 – 3.94	0.784
sex	-1.45	-6.88 – 3.97	0.593
strata zanbpd	-3.55	-7.26 – 0.16	0.060
Observations	54		
R ² / R ² adjusted	0.074 / 0.018		

...

...

Baseline GAF scores (0-48/49-100)

P-value for interaction: 0.6055547

Baseline GAF scores 0-48 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	15.36	9.63 – 21.10	< 0.001
group [2]	1.35	-1.38 – 4.08	0.329
sex	-2.79	-8.37 – 2.79	0.323
strata zanbpd	-3.93	-7.40 – -0.46	0.027
Observations	77		
R ² / R ² adjusted	0.081 / 0.043		

Baseline GAF scores 49-100 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	11.91	6.51 – 17.30	< 0.001
group [2]	0.25	-2.97 – 3.48	0.874
sex	-0.06	-5.61 – 5.49	0.983
strata zanbpd	-3.04	-6.29 – 0.21	0.066
Observations	53		
R ² / R ² adjusted	0.070 / 0.013		

...

...

Baseline selfharm

P-value for interaction: 0.1563246

Baseline no selfharm ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.87	8.62 – 19.12	< 0.001
group [2]	-0.25	-2.98 – 2.47	0.853
sex	-1.43	-6.52 – 3.66	0.578
strata zanbpd	-3.49	-6.38 – -0.60	0.018
Observations	77		
R ² / R ² adjusted	0.075 / 0.037		

Baseline selfharm ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.13	7.13 – 19.12	< 0.001
group [2]	2.74	-0.44 – 5.93	0.090
sex	-0.93	-7.03 – 5.18	0.762
strata zanbpd	-3.93	-7.59 – -0.27	0.036
Observations	53		
R ² / R ² adjusted	0.139 / 0.086		

...

...

Paused treatment due to COVID-19

P-value for interaction: 0.3758016

Baseline no pause ::: {.cell-output-display}

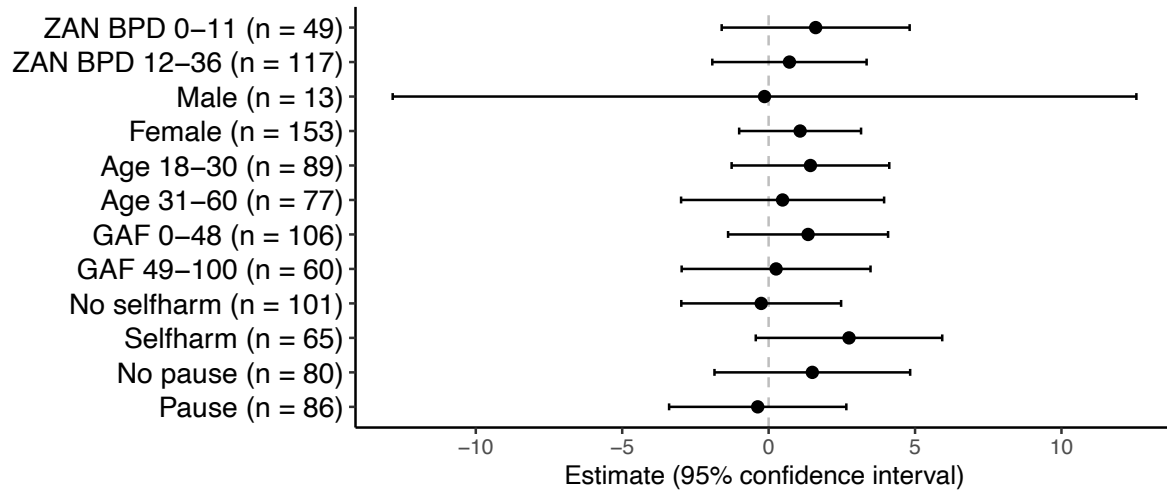
zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	6.92	-1.26 – 15.09	0.096
group [2]	1.49	-1.85 – 4.83	0.375
sex	5.52	-2.72 – 13.75	0.185
strata zanbpd	-5.83	-8.94 – -2.72	< 0.001
Observations	64		
R ² / R ² adjusted	0.222 / 0.183		

Baseline pause ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	15.87	11.28 – 20.46	< 0.001
group [2]	-0.37	-3.40 – 2.65	0.806
sex	-2.78	-7.13 – 1.58	0.207
strata zanbpd	-1.55	-4.65 – 1.56	0.323
Observations	66		
R ² / R ² adjusted	0.043 / -0.003		

... ..

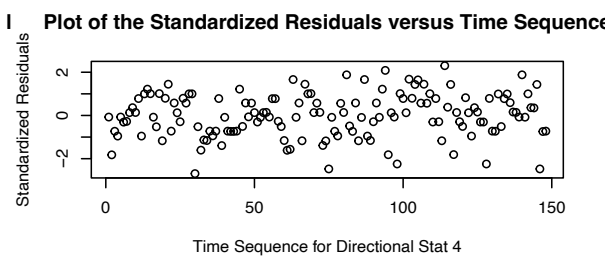
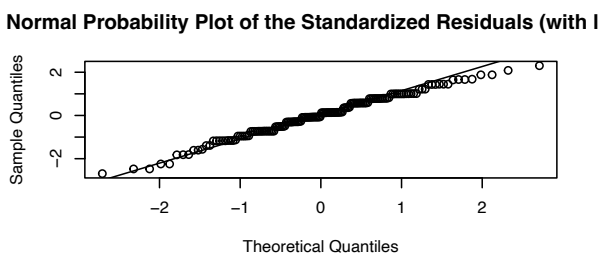
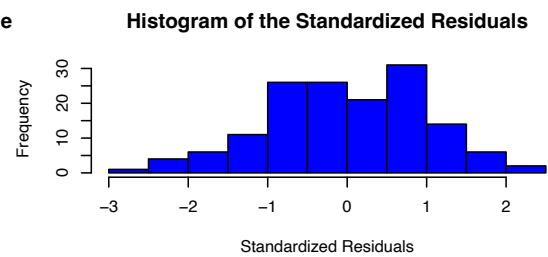
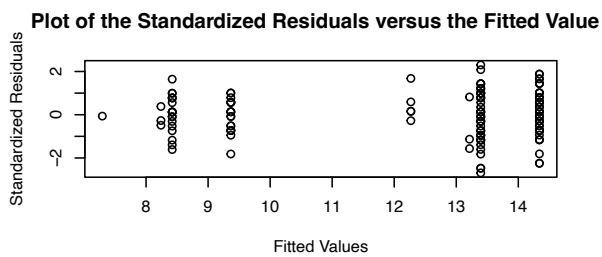
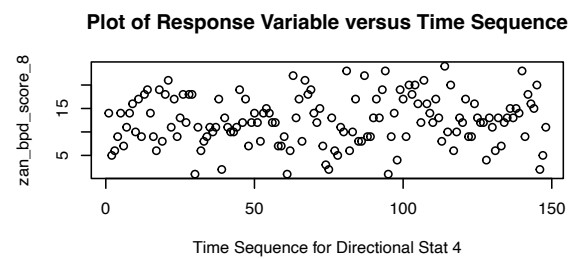
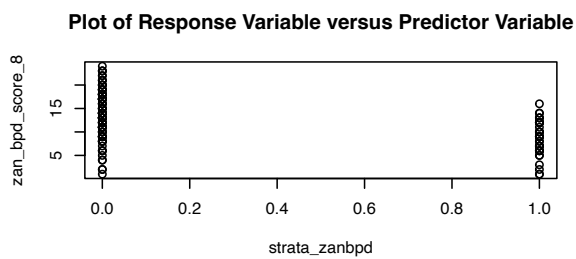
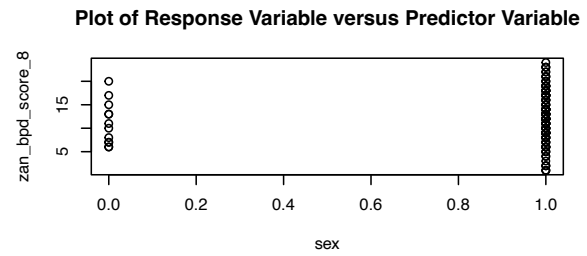
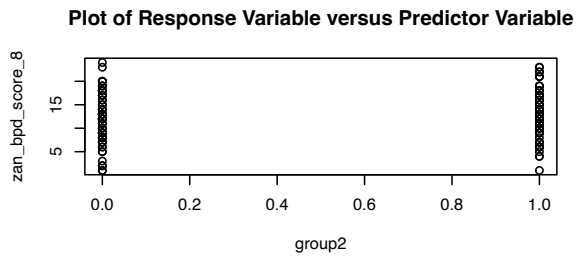
Forest plot



Exploratory 8-week analyses

Severity of borderline symptomatology (ZAN-BPD, 8-weeks)

Assumptions for linear regression



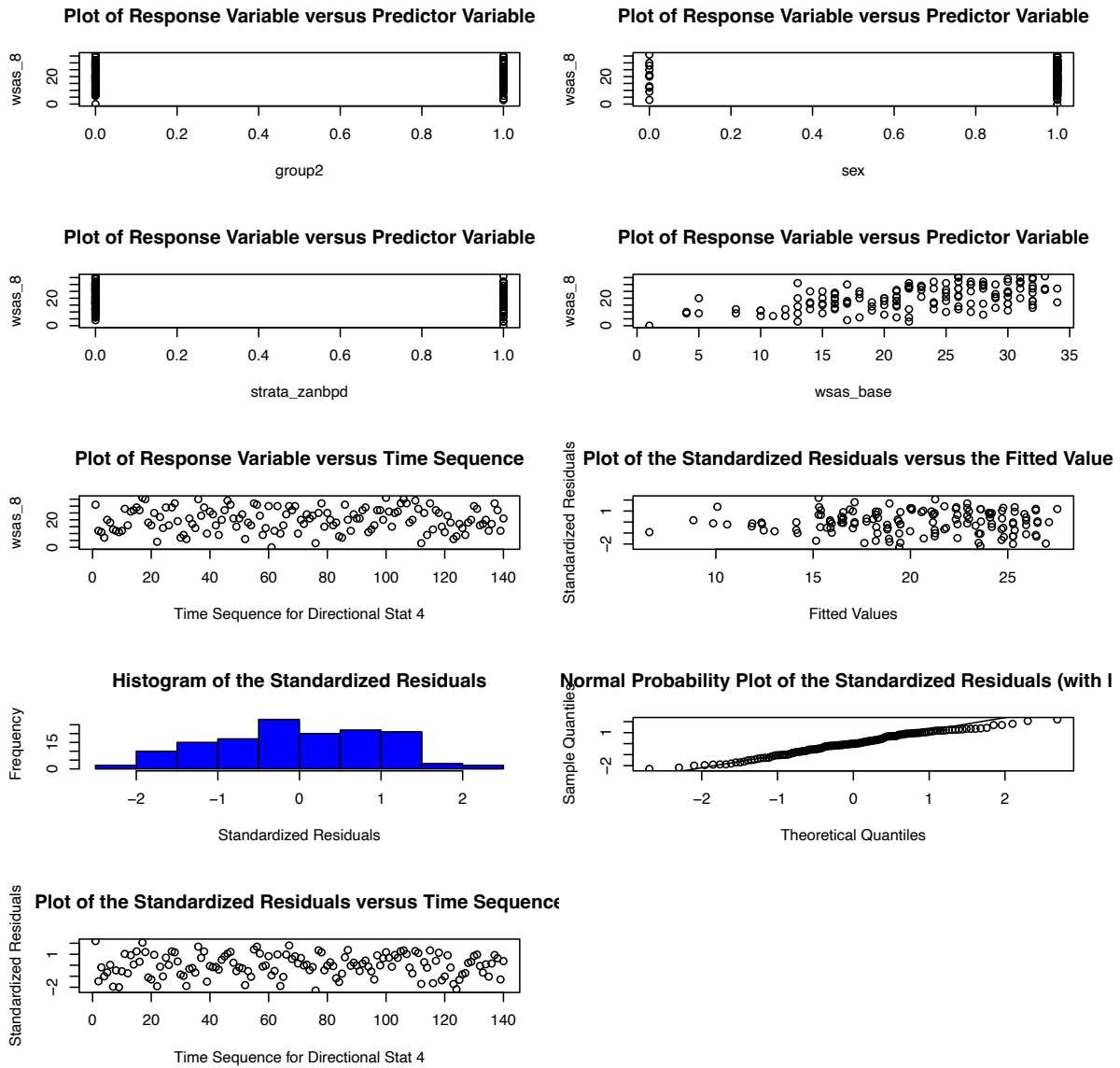
	Value	p-value	Decision
Global Stat	3.0192733	0.5546053	Assumptions acceptable.
Skewness	1.5539115	0.2125589	Assumptions acceptable.
Kurtosis	0.2087183	0.6477739	Assumptions acceptable.
Link Function	0.1037266	0.7474025	Assumptions acceptable.
Heteroscedasticity	1.1529169	0.2829394	Assumptions acceptable.

Analysis

zan_bpd_score_8			
Predictors	Estimates	CI	p
(Intercept)	12.27	9.45 – 15.09	<0.001
group [2]	0.95	-0.57 – 2.46	0.220
sex	1.13	-1.65 – 3.91	0.425
strata zanbpd	-4.97	-6.66 – -3.29	<0.001
Observations	148		
R ² / R ² adjusted	0.201 / 0.184		

Functional impairment (WSAS, 8-weeks)

Assumptions for linear regression



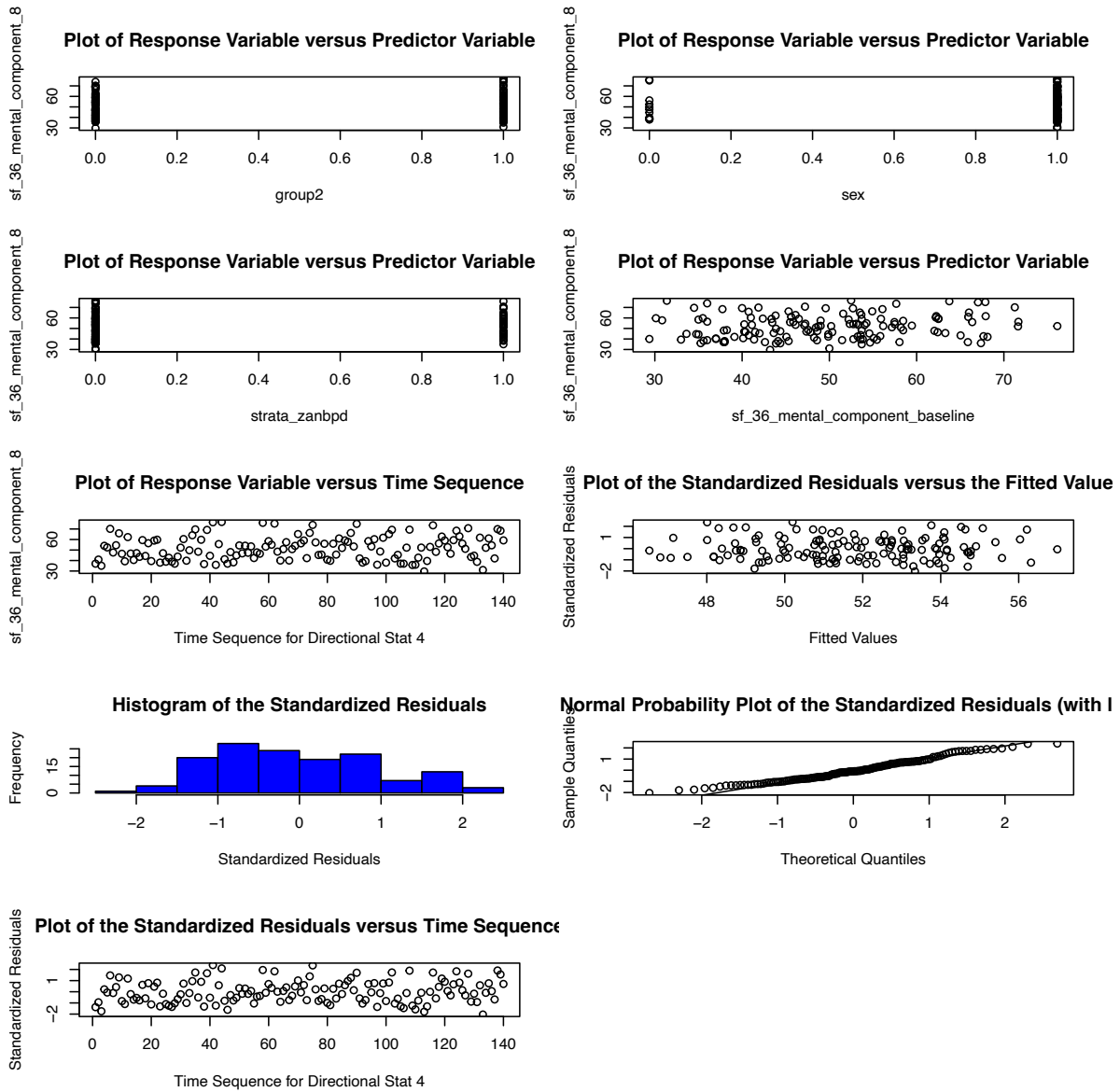
	Value	p-value	Decision
Global Stat	5.9132094	0.2057243	Assumptions acceptable.
Skewness	1.0530080	0.3048153	Assumptions acceptable.
Kurtosis	2.5693406	0.1089531	Assumptions acceptable.
Link Function	0.8933992	0.3445578	Assumptions acceptable.
Heteroscedasticity	1.3974615	0.2371491	Assumptions acceptable.

Analysis

wsas_8			
Predictors	Estimates	CI	p
(Intercept)	8.16	2.22 – 14.10	0.007
group [2]	0.50	-1.95 – 2.95	0.687
sex	-1.02	-5.56 – 3.53	0.659
strata zanbpd	-1.15	-3.92 – 1.63	0.415
wsas base	0.59	0.42 – 0.76	<0.001
Observations	140		
R ² / R ² adjusted	0.279 / 0.258		

Quality of life (SF-36 mental component, 8-weeks)

Assumptions for linear regression



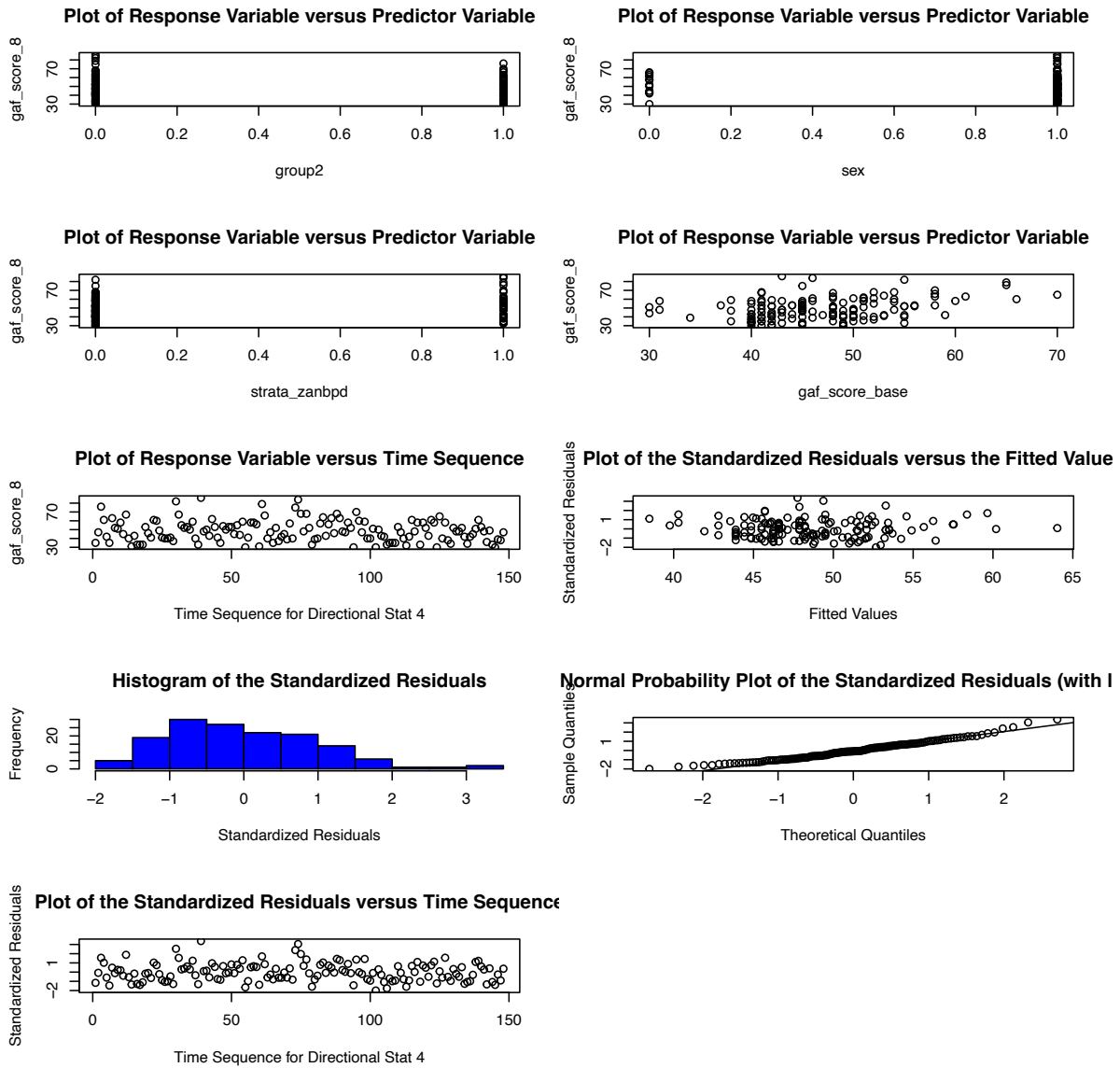
	Value	p-value	Decision
Global Stat	5.6989181	0.2227898	Assumptions acceptable.
Skewness	2.9612665	0.0852811	Assumptions acceptable.
Kurtosis	2.2065414	0.1374265	Assumptions acceptable.
Link Function	0.2972142	0.5856341	Assumptions acceptable.
Heteroscedasticity	0.2338959	0.6286500	Assumptions acceptable.

Analysis

sf_36_mental_component_8			
Predictors	Estimates	CI	p
(Intercept)	40.67	29.27 – 52.07	<0.001
group [2]	2.97	-0.78 – 6.72	0.120
sex	1.27	-5.67 – 8.21	0.718
strata zanbpd	0.14	-4.17 – 4.45	0.949
sf 36 mental component baseline	0.17	-0.02 – 0.36	0.078
Observations	140		
R ² / R ² adjusted	0.040 / 0.012		

Global functioning (GAF, 8-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	13.0374043	0.0110945	Assumptions NOT satisfied!
Skewness	7.7649331	0.0053270	Assumptions NOT satisfied!
Kurtosis	0.5953797	0.4403463	Assumptions acceptable.
Link Function	3.9536466	0.0467699	Assumptions NOT satisfied!
Heteroscedasticity	0.7234450	0.3950162	Assumptions acceptable.

Analysis

gaf_score_8			
Predictors	Estimates	CI	p
(Intercept)	26.17	10.41 – 41.93	0.001
group [2]	-1.28	-5.05 – 2.49	0.502
sex	-2.63	-9.56 – 4.29	0.454
strata zanbpd	0.96	-3.40 – 5.33	0.663
gaf score base	0.54	0.25 – 0.83	<0.001
Observations	148		
R ² / R ² adjusted	0.115 / 0.091		

Wilcoxon rank sum test with continuity correction

```
data: df$gaf_score_8[df$group == 2] and df$gaf_score_8[df$group == 1]
W = 2616, p-value = 0.641
alternative hypothesis: true location shift is not equal to 0
```


Severe self-harm (8-weeks)

Analysis

Relative risk derived by G-computation

RR: 1.4251 (sRR: 1.4227)

Confidence interval: 0.8744 - 2.4472

p-value: 0.1533

Statistical report for the MBT-RCT

Markus Harboe Olsen

13 november 2022 21:05

Table of contents

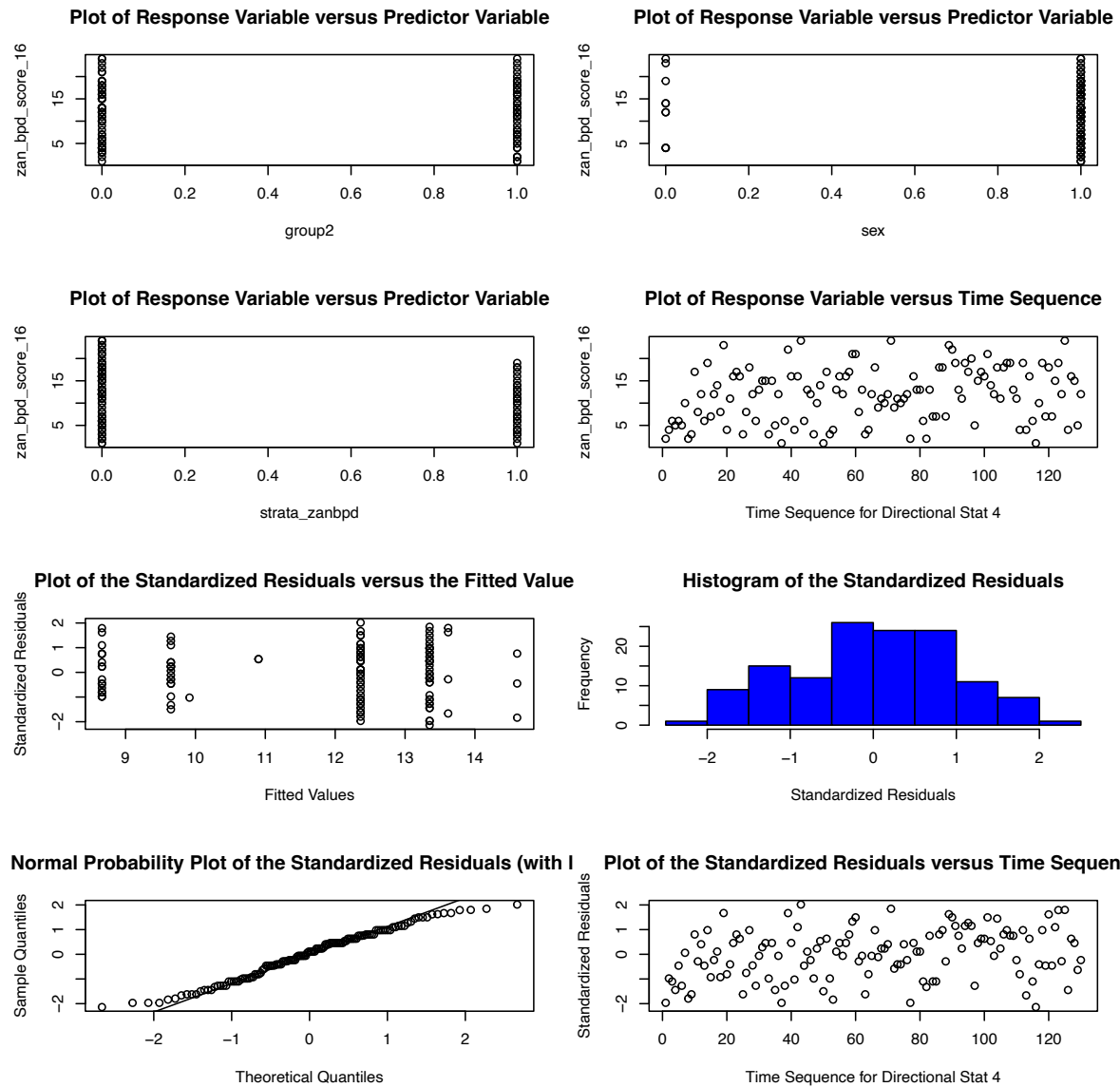
Primary outcome	2
Severity of borderline symptomatology (ZAN-BPD, 16-weeks)	2
Complete case analysis (primary analysis)	2
Multiple imputation (exploratory analysis)	4
Best-worst / worst-best (exploratory analysis)	5
8 vs. 16 weeks / 16 vs. 8 weeks (exploratory analysis)	6
Stratification variables (exploratory analysis)	7
Secondary outcomes	11
Functional impairment (WSAS, 16-weeks)	11
Quality of life (SF-36 mental component, 16-weeks)	13
Global functioning (GAF, 16-weeks)	15
Severe self-harm (16-weeks)	17
Serious adverse events (16-weeks)	17
Subgroup analysis	19
Baseline ZAN-BPD scores (0-11/12-36)	19
Sex (male/female)	20
Age (18-30/31-60)	21
Baseline GAF scores (0-48/49-100)	22
Baseline selfharm	23
Paused treatment due to COVID-19	24

Primary outcome

Severity of borderline symptomatology (ZAN-BPD, 16-weeks)

Complete case analysis (primary analysis)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	4.0541883	0.3987219	Assumptions acceptable.
Skewness	0.4777362	0.4894495	Assumptions acceptable.
Kurtosis	3.4208026	0.0643797	Assumptions acceptable.
Link Function	0.1556467	0.6931965	Assumptions acceptable.
Heteroscedasticity	0.0000027	0.9986988	Assumptions acceptable.

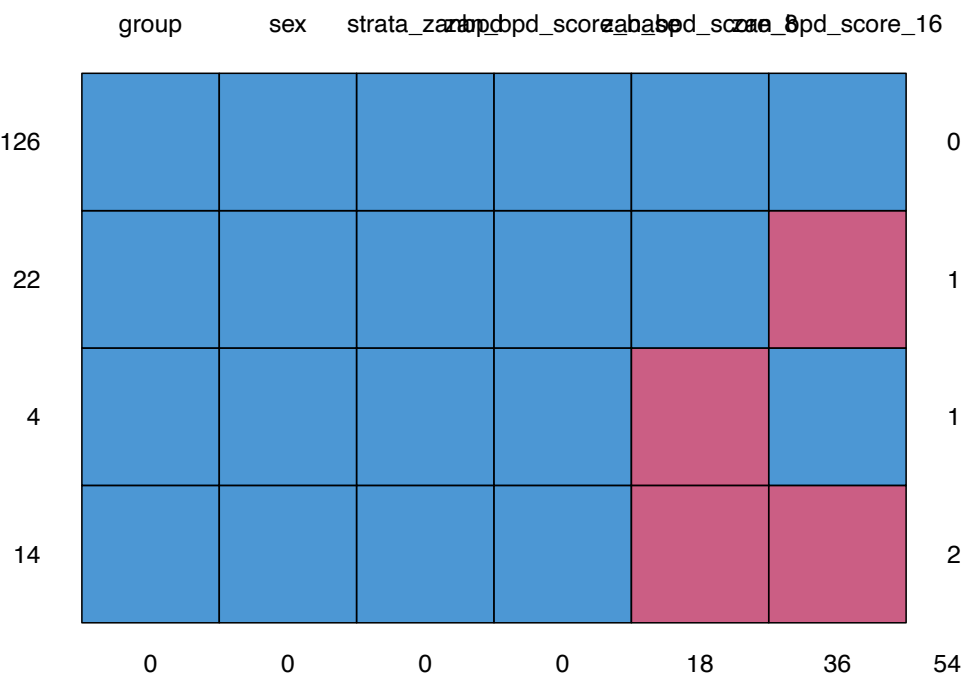
Summary data

	1	2
	(N=84)	(N=82)
zan_bpd_score_base		
Mean (SD)	13.9 (3.72)	13.8 (3.75)
zan_bpd_score_8		
Mean (SD)	12.0 (5.33)	12.8 (5.01)
Missing	10 (11.9%)	8 (9.8%)
zan_bpd_score_16		
Mean (SD)	11.3 (6.22)	12.3 (5.91)
Missing	14 (16.7%)	22 (26.8%)

Analysis

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.62	9.75 – 17.48	<0.001
group [2]	0.99	-1.06 – 3.03	0.341
sex	-1.25	-5.07 – 2.57	0.518
strata zanbpd	-3.70	-5.91 – -1.50	0.001
Observations	130		
R ² / R ² adjusted	0.089 / 0.067		

Multiple imputation (exploratory analysis)



Complete data summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.00	6.00	12.00	11.78	16.00	24.00	36

Imputed data summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	6.00	12.00	11.57	16.00	24.00

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	12.28	8.89 – 15.68	< 0.001
group [2]	1.53	-0.26 – 3.33	0.094
sex	-0.45	-3.80 – 2.89	0.789
strata zanbpd	-3.58	-5.55 – -1.60	< 0.001
Observations	166		
R ² / R ² adjusted	0.087 / 0.070		

Best-worst / worst-best (exploratory analysis)

Best 1 / Worst 2

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	11.64	8.28 – 15.00	< 0.001
group [2]	3.62	1.85 – 5.40	< 0.001
sex	-0.54	-3.86 – 2.77	0.746
strata zanbpd	-2.93	-4.88 – -0.98	0.003
Observations	166		
R ² / R ² adjusted	0.131 / 0.115		

Best 2 / Worst 1

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	15.12	11.78 – 18.46	< 0.001
group [2]	-1.60	-3.37 – 0.17	0.075
sex	-2.03	-5.32 – 1.26	0.226
strata zanbpd	-3.10	-5.04 – -1.16	0.002
Observations	166		
R ² / R ² adjusted	0.083 / 0.066		

8 vs. 16 weeks / 16 vs. 8 weeks (exploratory analysis)

1 is 8 weeks / 2 is 16 weeks

zan_816			
Predictors	Estimates	CI	p
(Intercept)	13.75	10.42 – 17.08	< 0.001
group [2]	0.50	-1.33 – 2.34	0.588
sex	-0.79	-4.12 – 2.53	0.637
strata zanbpd	-3.93	-5.94 – -1.92	< 0.001
Observations	134		
R ² / R ² adjusted	0.106 / 0.085		

2 is 8 weeks / 1 is 16 weeks

zan_816			
Predictors	Estimates	CI	p
(Intercept)	11.95	8.64 – 15.27	< 0.001
group [2]	1.42	-0.30 – 3.14	0.104
sex	0.88	-2.36 – 4.12	0.592
strata zanbpd	-4.76	-6.64 – -2.88	< 0.001
Observations	144		
R ² / R ² adjusted	0.168 / 0.151		

Stratification variables (exploratory analysis)

Age (18-30/31-60)

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.92	9.87 – 17.96	< 0.001
group [2]	1.04	-1.02 – 3.10	0.320
sex	-1.36	-5.21 – 2.49	0.486
strata zanbpd	-3.69	-5.91 – -1.48	0.001
covar ageTRUE	-0.55	-2.64 – 1.55	0.607
Observations	130		
R ² / R ² adjusted	0.091 / 0.062		

GAF (0-48/49-100)

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	14.13	10.15 – 18.10	<0.001
group [2]	0.93	-1.12 – 2.97	0.372
sex	-1.38	-5.20 – 2.45	0.478
strata zanbpd	-3.33	-5.64 – -1.02	0.005
covar gafTRUE	-1.18	-3.36 – 0.99	0.283
Observations	130		
R ² / R ² adjusted	0.097 / 0.068		

Selfharm ($\geq 1/0$)

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.21	9.24 – 17.17	<0.001
group [2]	0.97	-1.07 – 3.01	0.350
sex	-1.26	-5.09 – 2.56	0.515
strata zanbpd	-3.62	-5.83 – -1.41	0.002
self harm baseline [1]	0.99	-1.09 – 3.07	0.348
Observations	130		
R ² / R ² adjusted	0.095 / 0.066		

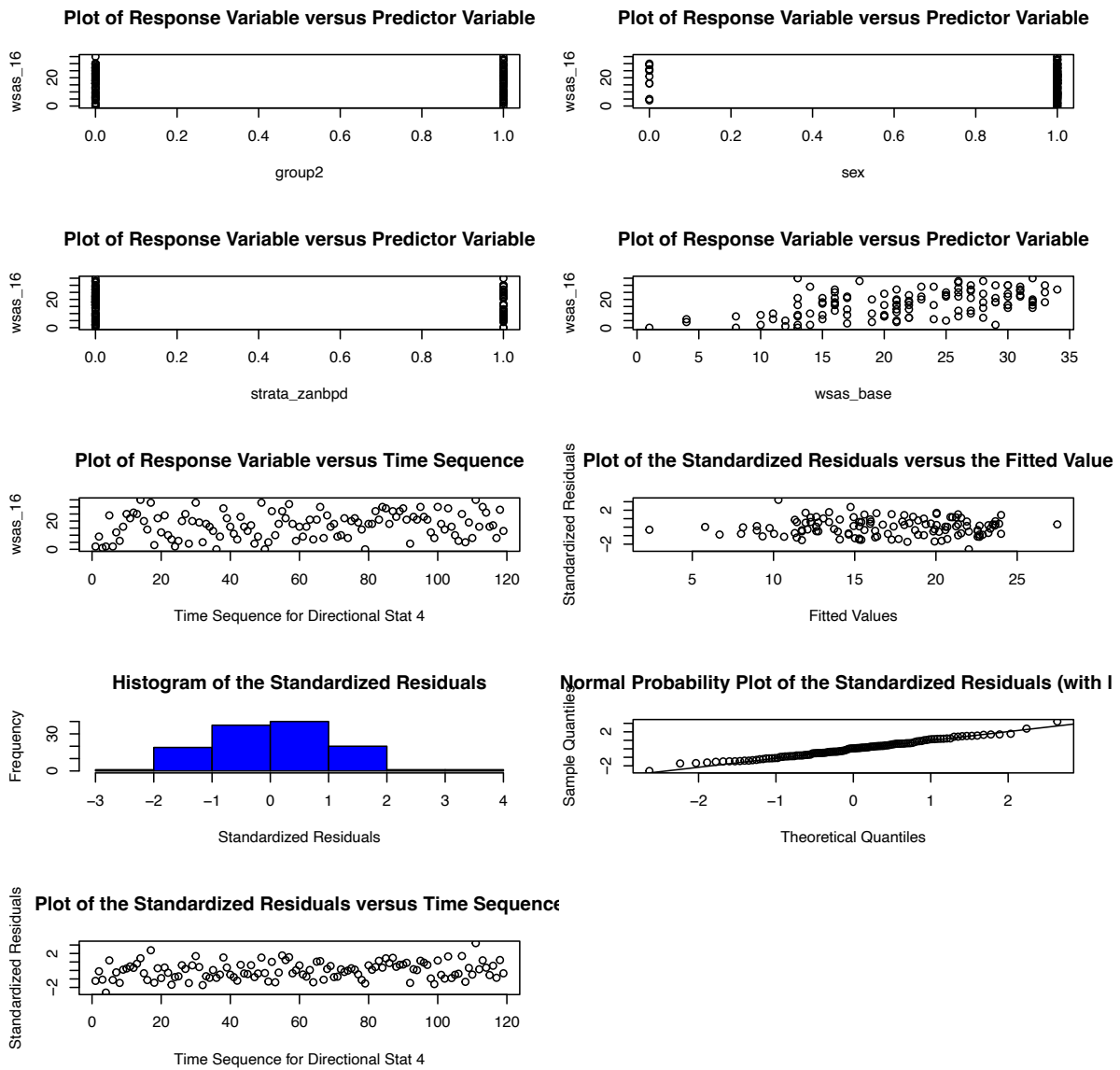
Paused treatment

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	14.29	10.34 – 18.24	< 0.001
group [2]	0.29	-1.95 – 2.53	0.798
sex	-0.75	-4.62 – 3.11	0.701
strata zanbpd	-3.67	-5.87 – -1.47	0.001
covid19 [Nej]	-1.68	-3.94 – 0.59	0.146
Observations	130		
R ² / R ² adjusted	0.104 / 0.076		

Secondary outcomes

Functional impairment (WSAS, 16-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	1.7158018	0.7878452	Assumptions acceptable.
Skewness	1.1646766	0.2804978	Assumptions acceptable.
Kurtosis	0.0103978	0.9187807	Assumptions acceptable.
Link Function	0.4868232	0.4853478	Assumptions acceptable.
Heteroscedasticity	0.0539042	0.8164038	Assumptions acceptable.

Summary data

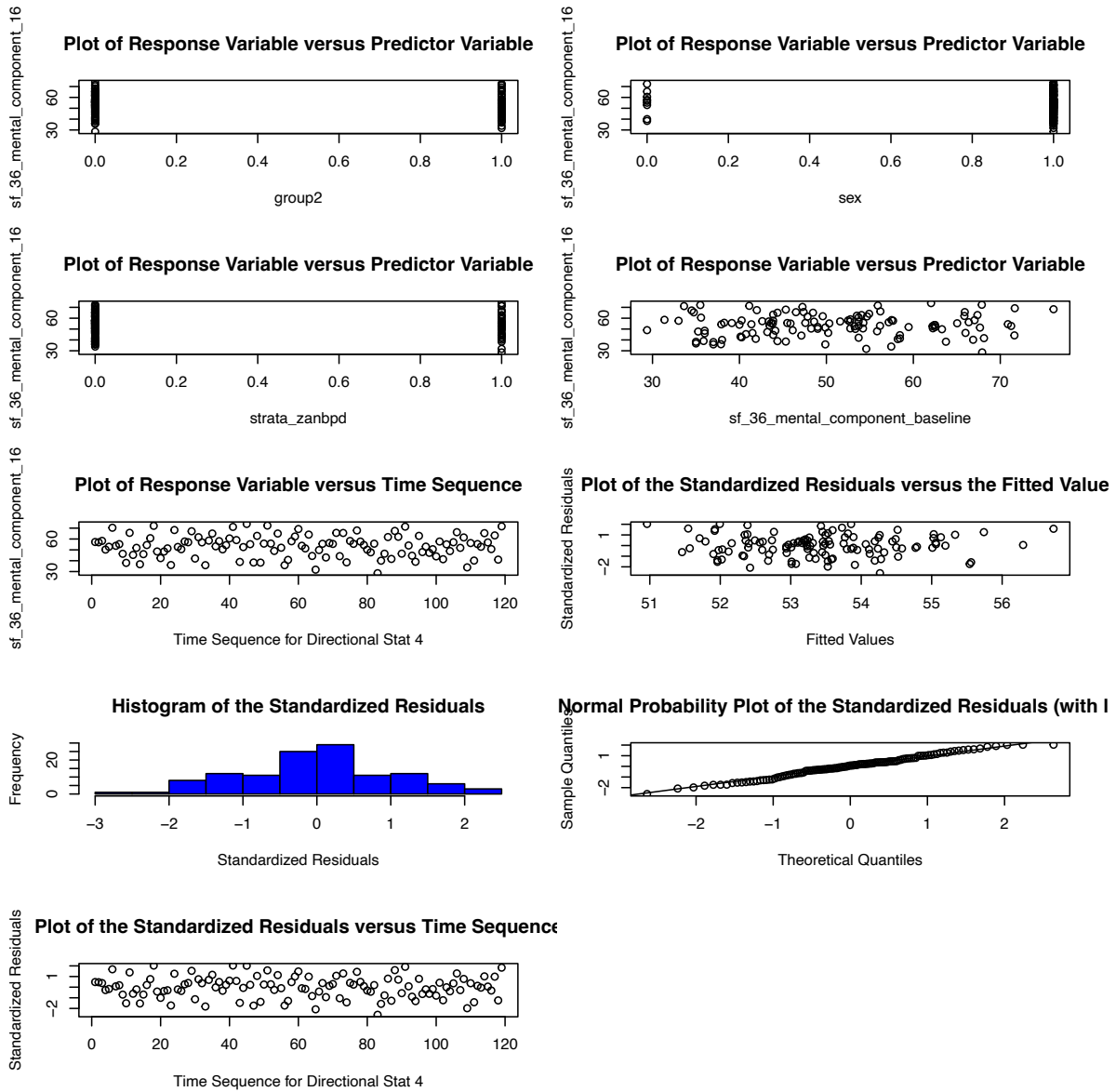
	1	2
	(N=84)	(N=82)
wsas_base		
Mean (SD)	22.9 (7.54)	21.9 (7.46)
wsas_8		
Mean (SD)	20.4 (8.59)	19.9 (8.36)
Missing	13 (15.5%)	13 (15.9%)
wsas_16		
Mean (SD)	16.7 (8.91)	17.2 (9.36)
Missing	21 (25.0%)	26 (31.7%)

Analysis

	wsas_16		
Predictors	Estimates	CI	p
(Intercept)	5.63	-1.33 – 12.58	0.112
group [2]	1.44	-1.43 – 4.32	0.321
sex	-4.24	-9.67 – 1.18	0.124
strata zanbpd	0.31	-2.82 – 3.45	0.844
wsas base	0.66	0.46 – 0.86	<0.001
Observations	119		
R ² / R ² adjusted	0.281 / 0.256		

Quality of life (SF-36 mental component, 16-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	6.4455556	0.1682533	Assumptions acceptable.
Skewness	0.2764293	0.5990512	Assumptions acceptable.
Kurtosis	0.8714568	0.3505523	Assumptions acceptable.
Link Function	5.2434878	0.0220291	Assumptions NOT satisfied!
Heteroscedasticity	0.0541817	0.8159404	Assumptions acceptable.

Summary data

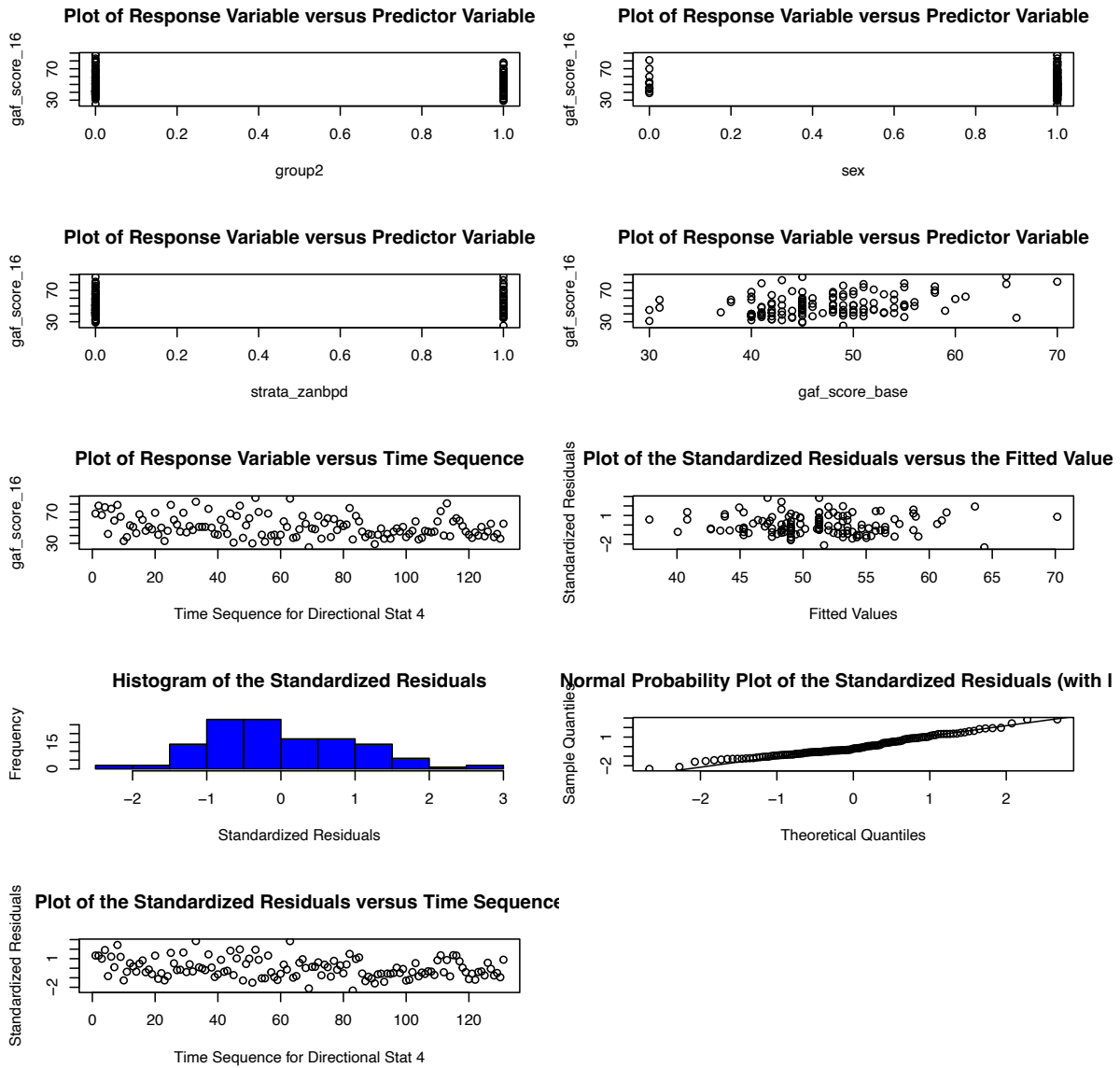
	1	2
	(N=84)	(N=82)
sf_36_mental_component_baseline		
Mean (SD)	50.3 (10.4)	49.2 (10.1)
sf_36_mental_component_8		
Mean (SD)	50.4 (10.8)	53.1 (11.6)
Missing	13 (15.5%)	13 (15.9%)
sf_36_mental_component_16		
Mean (SD)	53.9 (9.94)	52.9 (10.2)
Missing	21 (25.0%)	26 (31.7%)

Analysis

sf_36_mental_component_16			
Predictors	Estimates	CI	p
(Intercept)	52.91	41.94 – 63.88	<0.001
group [2]	-0.91	-4.62 – 2.79	0.626
sex	-2.48	-9.46 – 4.50	0.484
strata zanbpd	-0.90	-5.03 – 3.23	0.666
sf 36 mental component baseline	0.07	-0.11 – 0.25	0.439
Observations	119		
R ² / R ² adjusted	0.012 / -0.023		

Global functioning (GAF, 16-weeks)

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	10.3506099	0.0349182	Assumptions NOT satisfied!
Skewness	5.8210651	0.0158354	Assumptions NOT satisfied!
Kurtosis	0.0169192	0.8965081	Assumptions acceptable.
Link Function	1.4496393	0.2285858	Assumptions acceptable.
Heteroscedasticity	3.0629862	0.0800943	Assumptions acceptable.

Summary data

	1	2
	(N=84)	(N=82)
gaf_score_base		
Mean (SD)	46.5 (8.14)	46.3 (6.36)
gaf_score_8		
Mean (SD)	49.6 (13.5)	47.7 (10.5)
Missing	10 (11.9%)	8 (9.8%)
gaf_score_16		
Mean (SD)	52.2 (15.0)	49.7 (11.7)
Missing	14 (16.7%)	21 (25.6%)

Analysis

gaf_score_16			
Predictors	Estimates	CI	p
(Intercept)	17.82	-0.77 – 36.42	0.060
group [2]	-2.25	-6.70 – 2.20	0.318
sex	-0.19	-8.59 – 8.21	0.964
strata zanbpd	-2.60	-7.69 – 2.49	0.314
gaf score base	0.75	0.40 – 1.09	<0.001
Observations	131		
R ² / R ² adjusted	0.139 / 0.111		

Wilcoxon rank sum test with continuity correction

```
data: df$gaf_score_16[df$group == 2] and df$gaf_score_16[df$group == 1]
W = 2014, p-value = 0.578
alternative hypothesis: true location shift is not equal to 0
```

Severe self-harm (16-weeks)

Summary data

	1	2
	(N=84)	(N=82)
self_harm_baseline		
0	52 (61.9%)	49 (59.8%)
1	32 (38.1%)	33 (40.2%)
self_harm_8		
0	55 (65.5%)	47 (57.3%)
1	19 (22.6%)	27 (32.9%)
Missing	10 (11.9%)	8 (9.8%)
self_harm_16		
0	55 (65.5%)	45 (54.9%)
1	13 (15.5%)	16 (19.5%)
Missing	16 (19.0%)	21 (25.6%)

Analysis

Relative risk derived by G-computation
RR: 1.3717 (sRR: 1.3754)
Confidence interval: 0.6983 - 2.8412
p-value: 0.3353

Serious adverse events (16-weeks)

Summary data

	1	2
	(N=84)	(N=82)
as.factor(sae_total)		
0	69 (82.1%)	58 (70.7%)
1	9 (10.7%)	15 (18.3%)
2	4 (4.8%)	3 (3.7%)
3	1 (1.2%)	3 (3.7%)
4	0 (0%)	1 (1.2%)
5	1 (1.2%)	1 (1.2%)
7	0 (0%)	1 (1.2%)

Analysis

Fisher's Exact Test for Count Data

```
data: table(df$sae_total > 0, df$group)
p-value = 0.1003
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.8625312 4.2797517
sample estimates:
odds ratio
 1.896076
```

Wilcoxon rank sum test with continuity correction

```
data: df$sae_total[df$group == "1"] and df$sae_total[df$group == "2"]
W = 3046.5, p-value = 0.08356
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -3.403897e-05 5.726355e-05
sample estimates:
difference in location
 -1.998471e-06
```

Subgroup analysis

Baseline ZAN-BPD scores (0-11/12-36)

P-value for interaction: 0.6904671

Baseline ZAN-BPD scores 0-11 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	9.60	3.40 – 15.79	0.003
group [2]	1.61	-1.60 – 4.82	0.317
sex	-1.23	-7.31 – 4.86	0.685
Observations	40		
R ² / R ² adjusted	0.034 / -0.018		

Baseline ZAN-BPD scores 12-36 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.70	8.86 – 18.53	<0.001
group [2]	0.71	-1.92 – 3.35	0.593
sex	-1.20	-6.10 – 3.70	0.627
Observations	90		
R ² / R ² adjusted	0.006 / -0.017		

:::

...

Sex (male/female)

P-value for interaction: 0.7697972

Sex 0 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	14.06	4.87 – 23.25	0.009
group [2]	-0.14	-12.84 – 12.56	0.980
strata zanbpd	-3.30	-17.15 – 10.55	0.591
Observations	10		
R ² / R ² adjusted	0.046 / -0.226		

Sex 1 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	12.33	10.76 – 13.90	<0.001
group [2]	1.08	-1.01 – 3.16	0.308
strata zanbpd	-3.71	-5.96 – -1.47	0.001
Observations	120		
R ² / R ² adjusted	0.091 / 0.076		

...

...

Age (18-30/31-60)

P-value for interaction: 0.6602134

Age 18-30 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.81	7.98 – 19.65	< 0.001
group [2]	1.43	-1.26 – 4.12	0.294
sex	-1.43	-7.35 – 4.50	0.632
strata zanbpd	-3.68	-6.57 – -0.80	0.013
Observations	76		
R ² / R ² adjusted	0.105 / 0.068		

Age 31-60 ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.70	8.06 – 19.34	< 0.001
group [2]	0.48	-2.99 – 3.94	0.784
sex	-1.45	-6.88 – 3.97	0.593
strata zanbpd	-3.55	-7.26 – 0.16	0.060
Observations	54		
R ² / R ² adjusted	0.074 / 0.018		

...

...

Baseline GAF scores (0-48/49-100)

P-value for interaction: 0.6055547

Baseline GAF scores 0-48 :: { .cell-output-display }

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	15.36	9.63 – 21.10	< 0.001
group [2]	1.35	-1.38 – 4.08	0.329
sex	-2.79	-8.37 – 2.79	0.323
strata zanbpd	-3.93	-7.40 – -0.46	0.027
Observations	77		
R ² / R ² adjusted	0.081 / 0.043		

Baseline GAF scores 49-100 :: { .cell-output-display }

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	11.91	6.51 – 17.30	< 0.001
group [2]	0.25	-2.97 – 3.48	0.874
sex	-0.06	-5.61 – 5.49	0.983
strata zanbpd	-3.04	-6.29 – 0.21	0.066
Observations	53		
R ² / R ² adjusted	0.070 / 0.013		

...

...

Baseline selfharm

P-value for interaction: 0.1563246

Baseline no selfharm ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.87	8.62 – 19.12	< 0.001
group [2]	-0.25	-2.98 – 2.47	0.853
sex	-1.43	-6.52 – 3.66	0.578
strata zanbpd	-3.49	-6.38 – -0.60	0.018
Observations	77		
R ² / R ² adjusted	0.075 / 0.037		

Baseline selfharm ::: {.cell-output-display}

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	13.13	7.13 – 19.12	< 0.001
group [2]	2.74	-0.44 – 5.93	0.090
sex	-0.93	-7.03 – 5.18	0.762
strata zanbpd	-3.93	-7.59 – -0.27	0.036
Observations	53		
R ² / R ² adjusted	0.139 / 0.086		

...

⋮

Paused treatment due to COVID-19

P-value for interaction: 0.3758016

Baseline no pause :: { .cell-output-display }

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	6.92	-1.26 – 15.09	0.096
group [2]	1.49	-1.85 – 4.83	0.375
sex	5.52	-2.72 – 13.75	0.185
strata zanbpd	-5.83	-8.94 – -2.72	<0.001
Observations	64		
R ² / R ² adjusted	0.222 / 0.183		

Baseline pause :: { .cell-output-display }

zan_bpd_score_16			
Predictors	Estimates	CI	p
(Intercept)	15.87	11.28 – 20.46	<0.001
group [2]	-0.37	-3.40 – 2.65	0.806
sex	-2.78	-7.13 – 1.58	0.207
strata zanbpd	-1.55	-4.65 – 1.56	0.323
Observations	66		
R ² / R ² adjusted	0.043 / -0.003		

⋮ ⋮

```
-----
name: <unnamed>
log: /Users/janusjakobsen/Desktop/MBT_FINAL.log
log type: text
opened on: 10 Nov 2022, 21:11:08
```

```
. *SYNTAX MBT
```

```
.
. *ZAN16
```

```
. codebook zan_bpd_score_16 if group ==1
```

```
-----
zan_bpd_score_16                                     ZAN BPD Samlet Score 16 mV•neder
-----
```

```
          Type: Numeric (byte)
```

```
          Range: [1,24]                               Units: 1
Unique values: 22                               Missing .: 14/84
```

```
          Mean: 11.3429
          Std. dev.: 6.22032
```

```
          Percentiles:    10%      25%      50%      75%      90%
                           3         6       12       16       19
```

```
. codebook zan_bpd_score_16 if group ==2
```

```
-----
zan_bpd_score_16                                     ZAN BPD Samlet Score 16 mV•neder
-----
```

```
          Type: Numeric (byte)
```

```
          Range: [1,24]                               Units: 1
```

Unique values: 23

Missing .: 22/82

Mean: 12.2833

Std. dev.: 5.90631

Percentiles:	10%	25%	50%	75%	90%
	4	7.5	12	17	19

.
. by group, sort : summarize zan_bpd_score_16, detail

-> group = 1

ZAN BPD Samlet Score 16 mV•neder

Percentiles		Smallest		
1%	1	1		
5%	3	2		
10%	3	3	Obs	70
25%	6	3	Sum of wgt.	70
50%	12	Largest	Mean	11.34286
			Std. dev.	6.220317
75%	16	22		
90%	19	23	Variance	38.69234
95%	22	24	Skewness	.1647965
99%	24	24	Kurtosis	1.967706

-> group = 2

ZAN BPD Samlet Score 16 mV•neder

Percentiles		Smallest		
1%	1	1		
5%	2	1		
10%	4	2	Obs	60

25%	7.5	2	Sum of wgt.	60
50%	12		Mean	12.28333
		Largest	Std. dev.	5.906307
75%	17	21		
90%	19	22	Variance	34.88446
95%	21.5	23	Skewness	-.150302
99%	24	24	Kurtosis	2.192129

```
.
. regress zan_bpd_score_16 i.group i.sex i.ZAN_DIKO, vce(robust)
```

Linear regression	Number of obs	=	130
	F(3, 126)	=	4.67
	Prob > F	=	0.0040
	R-squared	=	0.0890
	Root MSE	=	5.8644

zan_bpd_s~16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	

2.group	.9872565	1.033165	0.96	0.341	-1.057348	3.031861
1.sex	-1.253016	2.299238	-0.54	0.587	-5.80314	3.297108
2.ZAN_DIKO	3.701827	1.013785	3.65	0.000	1.695576	5.708078
_cons	9.915093	2.365532	4.19	0.000	5.233775	14.59641

```
.
. regress zan_bpd_score_16 i.group i.sex i.ZAN_DIKO c.gaf_score_base i.AGE_DIKO i.covid19, vc
> e(robust)
```

Linear regression	Number of obs	=	130
	F(6, 123)	=	3.60
	Prob > F	=	0.0025
	R-squared	=	0.1322
	Root MSE	=	5.7931

zan_bpd_sco~16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
2.group	.3429771	1.139462	0.30	0.764	-1.912518	2.598472
1.sex	-1.32729	1.968426	-0.67	0.501	-5.223668	2.569089
2.ZAN_DIKO	2.929822	1.059601	2.77	0.007	.8324079	5.027237
gaf_score_base	-.1558485	.0836536	-1.86	0.065	-.3214357	.0097386
2.AGE_DIKO	-1.008759	1.099965	-0.92	0.361	-3.186073	1.168555
covid19						
Ja	1.733934	1.148397	1.51	0.134	-.5392488	4.007116
_cons	17.72583	5.002442	3.54	0.001	7.823805	27.62786

```
.
. *ZAN 8
```

```
.
. regress zan_bpd_score_8 i.group i.sex ZAN_DIKO, vce(robust)
```

```
Linear regression      Number of obs      =      148
                      F(3, 144)              =      16.55
                      Prob > F                =      0.0000
                      R-squared               =      0.2007
                      Root MSE              =      4.6696
```

zan_bpd_sc~8	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
2.group	.9452221	.7677486	1.23	0.220	-.5722906	2.462735
1.sex	1.126444	1.187089	0.95	0.344	-1.219926	3.472814
ZAN_DIKO	4.973218	.7388048	6.73	0.000	3.512915	6.433521
_cons	2.322026	1.553175	1.50	0.137	-.7479419	5.391994

```
.
. *Subgroup analyses
```

```
.
. regress zan_bpd_score_16 i.group##i.ZAN_DIKO i.sex, vce(robust)
```

```
Linear regression               Number of obs   =          130
                                F(4, 125)       =           3.53
                                Prob > F        =          0.0091
                                R-squared       =          0.0901
                                Root MSE    =          5.8841
```

zan_bpd_sco~16		Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
-----+-----							
2.group		1.607014	1.547793	1.04	0.301	-1.45626	4.670288
2.ZAN_DIKO		4.122151	1.43171	2.88	0.005	1.28862	6.955682
group#ZAN_DIKO							
2 2		-.8959394	2.033487	-0.44	0.660	-4.920462	3.128583
1.sex		-1.208758	2.28515	-0.53	0.598	-5.731352	3.313837
_cons		9.579769	2.459494	3.90	0.000	4.712126	14.44741

```
.
. regress zan_bpd_score_16 i.group##i.sex i.ZAN_DIKO, vce(robust)
```

```
Linear regression               Number of obs   =          130
                                F(4, 125)       =           3.47
                                Prob > F        =          0.0100
                                R-squared       =          0.0896
                                Root MSE    =          5.8858
```

zan_bpd_s~16		Coefficient	Robust std. err.	t	P> t	[95% conf. interval]

2.group	1	-0.0634023	4.481878	-0.01	0.989	-8.933596	8.806791
1.sex	1	-1.818755	3.743739	-0.49	0.628	-9.228079	5.590569
group#sex	2 1	1.138483	4.603252	0.25	0.805	-7.971923	10.24889
2.ZAN_DIKO	1	3.682989	1.022696	3.60	0.000	1.658947	5.707031
_cons	1	10.45361	3.728316	2.80	0.006	3.07481	17.83241

```
. regress zan_bpd_score_16 i.group##AGE_DIKO i.sex i.ZAN_DIKO, vce(robust)
```

Linear regression	Number of obs	=	130
	F(5, 124)	=	2.81
	Prob > F	=	0.0193
	R-squared	=	0.0923
	Root MSE	=	5.9006

zan_bpd_sco~16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]
2.group	1.433801	1.333303	1.08	0.284	-1.205178 4.072781
2.AGE_DIKO	-.1009442	1.550351	-0.07	0.948	-3.169524 2.967635
group#AGE_DIKO					
2 2	-.9438138	2.167995	-0.44	0.664	-5.234883 3.347256
1.sex	-1.443982	2.3155	-0.62	0.534	-6.027006 3.139041
2.ZAN_DIKO	3.627305	1.025408	3.54	0.001	1.597734 5.656875
_cons	10.18208	2.420876	4.21	0.000	5.390486 14.97367

```
. regress zan_bpd_score_16 i.group##GAF_DIKO i.sex i.ZAN_DIKO, vce(robust)
```

Linear regression

Number of obs = 130
F(5, 124) = 3.15
Prob > F = 0.0104
R-squared = 0.0993
Root MSE = 5.8779

zan_bpd_sco~16	Robust					
	Coefficient	std. err.	t	P> t	[95% conf. interval]	

2.group	1.368406	1.383108	0.99	0.324	-1.369152	4.105964
2.GAF_DIKO	-.6845177	1.527697	-0.45	0.655	-3.708258	2.339222

group#GAF_DIKO						
2 2	-1.093424	2.087586	-0.52	0.601	-5.225341	3.038493

1.sex	-1.335943	2.264589	-0.59	0.556	-5.818199	3.146312
2.ZAN_DIKO	3.339244	1.089468	3.07	0.003	1.182881	5.495607
_cons	10.53927	2.541589	4.15	0.000	5.508754	15.56979

.
. regress zan_bpd_score_16 i.group##self_harm_baseline i.sex i.ZAN_DIKO, vce(robust)

Linear regression

Number of obs = 130
F(5, 124) = 3.38
Prob > F = 0.0068
R-squared = 0.1100
Root MSE = 5.843

zan_bpd_score_16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]		

2.group	-.2502772	1.368693	-0.18	0.855	-2.959305	2.45875	
1.self_harm_baseline	-.3982646	1.483888	-0.27	0.789	-3.335294	2.538765	

group#self_harm_baseline							

2 1		2.982707	2.061626	1.45	0.150	-1.097828	7.063243
1.sex		-1.276951	2.297681	-0.56	0.579	-5.824705	3.270803
2.ZAN_DIKO		3.636319	1.034754	3.51	0.001	1.588251	5.684386
_cons		10.14248	2.390077	4.24	0.000	5.411848	14.87311

```
. regress zan_bpd_score_16 i.group##covid19 i.sex i.ZAN_DIKO, vce(robust)
```

Linear regression	Number of obs	=	130
	F(5, 124)	=	3.92
	Prob > F	=	0.0025
	R-squared	=	0.1100
	Root MSE	=	5.8429

zan_bpd_sc~16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]
2.group	1.415185	1.67451	0.85	0.400	-1.899141 4.72951
covid19					
Ja	2.562568	1.519528	1.69	0.094	-.4450042 5.57014
group#covid19					
2#Ja	-2.038032	2.27515	-0.90	0.372	-6.541191 2.465127
1.sex	-.7795747	2.154316	-0.36	0.718	-5.04357 3.48442
2.ZAN_DIKO	3.796214	.973144	3.90	0.000	1.87009 5.722339
_cons	8.60402	2.253566	3.82	0.000	4.143581 13.06446

```
.
.
.
.
.
```

```
. *WSAS
```

```
. codebook wsas_16 if group ==1
```

```
-----  
wsas_16                               WSAS 16 mV•nederTotal scoren-+  
-----
```

```
      Type: Numeric (byte)
```

```
      Range: [0,35]                      Units: 1  
Unique values: 29                      Missing .: 21/84
```

```
      Mean: 16.7143  
Std. dev.: 8.91434
```

```
Percentiles:    10%    25%    50%    75%    90%  
                5      9     17     24     28
```

```
. codebook wsas_16 if group ==2
```

```
-----  
wsas_16                               WSAS 16 mV•nederTotal scoren-+  
-----
```

```
      Type: Numeric (byte)
```

```
      Range: [0,35]                      Units: 1  
Unique values: 30                      Missing .: 26/82
```

```
      Mean: 17.1964  
Std. dev.: 9.3603
```

```
Percentiles:    10%    25%    50%    75%    90%  
                4      9     18    23.5    30
```

```
.  
. by group, sort : summarize wsas_16, detail
```

-> group = 1

WSAS 16 mV•nederTotal scoren~†

Percentiles		Smallest		
1%	0	0		
5%	2	0		
10%	5	1	Obs	63
25%	9	2	Sum of wgt.	63
50%	17	Largest	Mean	16.71429
75%	24		Std. dev.	8.914339
90%	28	30	Variance	79.46544
95%	30	30	Skewness	-.1290234
99%	35	35	Kurtosis	2.082277

-> group = 2

WSAS 16 mV•nederTotal scoren~†

Percentiles		Smallest		
1%	0	0		
5%	2	2		
10%	4	2	Obs	56
25%	9	3	Sum of wgt.	56
50%	18	Largest	Mean	17.19643
75%	23.5		Std. dev.	9.360302
90%	30	33	Variance	87.61526
95%	33	33	Skewness	.0359029
99%	35	35	Kurtosis	2.030963

```
. regress wsas_16 i.group i.sex ZAN_DIKO c.wsas_base, vce(robust)
```

```
Linear regression               Number of obs   =          119
                                F(4, 114)         =          15.71
                                Prob > F          =          0.0000
                                R-squared          =          0.2811
                                Root MSE        =          7.8425
```

		Robust				
wsas_16	Coefficient	std. err.	t	P> t	[95% conf. interval]	
2.group	1.444362	1.471335	0.98	0.328	-1.470342	4.359067
1.sex	-4.243747	2.262661	-1.88	0.063	-8.72606	.2385663
ZAN_DIKO	-.3113993	1.603885	-0.19	0.846	-3.488684	2.865885
wsas_base	.661873	.0909459	7.28	0.000	.4817098	.8420361
_cons	6.2521	3.850109	1.62	0.107	-1.374936	13.87914

```
. regress wsas_16 i.group i.sex i.ZAN_DIKO c.wsas_base c.gaf_score_base i.AGE_DIKO i.covid19,
> vce(robust)
```

```
Linear regression               Number of obs   =          119
                                F(7, 111)         =          15.22
                                Prob > F          =          0.0000
                                R-squared          =          0.3061
                                Root MSE        =          7.8083
```

		Robust				
wsas_16	Coefficient	std. err.	t	P> t	[95% conf. interval]	
2.group	.9241989	1.735036	0.53	0.595	-2.51389	4.362288
1.sex	-3.968578	2.222537	-1.79	0.077	-8.372685	.435528
2.ZAN_DIKO	-.6955572	1.664813	-0.42	0.677	-3.994495	2.603381
wsas_base	.631204	.0982636	6.42	0.000	.4364881	.8259199

```

gaf_score_base | -.0905957 .1188835 -0.76 0.448 -.3261713 .1449799
2.AGE_DIKO | -2.322508 1.462484 -1.59 0.115 -5.220518 .5755019
covid19 |
Ja | 2.012183 1.786744 1.13 0.263 -1.528369 5.552736
_cons | 11.11642 7.485101 1.49 0.140 -3.715804 25.94865
-----

```

```

.
. *WSAS 8

```

```

. regress wsas_8 i.group i.sex ZAN_DIKO, vce(robust)

```

```

Linear regression              Number of obs   =          140
                              F(3, 136)         =           1.11
                              Prob > F           =          0.3456
                              R-squared          =          0.0266
                              Root MSE       =          8.4295

```

```

-----
wsas_8 |      Robust
        | Coefficient std. err.      t    P>|t|    [95% conf. interval]
-----+-----
2.group |   -.3395716   1.42802   -0.24  0.812   -3.163567   2.484424
1.sex   |    .1089452   2.748673    0.04  0.968   -5.326723   5.544614
ZAN_DIKO |    3.014515   1.686959    1.79  0.076   -.3215487   6.350579
_cons   |   15.07056    3.96227    3.80  0.000    7.234929   22.90619
-----

```

```

.
.
. *SF36 MENTAL

.
. codebook sf_36_mental_component_16 if group ==1
-----

```

```
sf_36_mental_component_16
```

Type: Numeric (float)

Range: [28.5,73.875] Units: 1.000e-06
Unique values: 59 Missing .: 21/84

Mean: 53.8737
Std. dev.: 9.94485

Percentiles:	10%	25%	50%	75%	90%
	40	47.25	55.25	61.625	65.625

```
. codebook sf_36_mental_component_16 if group ==2
```

```
sf_36_mental_component_16
```

Type: Numeric (float)

Range: [31.75,72.375] Units: 1.000e-06
Unique values: 56 Missing .: 26/82

Mean: 52.907
Std. dev.: 10.17

Percentiles:	10%	25%	50%	75%	90%
	38.5	45.8958	53.3333	58.0833	67.4167

```
.  
. by group, sort : summarize sf_36_mental_component_16, detail
```

```
-> group = 1
```

SF-36 mental 16

```

-----
Percentiles      Smallest
1%              28.5      28.5
5%              36       35.875
10%             40       35.875      Obs              63
25%            47.25      36       Sum of wgt.        63

50%            55.25

75%            61.625      Largest
90%            65.625      68.25      Mean              53.87368
95%            68.25      70.375      Std. dev.         9.944855
99%            73.875      71.625      Variance          98.90014
                          73.875      Skewness          -.3026029
                          Kurtosis          2.61096
-----

```

```

-> group = 2

```

SF-36 mental 16

```

-----
Percentiles      Smallest
1%              31.75     31.75
5%              36.75     33.875
10%             38.5      36.75      Obs              56
25%            45.89583    38.25     Sum of wgt.        56

50%            53.33333

75%            58.08333      Largest
90%            67.41666      71.125      Mean              52.90699
95%            71.375      71.375      Std. dev.         10.16999
99%            72.375      72.125      Variance          103.4287
                          72.375      Skewness          .0625108
                          Kurtosis          2.413428
-----

```

```

.
. regress sf_36_mental_component_16 i.group i.sex ZAN_DIKO c.sf_36_mental_component_baseline,
> vce(robust)

```

```

Linear regression              Number of obs      =          119

```

```

F(4, 114)          =      0.27
Prob > F           =      0.8962
R-squared          =      0.0117
Root MSE          =      10.135

```

		Robust				
sf_36_mental_component_16	Coefficient	std. err.	t	P> t	[95% conf. interval]	
2.group	-.9136621	1.856351	-0.49	0.624	-4.59108	2.763756
1.sex	-2.477248	3.641619	-0.68	0.498	-9.691267	4.736771
ZAN_DIKO	.9024071	2.120605	0.43	0.671	-3.298494	5.103308
sf_36_mental_component_b~e	.0697354	.0975791	0.71	0.476	-.123568	.2630388
_cons	51.10482	7.075228	7.22	0.000	37.08885	65.1208

```

. regress sf_36_mental_component_16 i.group i.sex i.ZAN_DIKO c.gaf_score_base i.AGE_DIKO i.co
> vid19 c.sf_36_mental_component_baseline, vce(robust)

```

```

Linear regression      Number of obs   =      119
                      F(7, 111)       =      1.41
                      Prob > F         =      0.2100
                      R-squared        =      0.0639
                      Root MSE       =      9.9962

```

		Robust				
sf_36_mental_component_16	Coefficient	std. err.	t	P> t	[95% conf. interval]	
2.group	-2.129969	1.99188	-1.07	0.287	-6.077012	1.817074
1.sex	-.9036907	3.574006	-0.25	0.801	-7.985822	6.178441
2.ZAN_DIKO	1.772944	2.148397	0.83	0.411	-2.484249	6.030137
gaf_score_base	.2368787	.1460338	1.62	0.108	-.0524971	.5262544
2.AGE_DIKO	3.813009	1.88113	2.03	0.045	.0854234	7.540594
covid19						
Ja	1.454599	2.102585	0.69	0.490	-2.711813	5.621011


```
sf_36_mental_component_b~e | .0489562 .092867 0.53 0.599 -.1350658 .2329783
      _cons | 38.05843 8.497336 4.48 0.000 21.22039 54.89647
-----
```

```
.
. *8
```

```
.
. regress sf_36_mental_component_16 i.group i.sex ZAN_DIKO c.sf_36_mental_component_baseline,
> vce(robust)
```

```
Linear regression                Number of obs    =          119
                                F(4, 114)         =           0.27
                                Prob > F           =          0.8962
                                R-squared           =          0.0117
                                Root MSE        =          10.135
```

```
-----
sf_36_mental_component_16 |      Coefficient   Robust      t      P>|t|      [95% conf. interval]
-----+-----
      2.group |   -.9136621   1.856351   -0.49   0.624   -4.59108   2.763756
      1.sex |  -2.477248   3.641619   -0.68   0.498   -9.691267   4.736771
      ZAN_DIKO |   .9024071   2.120605    0.43   0.671   -3.298494   5.103308
sf_36_mental_component_b~e |   .0697354   .0975791    0.71   0.476   -.123568   .2630388
      _cons |   51.10482   7.075228    7.22   0.000   37.08885   65.1208
-----
```

```
.
.
. *GAF
```

```
.
. codebook gaf_score_16 if group ==1
```

```
-----
gaf_score_16                                Laveste GAF Score 16 mV•neder
-----
```

```

Type: Numeric (byte)
Range: [25,88]
Unique values: 40
Units: 1
Missing .: 14/84

```

```

Mean: 52.2286
Std. dev.: 15.0286

```

```

Percentiles:    10%    25%    50%    75%    90%
                37     41     49     62     76

```

```

. codebook gaf_score_16 if group ==2

```

```

-----
gaf_score_16                                     Laveste GAF Score 16 mV•neder
-----

```

```

Type: Numeric (byte)
Range: [29,78]
Unique values: 33
Units: 1
Missing .: 21/82

```

```

Mean: 49.7377
Std. dev.: 11.7373

```

```

Percentiles:    10%    25%    50%    75%    90%
                36     41     48     58     65

```

```

.
. by group, sort : summarize gaf_score_16, detail

```

```

-----
-> group = 1

```

```

Laveste GAF Score 16 mV•neder
-----

```

```

Percentiles    Smallest

```

1%	25	25		
5%	33	31		
10%	37	33	Obs	70
25%	41	33	Sum of wgt.	70
50%	49		Mean	52.22857
		Largest	Std. dev.	15.02864
75%	62	81		
90%	76	83	Variance	225.86
95%	81	87	Skewness	.654759
99%	88	88	Kurtosis	2.536603

-> group = 2

Laveste GAF Score 16 mV•neder

	Percentiles	Smallest		
1%	29	29		
5%	32	30		
10%	36	32	Obs	61
25%	41	32	Sum of wgt.	61
50%	48		Mean	49.7377
		Largest	Std. dev.	11.73726
75%	58	70		
90%	65	75	Variance	137.7634
95%	70	76	Skewness	.4857371
99%	78	78	Kurtosis	2.686481

.

. regress gaf_score_16 i.group i.sex ZAN_DIKO c. gaf_score_base, vce(robust)

Linear regression	Number of obs	=	131
	F(4, 126)	=	4.40
	Prob > F	=	0.0023
	R-squared	=	0.1388

Root MSE = 12.822

gaf_score_16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
2.group	-2.250969	2.205682	-1.02	0.309	-6.615948	2.114011
1.sex	-.1908501	3.031257	-0.06	0.950	-6.189618	5.807918
ZAN_DIKO	2.601935	2.699088	0.96	0.337	-2.73948	7.94335
gaf_score_base	.7476935	.1889833	3.96	0.000	.3737012	1.121686
_cons	12.6166	11.67247	1.08	0.282	-10.48289	35.71608

```
.
. regress gaf_score_16 i.group i.sex i.ZAN_DIKO c.gaf_score_base i.AGE_DIKO i.covid19, vce(ro
> bust)
```

Linear regression	Number of obs	=	131
	F(6, 124)	=	7.01
	Prob > F	=	0.0000
	R-squared	=	0.1962
	Root MSE	=	12.486

gaf_score_16	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
2.group	.4952688	2.333388	0.21	0.832	-4.12316	5.113697
1.sex	-2.812094	2.8466	-0.99	0.325	-8.446313	2.822126
2.ZAN_DIKO	2.462804	2.62453	0.94	0.350	-2.731875	7.657484
gaf_score_base	.7181524	.1771237	4.05	0.000	.3675749	1.06873
2.AGE_DIKO	-3.36284	2.236726	-1.50	0.135	-7.789949	1.064268
covid19						
Ja	-6.115811	2.438271	-2.51	0.013	-10.94183	-1.28979
_cons	22.3184	10.5346	2.12	0.036	1.467466	43.16933

```
. *8
```

```
. regress gaf_score_8 i.group i.sex ZAN_DIKO c. gaf_score_base, vce(robust)
```

```
Linear regression      Number of obs    =      148
                      F(4, 143)         =      5.34
                      Prob > F          =      0.0005
                      R-squared         =      0.1154
                      Root MSE       =     11.542
```

```
-----+-----
      |               Robust
gaf_score_8 | Coefficient  std. err.   t    P>|t|   [95% conf. interval]
-----+-----
      |
2.group | -1.281355   1.93958   -0.66  0.510   -5.115307   2.552597
1.sex | -2.633108   2.930063   -0.90  0.370   -8.424941   3.158725
ZAN_DIKO | -.9627205   2.468964   -0.39  0.697   -5.843102   3.917662
gaf_score_base | .5408578   .1478201    3.66  0.000    .2486629   .8330527
      _cons | 28.09435   10.20511    2.75  0.007    7.92199   48.2667
-----+-----
```

```
.
.
. *self_harm_16
```

```
. tabulate group self_harm_16
```

```
Randomiser |
ingsresult | Selvskade (ja/nej) 16
at 1, 2 |      0      1 |      Total
-----+-----+-----
      1 |      55      13 |      68
      2 |      45      16 |      61
-----+-----+-----
Total |      100      29 |      129
```

```
. by group, sort : inspect self_harm_16
```

```
-> group = 1
```

```
self_harm_16:  Selvskade (ja/nej) 16      Number of observations
```

		Total	Integers	Nonintegers
#	Negative	-	-	-
#	Zero	55	55	-
#	Positive	13	13	-
#				
#	Total	68	68	-
# #	Missing	16		
+-----				
0	1	84		
(2 unique values)				

```
-> group = 2
```

```
self_harm_16:  Selvskade (ja/nej) 16      Number of observations
```

		Total	Integers	Nonintegers
#	Negative	-	-	-
#	Zero	45	45	-
#	Positive	16	16	-
#				
# #	Total	61	61	-
# #	Missing	21		
+-----				
0	1	82		
(2 unique values)				

```
. logistic self_harm_16 i.group i.sex i.ZAN_DIKO, or
```

Logistic regression

Number of obs = 129

LR chi2(3) = 1.39

Prob > chi2 = 0.7080

Pseudo R2 = 0.0101

Log likelihood = -68.052394

self_harm_16	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
2.group	1.506078	.6400995	0.96	0.335	.6547528	3.464315
1.sex	.6570115	.4783031	-0.58	0.564	.1577248	2.736819
2.ZAN_DIKO	1.192751	.5617246	0.37	0.708	.4738901	3.002079
_cons	.3061566	.2463186	-1.47	0.141	.063257	1.481763

Note: _cons estimates baseline odds.

```
.  
. margins i.group, post
```

Predictive margins

Number of obs = 129

Model VCE: OIM

Expression: Pr(self_harm_16), predict()

	Margin	Delta-method std. err.	z	P> z	[95% conf. interval]	
group						
1	.1911976	.0476187	4.02	0.000	.0978668	.2845285
2	.262259	.0562108	4.67	0.000	.1520878	.3724301

```
.  
. nlcom _b[2bn. group]/_b[1. group]
```

_nl_1: _b[2bn. group]/_b[1. group]

	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
_nl_1	1.371664	.4507845	3.04	0.002	.4881427	2.255186

```
.
.
. *SAE

.
. tabulate group sae_yn, exact
```

Randomiser	Any SAE?		Total
ingsresult	Nej	Ja	
at 1, 2			
1	69	15	84
2	58	24	82
Total	127	39	166

```
Fisher's exact = 0.100
1-sided Fisher's exact = 0.060
```

```
.
. by group, sort : inspect sae_yn
```

```
-> group = 1
```

sae_yn: Any SAE?	Number of observations		
	Total	Integers	Nonintegers
# Negative	-	-	-
# Zero	69	69	-
# Positive	15	15	-
#			

	#		Total	84	84	-
	#	#	Missing	-		
+-----						
0		1		84		

(2 unique values)

sae_yn is labeled and all values are documented in the label.

-> group = 2

sae_yn: Any SAE?

			Number of observations		
			Total	Integers	Nonintegers
-----			-----	-----	-----
	#	Negative	-	-	-
	#	Zero	58	58	-
	#	Positive	24	24	-
	#		-----	-----	-----
	#	Total	82	82	-
	#	Missing	-		
+-----			-----		
0		1	82		

(2 unique values)

sae_yn is labeled and all values are documented in the label.

```
.
.
.
.
.
. log close
    name: <unnamed>
    log: /Users/janusjakobsen/Desktop/MBT_FINAL.log
    log type: text
    closed on: 10 Nov 2022, 21:11:58
-----
```

|

Posthoc statistical analyses for the MBT-RCT

Markus Harboe Olsen

2/28/23

Table of contents

SAE - overall	2
SAE - psychiatric hospitalization	2

SAE - overall

Summary data

	1	2
	(N=84)	(N=82)
as.factor(sae_binary)		
Ja	15 (17.9%)	24 (29.3%)
Nej	69 (82.1%)	58 (70.7%)

Analysis

Relative risk derived by G-computation
RR: 1.625 (sRR: 1.6173)
Confidence interval: 0.9377 - 3.0729
p-value: 0.0877

SAE - psychiatric hospitalization

Summary data

	1	2
	(N=84)	(N=82)
as.factor(psychiatrichospitalization_binary)		
0	75 (89.3%)	64 (78.0%)
1	9 (10.7%)	18 (22.0%)

Analysis

Relative risk derived by G-computation
RR: 2.0339 (sRR: 2.0222)
Confidence interval: 0.9883 - 5.0869
p-value: 0.056

Individual serious adverse events

Short-term MBT		Long-term MBT	
Participant	Events	Participant	Events
A1	1. Hospitalisation due to headache and pain in the neck and jaws (obs encephalitis) 2. Hospitalisation due to headache (2 nights)	B1	1. Admission to gynaecology ward (1 night)
A2	1. Hospitalisation*	B2	1. Hospitalisation due to appendicitis
A3	1. Hospitalisation (1 night) due to overdose of sertraline (2000 mg instead of 200 mg), dizziness and diarrhoea	B3	1. Hospitalisation (2 nights) due to medicine overdose (6 oxazepam) 2. Admission to a gastrointestinal surgery department
A4	1. Hospitalisation (21 nights) due to suicidal ideation	B4	1. Hospitalisation (1 night) due to abdominal pain 2. Hospitalisation (2 nights) due to suicide attempt (overdose of 70x tramadol 50 mg) 3. Admission to a gastrointestinal surgery department
A5	1. Hospitalisation (20 nights) due to suicidal ideation and self-harm	B5	1. Hospitalisation (1 night) due to suicide attempt (10-20 ibuprofen 400 mg)
A6	1. Hospitalisation (1 night) due to tick bite and hemiplegia	B6	1. Hospitalisation (1 month) due to suicidal ideation
A7	1. Hospitalisation (1 night) due to suicidal ideation. The patient is drunk and has taken cocaine 2. Hospitalisation (1 night) due to suicide attempt. The patient has taken 2 full packets of quetiapine 25 mg 3. Hospitalisation due to suicidal ideation 4. Hospitalisation due to suicide attempt (60x topiramate 100 mg)	B7	1. Suicidal ideation

	5. Suicide attempt (11 methylphenidate, 5 oxazepam, 1 tramadol).		
A8	1. Hospitalisation (1 night) due to drugs or alcohol	B8	1. Admission to psychiatric ward 2. Admission to psychiatric ward
A9	1. Hospitalisation (4 nights) due to suicidal ideation 2. Hospitalisation*	B9	1. Admission to neurological department (1 night)
A10	1. Hospitalisation (1 night) due to overdose of 60x sertraline 50 mg 2. Hospitalisation (10 nights) due to patient swallowing a battery 3. Hospitalisation (2 nights) due to suicidal ideation	B10	1. Hospitalisation (3 nights) due to anxiety and suicidal ideation
A11	1. Hospitalisation (2 nights) due to suicidal ideation	B11	1. Hospitalisation (2 nights) due to suicidal ideation 2. Hospitalisation (1 night) due to suicidal ideation 3. Hospitalisation (3 nights) due to suicidal ideation and severe depression
A12	1. Hospitalisation (1 night) due to seizures	B12	1. Hospitalisation (1 night) due to overdose of 14x sertraline 100 mg, 10x paracetamol, MDMA and cocaine. Drug-induced seizures. 2. Hospitalisation (1 night) due to suicide attempt and alcohol poisoning 3. Hospitalisation (1 night) due to suicide attempt 4. Hospitalisation (1 night) due to suicidal ideation and possibly psychosis. Under the influence of alcohol and drugs. 5. Hospitalisation (1 night) due to possible overdose of medicine. Unconscious at admission.

A13	1. Hospitalisation (1 night) due to suicide attempt (overdose) 2. Hospitalisation (2 nights) due to suicidal ideation	B13	1. Admission to psychiatric ward 2. Admission to psychiatric ward
A14	1. Hospitalisation (1 night) due to suicide attempt (overdose) 2. Hospitalisation (1 night) due to suicidal ideation	B14	1. Hospitalisation (1 night) due to abdominal pain 2. Hospitalisation (1 night) due to abdominal pain 3. Hospitalisation (1 night) due to abdominal pain and black stools 4. Hospitalisation (3 nights) due to suicidal ideation and anxiety 5. Hospitalisation (3 nights) due to suicidal ideation 6. Hospitalisation (3 nights) due to suicide attempt (75x promethazine 25 mg and 5x oxazepam 15 mg) 7. Hospitalisation (1 night) due to suicidal ideation
A15	1. Hospitalisation*	B15	1. Hospitalisation (4 nights) due to depressive symptoms
		B16	1. Hospitalisation (1 night) due to severe headache
		B17	1. Hospitalisation (3 nights) due to suicide attempt (overdose) 2. Hospitalisation (1 night) due to shortness of breath 3. Hospitalisation (1 night) due to suicidal ideation 4. Hospitalisation (8 nights) due to suicidal ideation

		B18	1. Hospitalisation (1 night) due to suicidal ideation
		B19	1. Hospitalisation (5 nights) due to suicidal ideation and depressive symptoms
		B20	1. Hospitalisation (6 nights) due to possible symptoms of bipolar disorder
		B21	1. Electric shock (230v)
		B22	1. Hospitalisation (1 night) due to chest pain
		B23	1. Admission to psychiatric ward
		B24	1. Hospitalisation due to suicidal ideation 2. Hospitalisation due to suicidal ideation 3. Hospitalisation due to suicide attempt (overdose of quetiapine)

*Unclear

Paper 7

Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder: 24 Months Follow-Up of a Randomized Clinical Trial

Sophie Juul^{a, b, c} Janus Christian Jakobsen^{a, d} Emilie Hestbaek^{b, c}
Caroline Barkholt Kamp^{a, d} Markus Harboe Olsen^{a, e} Marie Rishede^f
Frederik Weischer Frandsen^{b, c} Sune Bo^c Stig Poulsen^c Per Sørensen^b
Anthony Bateman^g Sebastian Simonsen^{b, c}

^aCopenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ^bMental Health Center Stolpegård, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark; ^cDepartment of Psychology, University of Copenhagen, Copenhagen, Denmark; ^dDepartment of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; ^eDepartment of Neuroanaesthesiology, The Neuroscience Centre, The Neuroscience Centre, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark; ^fDepartment of Psychiatry Mental Health Services East, Region Zealand, Roskilde, Denmark; ^gDepartment of Clinical, Educational, and Health Psychology, University College London, London, UK

Keywords

Mentalization-based therapy · Borderline personality disorder · Short-term psychotherapy · Randomized clinical trial

Abstract

Introduction: Borderline personality disorder (BPD) is a severe and prevalent psychiatric disorder. Mentalization-based therapy (MBT) is an evidence-based intervention for BPD, which is often delivered as a long-term psychotherapy program for BPD. We previously published a randomized clinical trial assessing short-term versus long-term MBT for BPD 16 months after randomization as the primary follow-up time point. The objective was to assess the long-term (24 months) results of short-term versus long-term MBT for outpatients with BPD. **Methods:** Adult out-

patients (≥ 18 years) with subthreshold or diagnosed BPD were randomly assigned (1:1) to short-term MBT (5 months) or long-term MBT (14 months). The primary outcome was BPD symptoms assessed with the Zanarini Rating Scale for Borderline Personality Disorder. Secondary outcomes were level of functioning (assessed using the Work and Social Adjustment Scale), quality of life (assessed using Short Form Health Survey, SF-36), global functioning (assessed using the Global Assessment of Functioning, GAF scale), and severe self-harm. All outcomes were assessed at 24 months after randomization. **Results:** Between October 4, 2018, and December 3, 2020, we randomly assigned 166 participants to short-term MBT ($n = 84$) or long-term MBT ($n = 82$). After

Clinical trial registration: This trial was prospectively registered at ClinicalTrials.gov, NCT03677037.

24 months, regression analyses showed no evidence of a difference when assessing the primary outcome, BPD symptoms (ZAN-BPD MD: -0.56 ; 95% CI: -2.67 to 1.54 ; $p = 0.598$), level of functioning (WSAS MD: -1.42 ; 95% CI: -5.04 to 2.21 ; $p = 0.440$), global functioning (GAF MD: 2.51 ; 95% CI: -1.65 to 6.67 ; $p = 0.234$), or severe self-harm (RR: 1.38 ; 95% CI: 0.88 – 2.21 ; $p = 0.149$). Regression analyses showed evidence of a beneficial effect of long-term MBT when assessing quality of life at 24 months (SF-36 MD: 5.52 ; 95% CI: 1.72 – 9.32 ; $p = 0.005$), but the result was potentially biased by a large proportion of missing data (40%). We found no evidence of a difference between short-term and long-term MBT on serious adverse events at 24 months (RR: 1.38 ; 95% CI: 0.88 – 2.21 ; $p = 0.149$). The serious adverse event result had no missing data. **Conclusion:** Long-term MBT did not lead to lower levels of BPD symptoms at 24 months after randomization compared with short-term MBT. More trials comparing short-term versus long-term treatment and with fewer missing data are needed to provide better evidence-based guidance.

© 2025 S. Karger AG, Basel

Introduction

Borderline personality disorder (BPD) is a severe mental health disorder affecting up to 6% of the general population [1–4]. BPD is characterized by a pervasive pattern of instability in interpersonal relationships and self-image, emotional dysregulation, intense anger and is associated with high rates of suicidal and self-harming behavior [5, 6]. Psychotherapy is recommended as the primary treatment of choice for BPD [7, 8]. Even though several short-term psychotherapies exist [9, 10], long-term psychotherapy (a year or more) is the most widely used and evidence-based treatment for BPD [11, 12].

Mentalization-based therapy (MBT) is an example of a long-term, evidence-based psychotherapy approach rooted in psychodynamic and attachment theory, developed specifically to treat BPD [13]. MBT is based on the theory of mentalizing, which refers to the process by which we make sense of ourselves and others, implicitly and explicitly, in terms of mental states [14]. MBT is originally designed as a long-term program, with 18 months of psychotherapy. In a recent Cochrane review of all psychological therapies for BP, the authors concluded that long-term MBT was more effective than treatment as usual in reducing self-harm, suicidality, and depression [11]. Six trials assessing the effects of MBT were included in the meta-analysis. The follow-up time points of these trials ranged from 12 months [15] after randomization to 8 years [16, 17] after randomization.

We recently published a randomized clinical trial comparing a new short-term MBT program (5 months) with the existing long-term MBT program (14 months) for outpatients with BPD implemented in the Mental Health Services of the Capital Region of Denmark [18]. The primary follow-up time point was 16 months after randomization. In this trial, we found no evidence of a difference between short-term versus long-term MBT for outpatients with BPD on any of our primary or secondary outcomes at 16 months after randomization [18]. More participants in the long-term MBT group had a serious adverse event primarily driven by a difference in psychiatric hospitalizations, but the result was not statistically significant. The long-term effects beyond 16 months after randomization is currently unclear. The aim of the present paper was to present long-term results (24 months after randomization) for our previously published randomized clinical trial assessing the beneficial and harmful effects of short-term versus long-term MBT for outpatients with BPD [18].

Materials and Methods

This trial is reported in accordance with the Consolidated Standards of Reporting Trials (CONSORT) 2010 statement [19] (online suppl. Material 1; for all online suppl. material, see <https://doi.org/10.1159/000544934>). The trial was prospectively registered at ClinicalTrials.gov (NCT03677037, registered September 19, 2018) before the first participant was randomized. The methodology was predefined and described in detail previously [20, 21].

Trial Design and Participants

The trial design and participants have been described in detail previously [18, 20, 21]. In short, the trial was as investigator-initiated, parallel-group, single-center, randomized clinical superiority trial. Outpatients from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre in the Mental Health Services of the Capital Region of Denmark were consecutively screened for eligibility to the trial during the trial period.

Eligible participants were adults (≥ 18 years) with subthreshold (defined as at least 4 criteria) or diagnosed BPD according to the diagnostic criteria provided in the Diagnostic and Statistical Manual for Mental Disorders-5th edition (DSM-5) [5], who provided written informed consent. Exclusion criteria were reduced intellectual functioning (IQ < 75), a full diagnosis of schizotypal personality disorder or antisocial personality disorder

according to DSM-5, a comorbid mental health disorder requiring specialist treatment, current (past 2 months) substance dependence assessed with the Mini International Neuropsychiatric Interview (MINI) [22], concurrent psychotherapeutic treatment, inability to understand Danish, and lack of informed consent.

The trial was overseen by a steering committee and by an independent data and safety monitoring committee. The trial was approved by the Danish Data Protection Agency (I-suite: 6553; August 30, 2018) and the Regional Ethics Committee of Health Research of the Capital Region of Denmark (H-18023136; August 27, 2018).

Randomization and Masking

The randomization and masking procedures have been described previously [18, 20, 21]. In short, an independent investigator randomly assigned participants (1:1) to short-term or long-term MBT according to a computer-generated allocation sequence with permuted blocks of various sizes. The allocation sequence was stratified by sex (male/female) and low or high (0–11/12–36) BPD symptom severity at baseline assessed with the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) [23].

Outcome assessors, the data safety and monitoring committee, data managers, statisticians, and the steering committee were blind to treatment allocation [20, 21, 24]. Data were unblinded on December 18, 2022, after the statistical report was finalized, and after the steering group had agreed on two blinded abstracts. The statistical analyses presented in this follow-up paper were conducted after the data had been unblinded.

Procedures

The procedures have been described in detail previously [18, 20, 21]. In short, patients were initially screened for eligibility during routine intake interviews at the Outpatient Clinic for Personality Disorders. Potential eligible and consenting patients were assessed with the MINI interview [22] and the Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD) [25] to confirm eligibility.

Short-term MBT was organized as a 5 months program of therapy in closed groups commencing with five sessions of introduction to MBT followed by 15 weekly sessions of group MBT. Group therapy was accompanied by individual sessions every second week. The participants were treated by two group therapists, one of them also being the individual therapist, i.e., conjoined therapy. The MBT therapists followed the treatment manual developed by Bateman and Fonagy [13].

Long-term MBT was organized as a 14-month program. All participants randomized to long-term MBT initially entered a 6-week introduction to MBT (primarily psychoeducation). When the introduction group finished, participants were allocated to one of eight slow-open MBT treatment groups. Treatment was then organized as 12 months of weekly group therapy sessions combined with individual therapy every second week. The long-term MBT group consisted of a total of 13.5 months of active treatment, but since there was some waiting time from the MBT introduction group to the beginning of the MBT treatment group, we pragmatically defined the long-term MBT intervention as 14 months of duration. The participants were treated by two group therapists and a third individual therapist, i.e., combined therapy. Both group and individual therapy was manualized by Bateman and Fonagy [13]. Three external MBT experts, who were blind to treatment allocation, assessed adherence to the MBT manual in both short-term and long-term MBT using the Mentalization-Based Treatment Adherence and Competence scale [26], yielding good adherence to the manual in both groups [18].

Outcomes

The primary outcome was severity of borderline symptomatology assessed with the ZAN-BPD interview by a blinded assessor [23]. Secondary outcomes were functional impairment assessed with the Work and Social Adjustment Scale (WSAS) [27], quality of life assessed with the Short-Form Health Survey (SF-36) mental component [28], global functioning assessed with the Global Assessment of Functioning (GAF) scale [29], and severe self-harm defined as the proportion of participants with one or more deliberate act of self-harm resulting in visible tissue damage (not including less severe self-harm, e.g., head banging, superficial cutting, or hitting oneself without tissue damage). All outcomes were assessed at 8, 16, and 24 months after randomization.

Safety was assessed as serious adverse events at 16 months and again at 24 months after randomization. Two blinded assessors independently went through patient medical records to assess serious adverse events using the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use-Good Clinical Practice (ICH-GCP) definition of a serious adverse event (any untoward medical occurrence that resulted in death, was life threatening, required hospitalization or prolonging of existing hospitalization and resulted in persistent or significant disability, or jeopardized the participant) [30].

Statistical Analysis

Our published and predefined statistical analysis plan, which was also used for this follow-up paper, provides a detailed explanation of our statistical methodology [21]. In short, we estimated that a total sample size of 166 participants would be needed to provide 80% power at a 5% significance level, with 3.5 points as the minimal important difference for the primary outcome, ZAN-BPD.

Missing data were handled according to the recommendations by Jakobsen et al. [31]. To assess if data were comparable at baseline, we visually inspected baseline tables to identify any imbalances between groups among participants with and without data available at follow-up. We did not perform significance testing for the baseline tables as this is not recommended by the CONSORT statement [19, 32, 33].

All continuous outcomes were analyzed using linear regression and dichotomous outcomes were analyzed using logistic regression. The analyses were based on an intention-to-treat population and were primarily adjusted for the baseline value of the outcome of interest and the stratification variables used in the randomization. We used a five-step procedure [34] to assess if the thresholds for statistical and clinical significance were crossed, and we handled missing data according to the procedure suggested by Jakobsen et al. [31].

Results

Between October 4, 2018, and December 3, 2020, of the 207 patients initially screened for trial eligibility, we randomly assigned 166 participants to short-term MBT ($n = 84$) or long-term MBT ($n = 82$) (Fig. 1). Baseline clinical and sociodemographic characteristics of the trial population have been published previously [18]. Visual inspection of baseline characteristics for participants with complete data at 24 months on the primary outcome revealed substantial imbalances between the groups, for example, regarding if they were in a relationship (living together) (short-term MBT: 20.6% vs. long-term MBT: 14%), living alone (41.3% vs. 35.1%), on social welfare (/unemployed) (41.3% vs. 59.6%), had co-occurring anxiety disorders (49.2% vs. 63.2%), and avoidant personality disorder (30.2% vs. 21.6%), as detailed in online supplementary Material 2. Similarly, visual inspection of baseline characteristics for participants lacking complete data for the secondary outcomes at 24 months showed apparent imbalances (online suppl. Material 2). These findings from the baseline tables suggest that the available

data at 24 months were based on participants who were not comparable at baseline.

At 24 months after randomization, linear regression showed no evidence of a difference between short-term and long-term MBT when assessing borderline symptoms (ZAN-BPD MD: -0.56 ; 95% CI: -2.67 to 1.54 ; $p = 0.598$) (Table 1). Twenty one (25%) of 84 participants in the short-term MBT group were lost to follow-up compared with 25 (30.5%) of 82 participants in the long-term MBT group. Sensitivity analyses showed that missing outcome data had the potential to alter the results (online suppl. Material 3). None of the pre-planned subgroup analyses (baseline severity of borderline symptoms; sex; age; baseline global functioning; baseline proportion of participants with severe self-harm; proportion of participants who had their group therapy temporarily paused due to COVID-19) showed evidence of a difference (Fig. 2).

At 24 months after randomization, linear regression showed no evidence of a difference between short-term and long-term MBT when assessing level of functioning (WSAS MD: -1.42% ; 95% CI: -5.04 to 2.21 ; $p = 0.440$). Twenty eight (33.3%) of 84 participants in the short-term MBT group were lost to follow-up compared with 35 (42.7%) of 82 participants in the long-term MBT group (Table 1).

At 24 months after randomization, linear regression showed evidence of a beneficial effect of long-term MBT when assessing quality of life (SF-36 MD: 5.52 ; 95% CI: 1.72 – 9.32 ; $p = 0.005$). Twenty eight (33.3%) of 84 participants in the short-term MBT group were lost to follow-up compared with 35 (42.7%) of 82 participants in the long-term MBT group. As mentioned, visual inspection of the baseline characteristics of participants with complete SF-36 data at 24 months showed considerable baseline imbalances (online suppl. Material 4).

At 24 months after randomization, a Wilcoxon rank-sum test showed no evidence of a difference between short-term and long-term MBT when assessing global functioning (GAF MD: 2.51 ; 95% CI: -1.65 to 6.67 ; $p = 0.234$). The median (interquartile range: IQR) in the short-term MBT group was 48 (IQR: 41.00 ; 54.0), and the median (IQR) in the long-term group was 49.5 (IQR: 40.75 ; 58.25) (online suppl. Material 3). Twenty one (25%) of 84 participants in the short-term MBT group were lost to follow-up compared with 26 (31.7%) of 82 participants in the long-term MBT group (Table 1).

At 24 months after randomization, 10 of 63 (15.9%) participants performed self-harm in the short-term MBT group compared with 11 of 57 (19.3%) participants in the

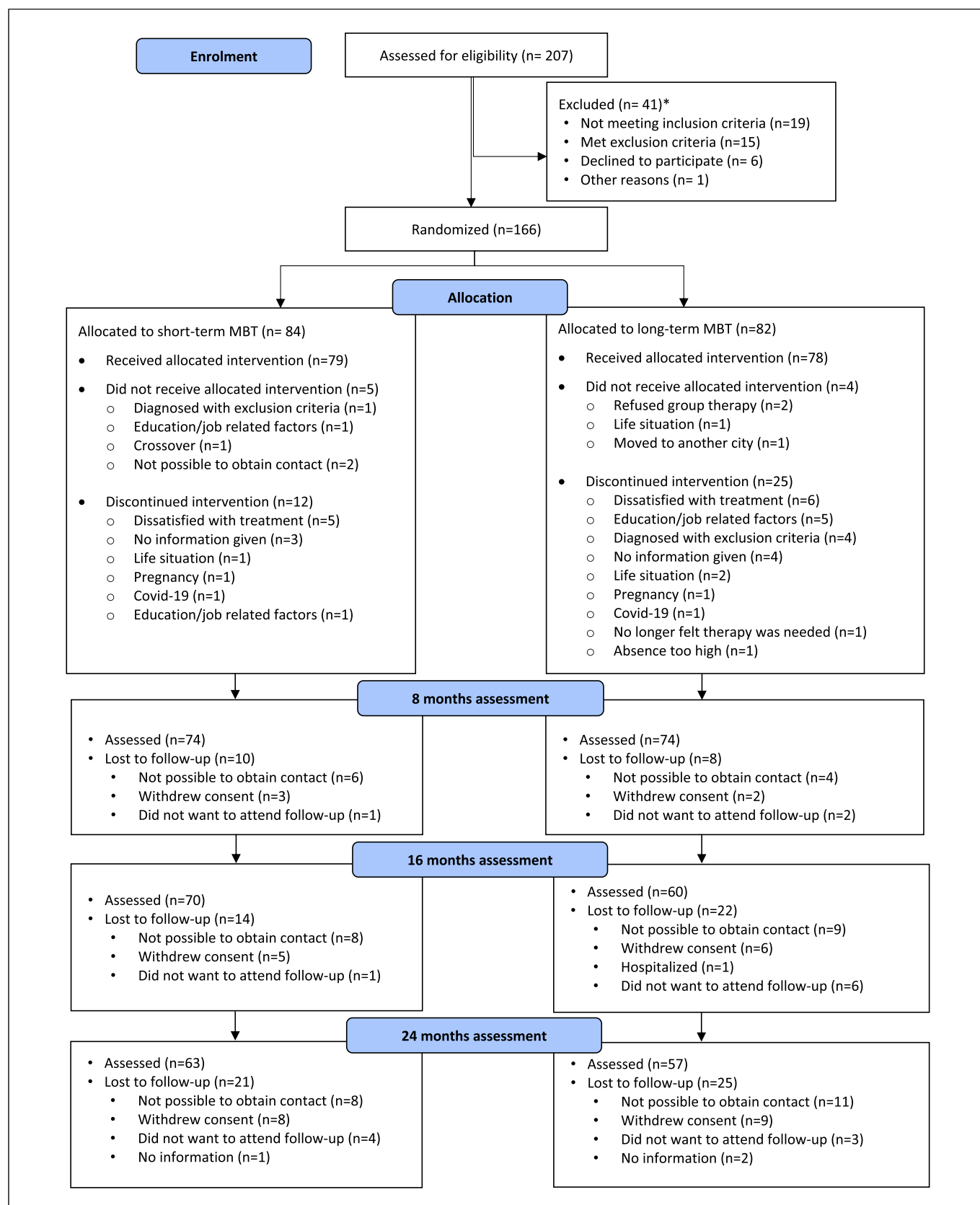


Fig. 1. CONSORT flowchart.

Table 1. Primary and secondary outcome results (ITT population)

Outcome	Short-term MBT (n = 84)		Long-term MBT (n = 82)		Estimate (95% CI)	p value
	No. analyzed	result	No. analyzed	result		
Primary						
Borderline symptoms, ZAN-BPD, mean (SD)						
Baseline	84	13.9 (3.72)	82	13.8 (3.75)		
8 months	74	12.0 (5.33)	74	12.8 (5.01)	MD 0.95 (−0.57 to 2.46)	0.22
16 months	70	11.3 (6.22)	60	12.3 (5.91)	MD 0.99 (−1.06 to 3.03)	0.341
24 months	63	10.9 (6.58)	57	10.2 (5.49)	MD −0.56 (−2.67 to 1.54)	0.598
Secondary						
Functional impairment, WSAS, mean (SD)						
Baseline	84	22.9 (7.54)	82	21.9 (7.46)		
8 months	71	20.4 (8.59)	69	19.9 (8.36)	MD 0.50 (−1.95 to 2.95)	0.687
16 months	63	16.7 (8.91)	56	17.2 (9.36)	MD 1.44 (−1.43 to 4.32)	0.321
24 months	56	17.5 (10.8)	47	15.1 (10.3)	MD −1.42 (−5.04 to 2.21)	0.440
Quality of life, SF-36 – mental component, mean (SD)						
Baseline	84	50.3 (10.4)	82	49.2 (10.1)		
8 months	71	50.4 (10.8)	69	53.1 (11.6)	MD 2.97 (−0.78 to 6.72)	0.120
16 months	63	53.9 (9.94)	56	52.9 (10.2)	MD −0.91 (−4.62 to 2.79)	0.626
24 months	56	49.1 (9.1)	47	54.4 (10.3)	MD 5.52 (1.72 to 9.32)	0.005
Global functioning, GAF, mean (SD)						
Baseline	84	46.5 (8.14)	82	46.3 (6.36)		
8 months	74	49.6 (13.5)	74	47.7 (10.5)	MD −1.28 (−5.05 to 2.49)	0.502
16 months	70	52.2 (15.0)	61	49.7 (11.7)	MD −2.25 (−6.70 to 2.20)	0.318
24 months	63	48.4 (11.2)	56	50.7 (13.4)	MD 2.51 (−1.65 to 6.67)	0.234
Proportion of participants with severe self-harm the past 8 months						
Baseline	84	32/84 (38.1%)	82	33/82 (40.2%)		
8 months	74	19/74 (25.7%)	74	27/74 (36.5%)	RR 1.43 (0.87 to 2.45)	0.153
16 months	68	13/68 (19.1%)	61	16/61 (26.2%)	RR 1.37 (0.70 to 2.84)	0.335
24 months	63	10/63 (15.9%)	57	11/57 (19.3%)	RR 1.24 (0.52 to 2.88)	0.588

GAF, Global Assessment of Functioning; ITT, intention to treat; MD, mean difference; OR, odds ratio; SF-36, Short Form Health Survey-36; SD, standard deviation; WSAS, Work and Social Adjustment Scale; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

long-term MBT group (RR: 1.24; 95% CI: 0.52–2.88; $p = 0.588$). Twenty one (25%) of 84 participants in the short-term MBT group were lost to follow-up compared with 25 (30.5%) of 82 participants in the long-term MBT group (Table 1).

At 24 months after randomization, 23 of 84 (27.4%) participants experienced a serious adverse event in the short-term MBT group compared with 31 of 82 (37.8%) of the participants in the long-term MBT group (RR: 1.38; 95% CI: 0.88–2.21; $p = 0.149$) (online suppl. Material 3). One participant in the short-term MBT group died by suicide versus no participants in the long-term group. A table of individual serious adverse events can be found in online supplementary Material 5. The serious adverse event result had no missing data.

Discussion

This is the first randomized clinical trial assessing short-term versus long-term MBT for adult outpatients with BPD. Our analyses showed that at 24 months after randomization, no evidence of a difference was found when assessing BPD symptomatology, level of functioning, global functioning, and self-harm. We found indications of a beneficial effect of long-term MBT when assessing quality of life at 24 months after randomization, and the result reached our predefined threshold for a minimal clinically important difference. This result was the only statistically significant result. Given that our results otherwise have been consistent across time points and outcomes, we hypothesize that this result on quality

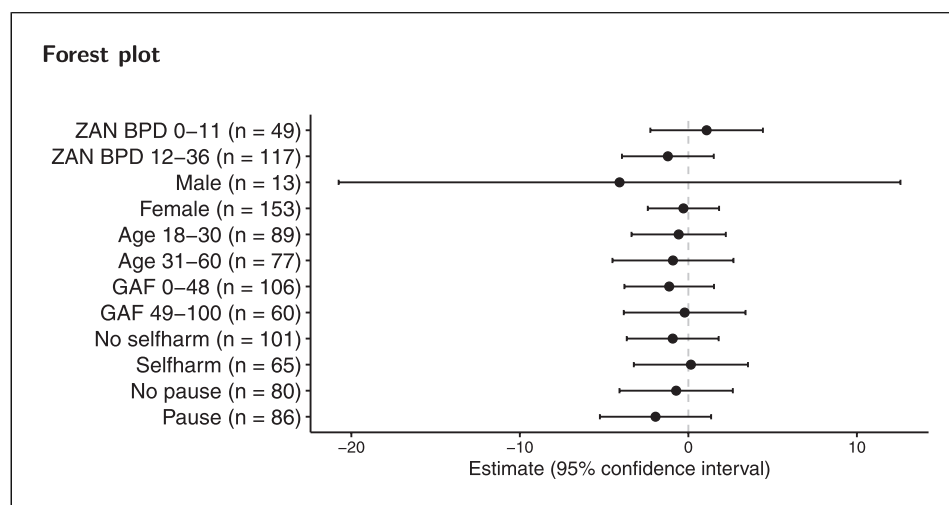


Fig. 2. Forest plot of subgroup analyses for the primary outcome at 24 months.

of life at 24 months after randomization may be a type 1 error. Another explanation for the significant results could be that missing data approached 40%. Visual inspection of baseline characteristics for participants with complete data at 24 months revealed substantial imbalances between the groups. This indicates that missing outcome data, rather than the trial interventions, may account for the difference between the groups on quality of life at 24 months follow-up. We found no evidence of a difference between short-term and long-term MBT on serious adverse events at 24 months. There were no missing data for the serious adverse event result.

Our present results have several strengths. First, we used the same rigorous trial methodology as in our main publication to reduce the risks of bias, and our methodology was pre-defined in our published protocol and detailed statistical analysis plan [20, 21]. Second, we assessed both beneficial and harmful effects and included objective assessments of serious adverse events, which is rare in psychotherapy trials for BPD [11]. Third, our trial was conducted in a real-world outpatient setting, and we followed the routine eligibility criteria used in the clinic in everyday practice to increase the external validity of the results. Therefore, our results do not apply to patients needing acute care for other disorders than BPD. Furthermore, the excellent compliance with MBT suggests that this approach could be acceptable for implementation in routine clinical use with the appropriate training and supervision. Fourth, all participants were assessed at the same 24-month time point regardless of their allocated intervention.

The primary limitation of our trial is, as mentioned, the large proportion of missing data 24 months after

randomization on our primary and secondary outcomes. The higher the proportions of missing outcome data in a trial, the more problematic it is to handle statistically and the more biased the conclusions potentially become [35]. When data are missing in a randomized clinical trial, the primary concern is whether the available data lead to biased results, and the large proportions of missing data can potentially bias the present results. Second, the participants only received individual therapy every second week compared to weekly sessions in the original program [18, 20]. However, this scheduling was consistent for both the short- and long-term participants in our trial, ensuring any effects of the less intensive program would balance between groups and not influence the relative effect, i.e., the effect estimates [36]. Third, due to the COVID-19 pandemic, some participants in both groups received less group therapy than initially intended. However, our subgroup analyses did not suggest that the COVID-19 pause of group therapy influenced the results. Fourth, during the trial period, once the participants finalized their allocated trial intervention, they were not offered any additional specialized care for BPD at the clinic. However, we did not systematically record what other types of services they received (e.g., social services, supportive care, and psychotherapy for comorbidities), which is a limitation.

Six previous smaller trials have assessed MBT versus control interventions for people with BPD [15, 37–41]. Some of the previous trials have indicated that longer-term MBT reduces the risk of self-harm and suicidality compared with control interventions [11], but these previous trials are at high risk of bias [11], had small

sample sizes [15, 37–41], and are likely underpowered to assess hard outcomes [15, 37–42].

Our trial of short-term versus long-term MBT is the first to use blinded outcome assessments of BPD severity both at baseline and 24 months after randomization. It is the largest trial of MBT until today. Overall, the results of our trial are similar to another recent larger trial assessing the effects of 6 versus 12 months of DBT for BPD [43], which also assessed the effects at 24 months after randomization. This trial randomized a total of 240 participants. The authors concluded that the noninferiority hypothesis was supported by the primary outcome, total self-harm, and several secondary outcomes. We plan to conduct an individual patient data meta-analysis with the results of both our trials to investigate the potential differential effects of short-term versus long-term psychotherapy for outpatients with BPD in different subgroups of participants.

Conclusions

Long-term MBT did not lead to lower levels of BPD symptoms at 24 months after randomization compared with short-term MBT. More trials comparing short-term versus long-term treatment and with fewer missing data are needed to provide better evidence-based guidance.

Acknowledgments

We thank the trial participants and the trial clinicians from the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark; the independent members of the Data Safety and Monitoring Committee (Prof. Ole Jakob Storebø and Prof. Mickey Kongerslev); our student assistants (Mathilde Hasselby-Andersen, Anders Sonne Munch, Amanda Ark Søndergaard, Laura Alsing Juul, Rebecca Engel Thomas, Rebecca Kjær Andersen, and Olivia Frost); and the independent coders of MBT adherence (Nana Lund Nørgaard and Anne Blom Corlin).

Statement of Ethics

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Research Boards (IRBs) of the Mental Health Services of the Capital region of Denmark (Approval No. H-18023136; August 27, 2018) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All participants provided written informed consent.

Conflict of Interest Statement

A.B. declares receiving book royalties from Oxford University Press and American Psychiatric Publishing Inc. for publishing books on mentalizing and payment for training courses in MBT from Anna Freud Centre, London. No other disclosures were reported.

Funding Sources

The trial was funded by Trygfonden (S.J., Grant No. 123488) and by the Mental Health Services in the Capital Region of Denmark Research Fund (S.J. and S.S., Grant No. N/A). The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had access to all study data and had final responsibility for the decision to submit for publication.

Author Contributions

S.J. and S.S. conceptualized the trial, secured the funding, designed the short-term MBT intervention in collaboration with A.B. and clinicians at the outpatient clinic, and was the principal investigator (S.J.) and sponsor (S.S.) of the trial. P.S. provided organizational support for the trial. F.W.F. assisted in the clinical implementation of the trial. The trial was overseen by a steering committee consisting of S.J., S.P., P.S., J.C.J., S.L., and S.S. S.J. recruited participants and performed baseline assessments assisted by E.H., C.B.K., and M.R. Blinded outcome assessments were performed by E.H., C.B.K., M.R., and several student assistants. A.B. and S.B. trained and supervised the staff in both short-term and long-term MBT. J.C.J. supervised the design of the trial protocol and statistical analysis plan. M.H.O. and J.C.J. independently performed blinded statistical analyses. S.J., J.C.J., and M.H.O. had full access to all the blinded trial data and take responsibility for the integrity of the data and the accuracy of the data analysis. S.B. assisted in the coding of adherence to MBT. S.J., J.C.J., and S.S. drafted the paper. All authors commented on and approved the final version of the paper.

Data Availability Statement

Data from this trial are not publicly available due to ethical grounds. Access to deidentified data from this trial may be obtained by contacting the corresponding author. All requests for data access must be approved by the IRBs of the Mental Health Services in the Capital Region of Denmark and the institution with whom the primary investigator is affiliated. Data access will be governed by a data sharing agreement executed between the Capital Region of Denmark and the institution with whom the primary investigator is affiliated.

References

- 1 Torgersen J. Prevalence, sociodemographics, and functional impairment. In: Oldham J, Skodol AE, Bender DS, editors. *Textbook of personality disorders*. Washington DC: American Psychiatric Publishing; 2014. Vol. 2. p. 109–31.
- 2 Grant BF, Chou SP, Goldstein RB, Huang B, Stinson FS, Saha TD, et al. Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the wave 2 national epidemiologic survey on alcohol and related conditions. *J Clin Psychiatry*. 2008;69(4):533–45. <https://doi.org/10.4088/jcp.v69n0404>
- 3 Ellison WD, Rosenstein LK, Morgan TA, Zimmerman M. Community and clinical epidemiology of borderline personality disorder. *Psychiatr Clin North Am*. 2018;41(4):561–73. <https://doi.org/10.1016/j.psc.2018.07.008>
- 4 Zimmerman M, Rothschild L, Chelminski I. The prevalence of DSM-IV personality disorders in psychiatric outpatients. *Am J Psychiatry*. 2005;162(10):1911–8. <https://doi.org/10.1176/appi.ajp.162.10.1911>
- 5 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub; 2013. Report No.: 0890425574.
- 6 Black DW, Blum N, Pfohl B, Hale N. Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *J Pers Disord*. 2004;18(3):226–39. <https://doi.org/10.1521/pedi.18.3.226.35445>
- 7 Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385(9969):735–43. [https://doi.org/10.1016/S0140-6736\(14\)61394-5](https://doi.org/10.1016/S0140-6736(14)61394-5)
- 8 National Institute for Health and Care Excellence. *Borderline personality disorder: recognition and management [CG78]*; 2009. Available from: <https://www.nice.org.uk/guidance/cg78/resources/2018-surveillance-of-personality-disorders-nice-guidelines-cg77-and-cg78-4906490080/chapter/Surveillance-decision?tab=evidence>
- 9 Gratz KL, Tull MT, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099–112. <https://doi.org/10.1017/S0033291713002134>
- 10 Blum N, St John D, Pfohl B, Stuart S, McCormick B, Allen J, et al. Systems training for emotional predictability and problem solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *Am J Psychiatry*. 2008;165(4):468–78. <https://doi.org/10.1176/appi.ajp.2007.07071079>
- 11 Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020(5):CD012955. <https://doi.org/10.1002/14651858.CD012955.pub2>
- 12 Simonsen S, Bateman A, Bohus M, Dalewijk HJ, Doering S, Kaera A, et al. European guidelines for personality disorders: past, present and future. *Borderline Personal Disord Emot Dysregul*. 2019;6(1):9–10. <https://doi.org/10.1186/s40479-019-0106-3>
- 13 Bateman A, Fonagy P. *Mentalization-based treatment for personality disorders: a practical guide*. Oxford University Press; 2016.
- 14 Bateman A, Fonagy P. Mentalization based treatment for borderline personality disorder. *World Psychiatry*. 2010;9(1):11–5. <https://doi.org/10.1002/j.2051-5545.2010.tb00255.x>
- 15 Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2012;51(12):1304–13. e3. <https://doi.org/10.1016/j.jaac.2012.09.018>
- 16 Bateman A, Fonagy P. 8-year follow-up of patients treated for borderline personality disorder: mentalization-based treatment versus treatment as usual. *Am J Psychiatry*. 2008;165(5):631–8. <https://doi.org/10.1176/appi.ajp.2007.07040636>
- 17 Bateman A, Constantinou MP, Fonagy P, Holzer S. Eight-year prospective follow-up of mentalization-based treatment versus structured clinical management for people with borderline personality disorder. *Personal Disord*. 2021;12(4):291–9. <https://doi.org/10.1037/per0000422>
- 18 Juul S, Jakobsen JC, Hestbaek E, Jørgensen CK, Olsen MH, Rishede M, et al. Short-term versus long-term mentalization-based therapy for borderline personality disorder: a randomized clinical trial (MBT-RCT). *Psychother Psychosom*. 2023;92(5):329–39. <https://doi.org/10.1159/000534289>
- 19 Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*. 2010;152(11):726–32. <https://doi.org/10.7326/0003-4819-152-11-201006010-00232>
- 20 Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196. <https://doi.org/10.1186/s13063-019-3306-7>
- 21 Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, 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;22(1):497. <https://doi.org/10.1186/s13063-021-05450-y>
- 22 Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59(Suppl 20):22–57.
- 23 Zanarini MC, Vujanovic AA, Parachini EA, Boulanger JL, Frankenburg FR, Hennen J. Zanarini rating scale for borderline personality disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Pers Disord*. 2003;17(3):233–42. <https://doi.org/10.1521/pedi.17.3.233.22147>
- 24 Juul S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med*. 2021;26(3):109. <https://doi.org/10.1136/bmjebm-2020-111407>
- 25 First MB, Williams JB, Karg RS, Spitzer RL. *User's guide for the SCID-5-CV structured clinical interview for DSM-5® disorders: clinical version*. American Psychiatric Publishing, Inc.; 2016.
- 26 Bateman A, Bales D. MBT. HJAQMf. A quality manual for MBT; 2020. Available from: <https://www.annafreud.org/media/14278/adherence-scale-august-2020.pdf> (Accessed January 10, 2023).
- 27 Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personal Ment Health*. 2017;11(4):215–28. <https://doi.org/10.1002/pmh.1382>
- 28 Brazier JE, Harper R, Jones N, O'cathain A, Thomas K, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160–4. <https://doi.org/10.1136/bmj.305.6846.160>
- 29 Hall RC. Global assessment of functioning: a modified scale. *Psychosom*. 1995;36(3):267–75. [https://doi.org/10.1016/S0033-3182\(95\)71666-8](https://doi.org/10.1016/S0033-3182(95)71666-8)
- 30 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised guideline: integrated addendum to ICH E6(R1): guideline for good clinical practice (ICH-GCP); 2015.
- 31 Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162. <https://doi.org/10.1186/s12874-017-0442-1>

- 32 Sherry AD, Msaouel P, McCaw ZR, Abi Jaoude J, Hsu EJ, Kouzy R, et al. Prevalence and implications of significance testing for baseline covariate imbalance in randomised cancer clinical trials: the Table 1 Fallacy. *Eur J Cancer*. 2023;194:113357. <https://doi.org/10.1016/j.ejca.2023.113357>
- 33 de Boer MR, Waterlander WE, Kuijper LDJ, Steenhuis IHM, Twisk JWR. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *Int J Behav Nutr Phys Act*. 2015;12(1):4. <https://doi.org/10.1186/s12966-015-0162-z>
- 34 Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol*. 2014;14(34):34. <https://doi.org/10.1186/1471-2288-14-34>
- 35 Altman DG, Bland JM. Missing data. *BMJ*. 2007;334(7590):424. <https://doi.org/10.1136/bmj.38977.682025.2C>
- 36 Juul S, Jakobsen JC, Hestbaek E, Kamp CB, Olsen MH, Rishede M, et al. Response to Philips et al.: Several Fundamental Misconceptions about Evidence-Based Practice. *Psychother Psychosom*. 2024;93(5):348–9. <https://doi.org/10.1159/000540903>
- 37 Philips B, Wennberg P, Konradsson P, Franck J. Mentalization-based treatment for concurrent borderline personality disorder and substance use disorder: a randomized controlled feasibility study. *Eur Addict Res*. 2018;24(1):1–8. <https://doi.org/10.1159/000485564>
- 38 Jørgensen CR, Freund C, Bøye R, Jordet H, Andersen D, Kjølbye M. Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. *Acta Psychiatr Scand*. 2013;127(4):305–17. <https://doi.org/10.1111/j.1600-0447.2012.01923.x>
- 39 Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563–9. <https://doi.org/10.1176/ajp.156.10.1563>
- 40 Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355–64. <https://doi.org/10.1176/appi.ajp.2009.09040539>
- 41 Robinson P, Hellier J, Barrett B, Barzdaitiene D, Bateman A, Bogaardt A, et al. The NOURISHED randomised controlled trial comparing mentalisation-based treatment for eating disorders (MBT-ED) with specialist supportive clinical management (SSCM-ED) for patients with eating disorders and symptoms of borderline personality disorder. *Trials*. 2016;17(1):549. <https://doi.org/10.1186/s13063-016-1606-8>
- 42 Juul S, Siddiqui F, Faltermeier P, Petersen JJ, Kamp CB, Jakobsen RH, et al. Less focus on symptom scales in psychiatric trials: it is time to ensure research equality between psychiatry and other medical specialties. *Lancet Reg Health Eur*. 2024;43:100993. <https://doi.org/10.1016/j.lanepe.2024.100993>
- 43 McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022;91:382–97. <https://doi.org/10.1159/000525102>



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	3
	2b	Specific objectives or hypotheses	5
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	Not applicable
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	8
	6b	Any changes to trial outcomes after the trial commenced, with reasons	Not applicable
Sample size	7a	How sample size was determined	6
	7b	When applicable, explanation of any interim analyses and stopping guidelines	Not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	6
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	6-7

		assessing outcomes) and how	
Statistical methods	11b	If relevant, description of the similarity of interventions	Not applicable
	12a	Statistical methods used to compare groups for primary and secondary outcomes	9
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	9
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	10
	13b	For each group, losses and exclusions after randomisation, together with reasons	10
Recruitment	14a	Dates defining the periods of recruitment and follow-up	10
	14b	Why the trial ended or was stopped	Not applicable
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	10 (+ supplementary)
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	10-12
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	10-12
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	10-12 (+ supplementary)
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	12 (+ supplementary)
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	12-15
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	12-15
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	12-15
Other information			
Registration	23	Registration number and name of trial registry	2
Protocol	24	Where the full trial protocol can be accessed, if available	16
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

Citation: Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine. 2010;8:18. © 2010 Schulz et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up-to-date references relevant to this checklist, see www.consort-statement.org.

Baseline sociodemographic and clinical characteristics for patients **with** ZAN-BPD outcome at 24M

Characteristics	Short-term MBT (n=63)	Long-term MBT (n=57)
Sociodemographic characteristics		
Age, mean (SD), years	30,6 (10,4)	33,0 (10,8)
Female sex – no. (%)	59 (93,7%)	52 (91,2%)
Civil status – no. (%)		
Single	35 (55,6%)	30 (52,6%)
In a relationship (living together)	13 (20,6%)	8 (14,0%)
In a relationship (living separately)	8 (12,7%)	13 (22,8%)
Married	7 (11,1%)	6 (10,5%)
Living situation – no. (%)		
With others	37 (58,7%)	36 (63,2%)
Alone	26 (41,3%)	20 (35,1%)
Other	0 (0,0%)	1 (1,8%)
Education level after high school – no. (%)		
Short education	12 (19,0%)	6 (10,5%)
Medium education	14 (22,2%)	11 (19,3%)
Long education	5 (7,9%)	4 (7,0%)
No higher education	25 (39,7%)	29 (50,9%)
Other education	7 (11,1%)	7 (12,3%)
Employment status – no. (%)		
On social welfare (/unemployed)	26 (41,3%)	34 (59,6%)
Currently studying	19 (30,2%)	12 (21,1%)
Employed	11 (17,5%)	11 (19,3%)
Unemployed/homemaker	2 (3,2%)	0 (0,0%)
Retired	1 (1,6%)	0 (0,0%)
Self-employed	1 (1,6%)	0 (0,0%)
Other	3 (4,8%)	0 (0,0%)
Clinical characteristics		
Psychiatric comorbidity (MINI), no. (%)		
Anxiety disorders	31 (49,2%)	36 (63,2%)
Major depressive disorder	23 (36,5%)	18 (31,6%)
Post-traumatic stress disorder	14 (22,2%)	18 (31,6%)
Eating disorders	10 (15,3%)	7 (12,3%)
Obsessive compulsive disorder	5 (7,9%)	7 (12,3%)
Proportion of participants with subthreshold BPD, no. (%)	8 (12,7%)	2 (3,5%)
No. of BDP criteria in total (SCID-5-PD), mean (SD)	6,2 (1,4)	6,4 (1,1)
No. of PD criteria in total (SCID-5-PD), mean (SD)	17,7 (4,9)	17,6 (4,7)
Personality disorder comorbidity (SCID-5-PD), no. (%)		
Avoidant personality disorder	19 (30,2%)	12 (21,6%)
Obsessive-compulsive personality disorder	10 (15,9%)	19 (33,3%)
Paranoid personality disorder	9 (14,3%)	13 (22,8%)
Dependent personality disorder	5 (7,9%)	4 (7,0%)
Proportion of participants with one or more suicide-attempts the past 8 months, no. (%)	5 (7,9%)	1 (1,8%)
Proportion of participants with one or more acts of severe self-harm the past 8 months, no. (%)	25 (39,7%)	22 (38,6%)
Proportion of participants on antidepressants	26 (46,4%)	24 (42,1%)
Proportion of participants on antipsychotics	8 (14,3%)	5 (8,8%)
Proportion of participants on other psychoactive drugs	15 (26,8%)	13 (22,8%)

Baseline sociodemographic and clinical characteristics for patients **without** ZAN-BPD outcome at 24M

Characteristics	Short-term MBT (n=21)	Long-term MBT (n=25)
Sociodemographic characteristics		
Age, mean (SD), years	34,9 (8,6)	30,4 (8,5)
Female sex – no. (%)	18 (85,7%)	24 (96,0%)
Civil status – no. (%)		
Single	10 (47,6%)	11 (44,0%)
In a relationship (living together)	7 (33,3%)	5 (20,0%)
In a relationship (living separately)	3 (14,3%)	6 (24,0%)
Married	1 (4,1%)	3 (12,0%)
Living situation – no. (%)		
With others	11 (52,4%)	17 (68,0%)
Alone	10 (47,6%)	8 (32,0%)
Other	0 (0%)	0 (0,0%)
Education level after high school – no. (%)		
Short education	3 (14,3%)	5 (20,0%)
Medium education	6 (28,6%)	6 (24,0%)
Long education	1 (4,8%)	1 (4,0%)
No higher education	8 (38,1%)	9 (36,0%)
Other education	3 (14,3%)	4 (16,0%)
Employment status – no. (%)		
On social welfare (/unemployed)	12 (57,1%)	13 (52,0%)
Currently studying	2 (9,5%)	7 (28,0%)
Employed	7 (33,3%)	3 (12,0%)
Unemployed/homemaker	0 (0,0%)	0 (0%)
Retired	0 (0,0%)	0 (0%)
Self-employed	0 (0,0%)	1 (4,0%)
Other	0 (0,0%)	1 (4,0%)
Clinical characteristics		
Psychiatric comorbidity (MINI), no. (%)		
Anxiety disorders	10 (47,6%)	17 (68,0%)
Major depressive disorder	8 (38,1%)	10 (40,0%)
Post-traumatic stress disorder	4 (19,0%)	6 (24,0%)
Eating disorders	2 (9,5%)	5 (20,0%)
Obsessive compulsive disorder	2 (9,5%)	3 (12,0%)
Proportion of participants with subthreshold BPD, no. (%)	3 (14,3%)	0 (0,0%)
No. of BDP criteria in total (SCID-5-PD), mean (SD)	6,2 (1,3)	6,8 (1,3)
No. of PD criteria in total (SCID-5-PD), mean (SD)	18,8 (4,9)	18,9 (6,7)
Personality disorder comorbidity (SCID-5-PD), no. (%)		
Avoidant personality disorder	5 (23,8%)	9 (36,0%)
Obsessive-compulsive personality disorder	4 (19,0%)	8 (32,0%)
Paranoid personality disorder	5 (23,8%)	7 (28,0%)
Dependent personality disorder	3 (14,3%)	3 (12,0%)
Proportion of participants with one or more suicide-attempts the past 8 months, no. (%)	0 (0,0%)	5 (20,0%)
Proportion of participants with one or more acts of severe self-harm the past 8 months, no. (%)	7 (33,3%)	11 (44,0%)
Proportion of participants on antidepressants	9 (42,9%)	11 (44,0%)
Proportion of participants on antipsychotics	3 (14,3%)	6 (24,0%)
Proportion of participants on other psychoactive drugs	6 (28,6%)	7 (28,0%)

Statistical report for the MBT-RCT 24 month follow-up

Markus Harboe Olsen

1/23/24

Table of contents

Primary outcome	2
Severity of borderline symptomatology (ZAN-BPD, 24-months)	2
Complete case analysis (primary analysis)	2
Multiple imputation (exploratory analysis)	5
Best-worst / worst-best (exploratory analysis)	7
Stratification variables (exploratory analysis)	7
Secondary outcomes	12
Functional impairment (WSAS, 24-months)	12
Quality of life (SF-36 mental component, 24-months)	15
Global functioning (GAF, 24-months)	18
Severe self-harm (24-months)	22
Serious adverse events (24-months)	23
Subgroup analysis	24
Baseline ZAN-BPD scores (0-11/12-36)	24
Sex (male/female)	25
Age (18-30/31-60)	26
Baseline GAF scores (0-48/49-100)	27
Baseline selfharm	28
Paused treatment due to COVID-19	29
Forest plot	30

Primary outcome

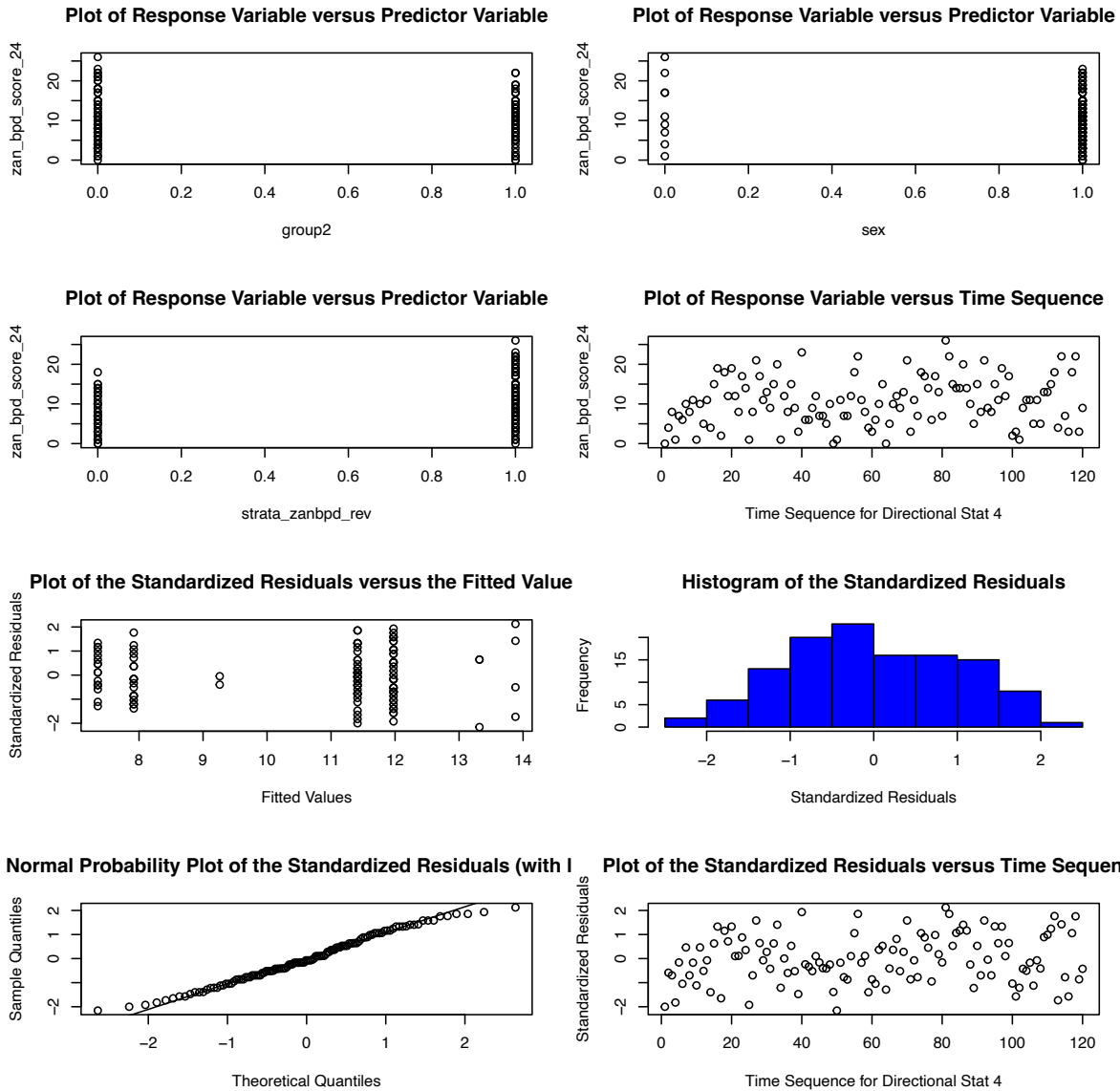
Severity of borderline symptomatology (ZAN-BPD, 24-months)

Complete case analysis (primary analysis)

	1 <i>n</i> = 84	2 <i>n</i> = 82
zan_bpd_score_base		
mean (95%CI) [SD]	13.90 (13.10;14.71) [3.72]	13.80 (12.98;14.63) [3.75]
zan_bpd_score_8		
mean (95%CI) [SD]	11.96 (10.72;13.19) [5.33]	12.77 (11.61;13.93) [5.01]
missing	10 (11.9%)	8 (9.8%)
zan_bpd_score_16		
mean (95%CI) [SD]	11.34 (9.86;12.83) [6.22]	12.28 (10.78;13.78) [5.86]
missing	14 (16.7%)	21 (25.6%)
zan_bpd_score_24		
mean (95%CI) [SD]	10.87 (9.22;12.53) [6.58]	10.23 (8.77;11.69) [5.49]
missing	21 (25%)	25 (30.5%)

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	9.82	5.43 – 14.21	<0.001
group [2]	-0.56	-2.67 – 1.54	0.598
sex	-1.90	-5.90 – 2.09	0.347
strata zanbpd rev	4.06	1.80 – 6.32	0.001
Observations	120		
R ² / R ² adjusted	0.110 / 0.087		

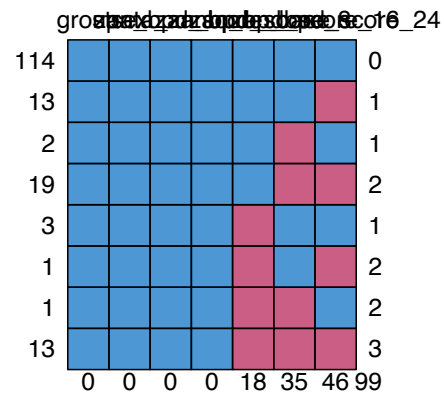
Assumptions for linear regression



	Value	p-value	Decision
Global Stat	4.2183369	0.3772628	Assumptions acceptable.
Skewness	0.0646554	0.7992836	Assumptions acceptable.
Kurtosis	2.8992158	0.0886227	Assumptions acceptable.
Link Function	0.8823338	0.3475633	Assumptions acceptable.

	Value	p-value	Decision
Heteroscedasticity	0.3721319	0.5418445	Assumptions acceptable.

Multiple imputation (exploratory analysis)



Complete data summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.00	6.00	10.00	10.57	15.00	26.00	46

Imputed data summary

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	7.00	11.00	10.78	15.00	26.00

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.15	10.91 – 17.39	< 0.001
group [2]	-0.63	-2.34 – 1.09	0.471
sex	-1.96	-5.16 – 1.24	0.228
strata zanbpd	-4.23	-6.12 – -2.35	< 0.001
Observations	166		
R ² / R ² adjusted	0.116		
	/		
	0.100		

Best-worst / worst-best (exploratory analysis)**Best 1 / Worst 2**

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	8.32	4.03 – 12.60	< 0.001
group [2]	6.06	3.79 – 8.32	< 0.001
sex	0.20	-4.02 – 4.43	0.925
strata zanbpd	-3.23	-5.71 – -0.74	0.011
Observations	166		
R ² / R ² adjusted	0.173		
	/		
	0.158		

Best 2 / Worst 1

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	18.77	14.52 – 23.02	< 0.001
group [2]	-7.19	-9.43 – -4.94	< 0.001
sex	-4.05	-8.24 – 0.14	0.058
strata zanbpd	-3.10	-5.57 – -0.63	0.014
Observations	166		
R ² / R ² adjusted	0.237		
	/		
	0.223		

Stratification variables (exploratory analysis)**Age (18-30/31-60)**

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.07	9.89 – 18.25	< 0.001
group [2]	-0.51	-2.64 – 1.61	0.633
sex	-1.96	-5.99 – 2.06	0.335
strata zanbpd	-4.05	-6.32 – -1.78	0.001
covar ageTRUE	-0.40	-2.56 – 1.77	0.718

zan_bpd_score_24	
Observations	120
R ² / R ² adjusted	0.111
	/
	0.080

GAF (0-48/49-100)

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.56	10.36 – 18.77	<0.001
group [2]	-0.68	-2.79 – 1.44	0.527
sex	-2.15	-6.17 – 1.86	0.290
strata zanbpd	-3.68	-6.04 – -1.33	0.002
covar gafTRUE	-1.26	-3.50 – 0.98	0.266
Observations	120		
R ² / R ² adjusted	0.120		
	/		
	0.089		

Selfharm ($\geq 1/0$)

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	13.47	9.31 – 17.62	<0.001
group [2]	-0.55	-2.66 – 1.56	0.605
sex	-1.86	-5.86 – 2.15	0.360
strata zanbpd	-4.03	-6.29 – -1.76	0.001
self harm baseline [1]	0.90	-1.25 – 3.06	0.410
Observations	120		
R ² / R ² adjusted	0.115		
	/		
	0.085		

Paused treatment

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.63	10.57 – 18.70	< 0.001
group [2]	-1.47	-3.76 – 0.82	0.206
sex	-1.05	-5.10 – 3.00	0.609
strata zanbpd	-4.05	-6.29 – -1.82	< 0.001
covid19 [Nej]	-2.22	-4.57 – 0.12	0.063
Observations	120		
R ² / R ² adjusted	0.137		
	/		
	0.107		

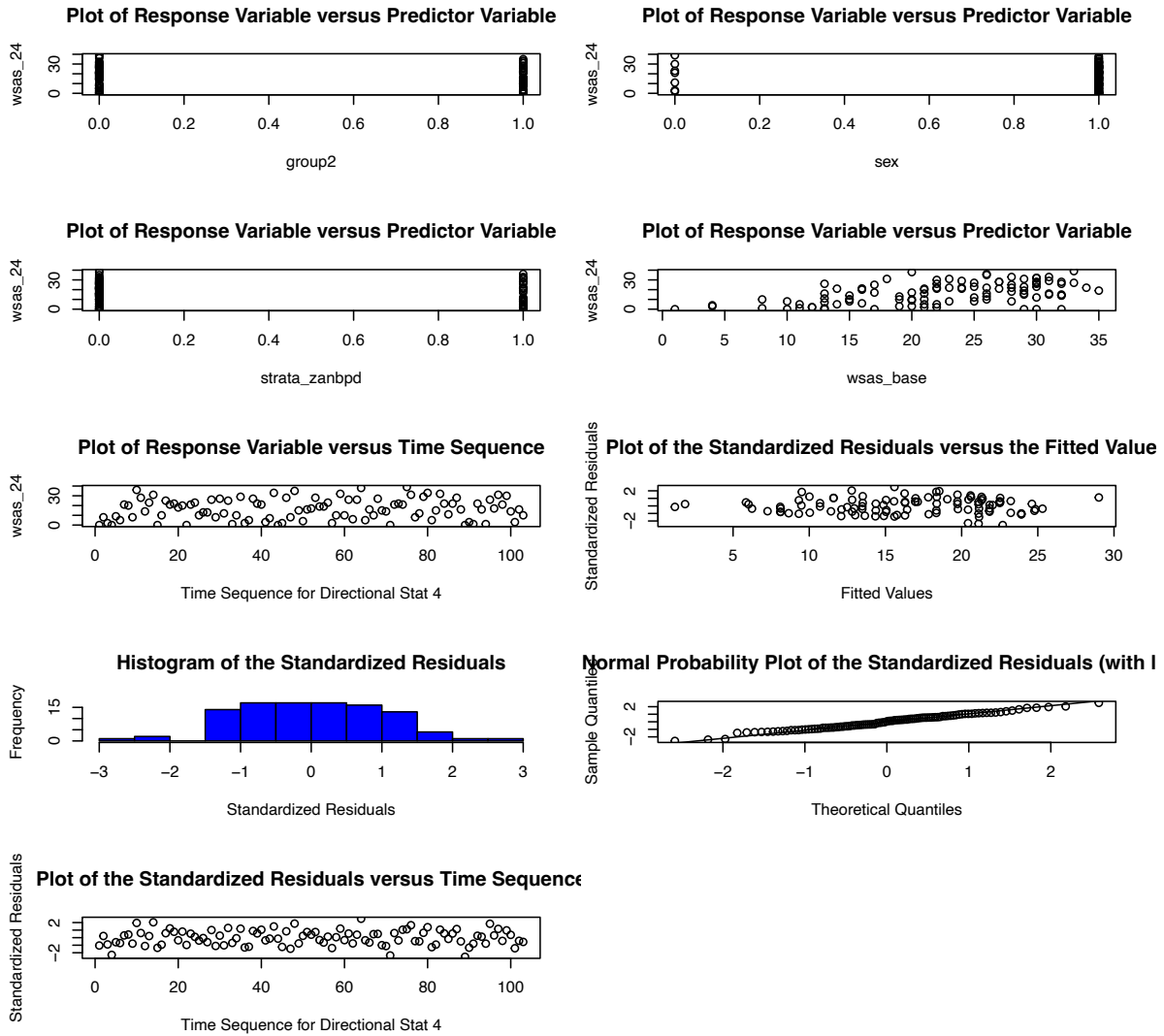
Secondary outcomes

Functional impairment (WSAS, 24-months)

	1 <i>n</i> = 84	2 <i>n</i> = 82
wsas_base		
mean (95%CI) [SD]	22.94 (21.30;24.58) [7.54]	21.88 (20.24;23.52) [7.46]
wsas_8		
mean (95%CI) [SD]	20.44 (18.40;22.47) [8.59]	19.94 (17.93;21.95) [8.36]
missing	13 (15.5%)	13 (15.9%)
wsas_16		
mean (95%CI) [SD]	16.71 (14.47;18.96) [8.91]	17.20 (14.69;19.70) [9.36]
missing	21 (25%)	26 (31.7%)
wsas_24		
mean (95%CI) [SD]	17.50 (14.61;20.39) [10.80]	15.06 (12.05;18.08) [10.28]
missing	28 (33.3%)	35 (42.7%)

wsas_24			
Predictors	Estimates	CI	p
(Intercept)	6.10	-2.82 – 15.02	0.178
group [2]	-1.42	-5.04 – 2.21	0.440
sex	-4.39	-11.57 – 2.78	0.227
strata zanbpd	-1.22	-5.17 – 2.73	0.541
wsas base	0.69	0.45 – 0.94	<0.001
Observations	103		
R ² / R ² adjusted	0.280 / 0.250		

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	0.8250511	0.9350584	Assumptions acceptable.
Skewness	0.0007554	0.9780740	Assumptions acceptable.
Kurtosis	0.4132212	0.5203390	Assumptions acceptable.
Link Function	0.3933525	0.5305424	Assumptions acceptable.
Heteroscedasticity	0.0177221	0.8940951	Assumptions acceptable.

Quality of life (SF-36 mental component, 24-months)

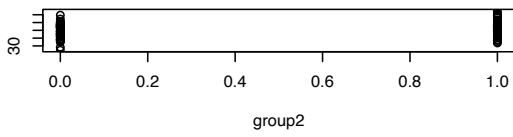
	1 <i>n</i> = 84	2 <i>n</i> = 82
sf_36_mental_component_baseline		
mean (95%CI) [SD]	50.26 (48.01;52.51) [10.39]	49.21 (47.00;51.43) [10.09]
sf_36_mental_component_8		
mean (95%CI) [SD]	50.36 (47.81;52.91) [10.76]	53.07 (50.29;55.85) [11.58]
missing	13 (15.5%)	13 (15.9%)
sf_36_mental_component_16		
mean (95%CI) [SD]	53.87 (51.37;56.38) [9.94]	52.91 (50.18;55.63) [10.17]
missing	21 (25%)	26 (31.7%)
sf_36_mental_component_24		
mean (95%CI) [SD]	49.05 (46.62;51.48) [9.07]	54.44 (51.41;57.46) [10.30]
missing	28 (33.3%)	35 (42.7%)

sf_36_mental_component_24			
Predictors	Estimates	CI	p
(Intercept)	41.12	29.02 – 53.22	<0.001
group [2]	5.52	1.72 – 9.32	0.005
sex	1.68	-5.81 – 9.17	0.658
strata zanbpd	2.24	-1.97 – 6.46	0.293
sf 36 mental component baseline	0.11	-0.08 – 0.30	0.249
Observations	103		
R ² / R ² adjusted	0.106 / 0.070		

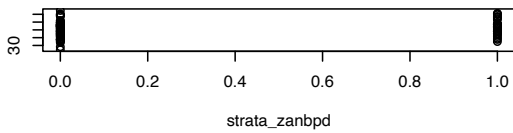
Assumptions for linear regression

sf_36_mental_component_2

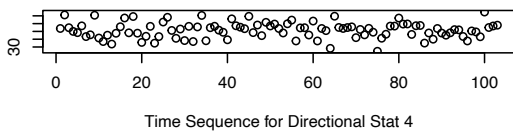
Plot of Response Variable versus Predictor Variable



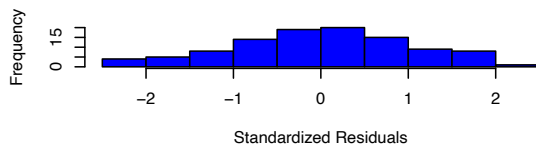
Plot of Response Variable versus Predictor Variable



Plot of Response Variable versus Time Sequence

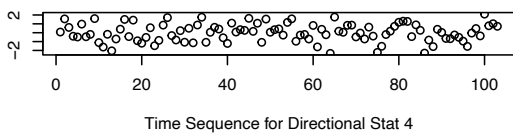


Histogram of the Standardized Residuals



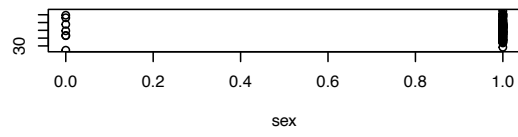
Standardized Residuals

Plot of the Standardized Residuals versus Time Sequence

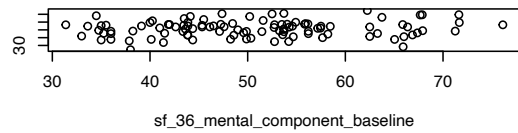


sf_36_mental_component_2

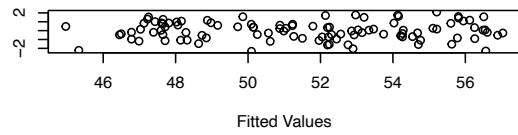
Plot of Response Variable versus Predictor Variable



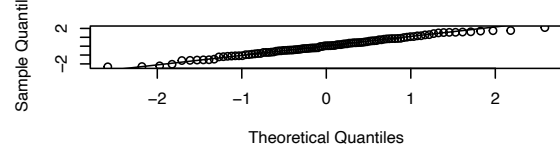
Plot of Response Variable versus Predictor Variable



Plot of the Standardized Residuals versus the Fitted Value



Normal Probability Plot of the Standardized Residuals (with I



	Value	p-value	Decision
Global Stat	1.5293684	0.8214268	Assumptions acceptable.
Skewness	0.3624164	0.5471673	Assumptions acceptable.
Kurtosis	1.1106677	0.2919369	Assumptions acceptable.
Link Function	0.0455091	0.8310707	Assumptions acceptable.
Heteroscedasticity	0.0107753	0.9173249	Assumptions acceptable.

Global functioning (GAF, 24-months)

	1 <i>n</i> = 84	2 <i>n</i> = 82
gaf_score_base		
median (IQR)	45.00 (42.00;51.00)	45.00 (41.00;50.00)
gaf_score_8		
median (IQR)	47.50 (40.00;58.00)	48.00 (40.00;55.50)
missing	10 (11.9%)	8 (9.8%)
gaf_score_16		
median (IQR)	49.00 (41.00;61.50)	48.00 (41.00;58.00)
missing	14 (16.7%)	21 (25.6%)
gaf_score_24		
median (IQR)	48.00 (41.00;54.00)	49.50 (40.75;58.25)
missing	21 (25%)	26 (31.7%)

	1 <i>n</i> = 84	2 <i>n</i> = 82
gaf_score_base		
mean (95%CI) [SD]	46.55 (44.78;48.31) [8.14]	46.28 (44.88;47.68) [6.36]
gaf_score_8		
mean (95%CI) [SD]	49.55 (46.42;52.69) [13.55]	47.74 (45.31;50.17) [10.48]
missing	10 (11.9%)	8 (9.8%)
gaf_score_16		
mean (95%CI) [SD]	52.23 (48.65;55.81) [15.03]	49.74 (46.73;52.74) [11.74]
missing	14 (16.7%)	21 (25.6%)
gaf_score_24		
mean (95%CI) [SD]	48.40 (45.58;51.22) [11.20]	50.70 (47.10;54.29) [13.43]
missing	21 (25%)	26 (31.7%)

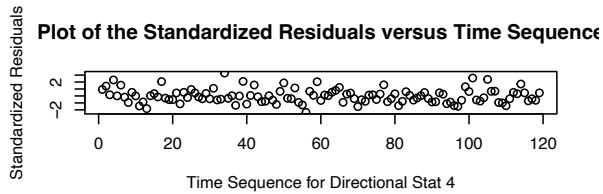
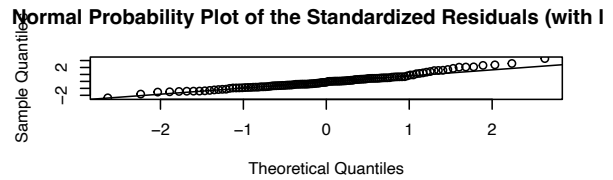
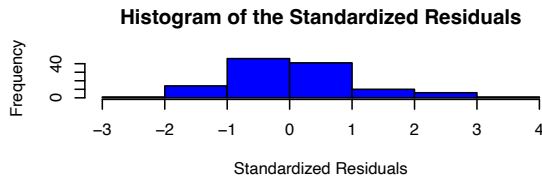
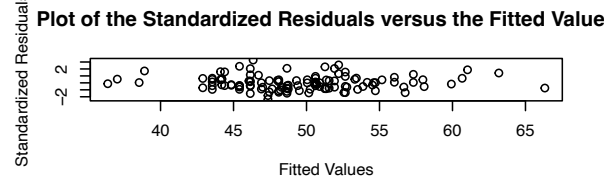
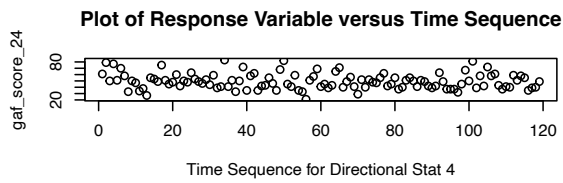
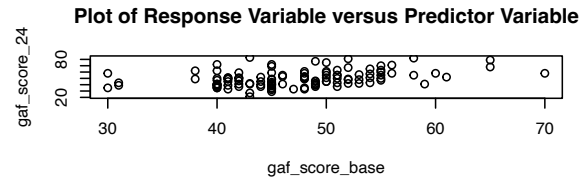
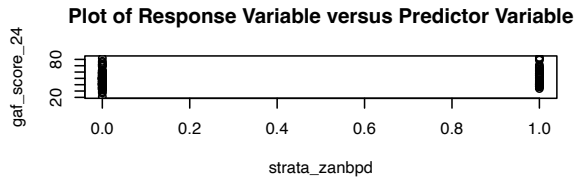
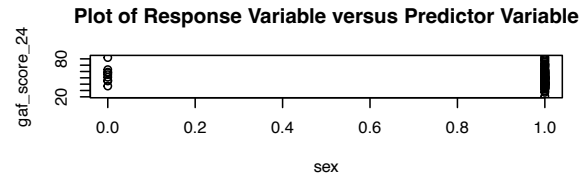
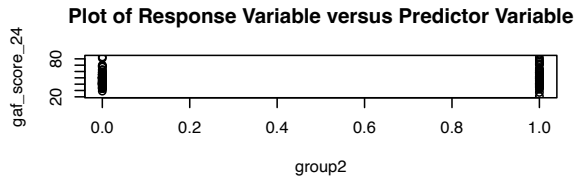
gaf_score_24			
Predictors	Estimates	CI	p
(Intercept)	20.77	3.11 – 38.43	0.022
group [2]	2.51	-1.65 – 6.67	0.234
sex	-3.91	-11.85 – 4.03	0.332

gaf_score_24			
strata zanbpd	1.50	-3.20 – 6.20	0.529
gaf score base	0.65	0.33 – 0.97	<0.001
Observations	119		
R ² / R ² adjusted	0.169		
	/		
	0.139		

Wilcoxon rank sum test with continuity correction

```
data:  df$gaf_score_24[df$group == 2] and df$gaf_score_24[df$group == 1]
W = 1954.5, p-value = 0.3114
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -2.000025  6.999956
sample estimates:
difference in location
      2.000038
```

Assumptions for linear regression



	Value	p-value	Decision
Global Stat	14.0562263	0.0071177	Assumptions NOT satisfied!
Skewness	9.3563290	0.0022222	Assumptions NOT satisfied!
Kurtosis	1.5650614	0.2109257	Assumptions acceptable.
Link Function	2.7777565	0.0955820	Assumptions acceptable.
Heteroscedasticity	0.3570795	0.5501327	Assumptions acceptable.

Severe self-harm (24-months)

	1 <i>n</i> = 84	2 <i>n</i> = 82	Logistic regression <i>RR</i> (95% <i>CI</i>)	<i>p</i>
self_harm_baseline				
0	52 (61.9%)	49 (59.8%)		
1	32 (38.1%)	33 (40.2%)		
self_harm_8				
0	55 (65.5%)	47 (57.3%)		
1	19 (22.6%)	27 (32.9%)		
missing	10 (11.9%)	8 (9.8%)		
self_harm_16				
0	55 (65.5%)	45 (54.9%)		
1	13 (15.5%)	16 (19.5%)		
missing	16 (19.0%)	21 (25.6%)		
self_harm_24			1.24 (0.56;2.77)	0.596
0	53 (63.1%)	46 (56.1%)		
1	10 (11.9%)	11 (13.4%)		
missing	21 (25.0%)	25 (30.5%)		

Sensitivity analysis using G-computation

Relative risk derived by G-computation

RR: 1.24 (sRR: 1.22)

Confidence interval: 0.52 - 2.88

p-value: 0.5878

It took 16 seconds.

Serious adverse events (24-months)

	1 <i>n</i> = 84	2 <i>n</i> = 82	Logistic regression <i>RR</i> (95% <i>CI</i>)	<i>p</i>
sae_yn_24m			1.38 (0.88;2.18)	0.161
0	61 (72.6%)	51 (62.2%)		
1	23 (27.4%)	31 (37.8%)		

Sensitivity analysis using G-computation

Relative risk derived by G-computation

RR: 1.38 (sRR: 1.37)

Confidence interval: 0.88 - 2.21

p-value: 0.1498

It took 15 seconds.

Subgroup analysis

Baseline ZAN-BPD scores (0-11/12-36)

P-value for interaction: 0.3627107

Baseline ZAN-BPD scores 0-11 ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	6.92	-0.91 – 14.74	0.081
group [2]	1.08	-2.26 – 4.42	0.515
sex	0.29	-7.19 – 7.78	0.937
Observations	38		
R ² / R ² adjusted	0.012		
	/ -		
	0.044		

Baseline ZAN-BPD scores 12-36 ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.52	9.73 – 19.32	< 0.001
group [2]	-1.22	-3.94 – 1.51	0.377
sex	-2.27	-7.14 – 2.59	0.355
Observations	82		
R ² / R ² adjusted	0.021		
	/ -		
	0.004		

:::

...

Sex (male/female)

P-value for interaction: 0.3682704

Sex 0 ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	15.75	4.84 – 26.66	0.012
group [2]	-4.08	-20.75 – 12.59	0.571
strata zanbpd	-3.67	-23.59 – 16.26	0.668
Observations	9		
R ² / R ² adjusted	0.151		
	/ -		
	0.133		

Sex 1 ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	11.82	10.20 – 13.44	<0.001
group [2]	-0.29	-2.41 – 1.82	0.785
strata zanbpd	-3.96	-6.21 – -1.70	0.001
Observations	111		
R ² / R ² adjusted	0.101		
	/		
	0.085		

...

...

Age (18-30/31-60)

P-value for interaction: 0.9674119

Age 18-30 ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	13.91	8.03 – 19.79	< 0.001
group [2]	-0.58	-3.37 – 2.22	0.683
sex	-1.43	-7.41 – 4.55	0.634
strata zanbpd	-5.06	-8.04 – -2.08	0.001
Observations	71		
R ² / R ² adjusted	0.151		
	/		
	0.113		

Age 31-60 ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.04	7.96 – 20.12	< 0.001
group [2]	-0.91	-4.51 – 2.68	0.611
sex	-2.67	-8.49 – 3.14	0.359
strata zanbpd	-2.53	-6.30 – 1.25	0.184
Observations	49		
R ² / R ² adjusted	0.071		
	/		
	0.009		

...

...

Baseline GAF scores (0-48/49-100)

P-value for interaction: 0.7313637

Baseline GAF scores 0-48 :: { .cell-output-display }

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	17.50	11.55 – 23.46	< 0.001
group [2]	-1.14	-3.80 – 1.52	0.396
sex	-5.04	-10.81 – 0.72	0.086
strata zanbpd	-3.59	-6.82 – -0.35	0.030
Observations	71		
R ² / R ² adjusted	0.110		
	/		
	0.070		

Baseline GAF scores 49-100 :: { .cell-output-display }

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	11.11	5.32 – 16.90	< 0.001
group [2]	-0.22	-3.83 – 3.39	0.903
sex	0.29	-5.69 – 6.27	0.923
strata zanbpd	-4.11	-7.75 – -0.46	0.028
Observations	49		
R ² / R ² adjusted	0.107		
	/		
	0.047		

...

...

Baseline selfharm

P-value for interaction: 0.592711

Baseline no selfharm ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	10.44	4.88 – 16.00	<0.001
group [2]	-0.93	-3.66 – 1.80	0.500
sex	1.30	-4.10 – 6.70	0.633
strata zanbpd	-3.21	-6.11 – -0.31	0.030
Observations	73		
R ² / R ² adjusted	0.078		
	/		
	0.038		

Baseline selfharm ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	17.67	11.65 – 23.70	<0.001
group [2]	0.15	-3.24 – 3.54	0.929
sex	-5.57	-11.74 – 0.60	0.076
strata zanbpd	-4.82	-8.59 – -1.05	0.013
Observations	47		
R ² / R ² adjusted	0.223		
	/		
	0.169		

...

...

Paused treatment due to COVID-19

P-value for interaction: 0.6145096

Baseline no pause ::: {.cell-output-display}

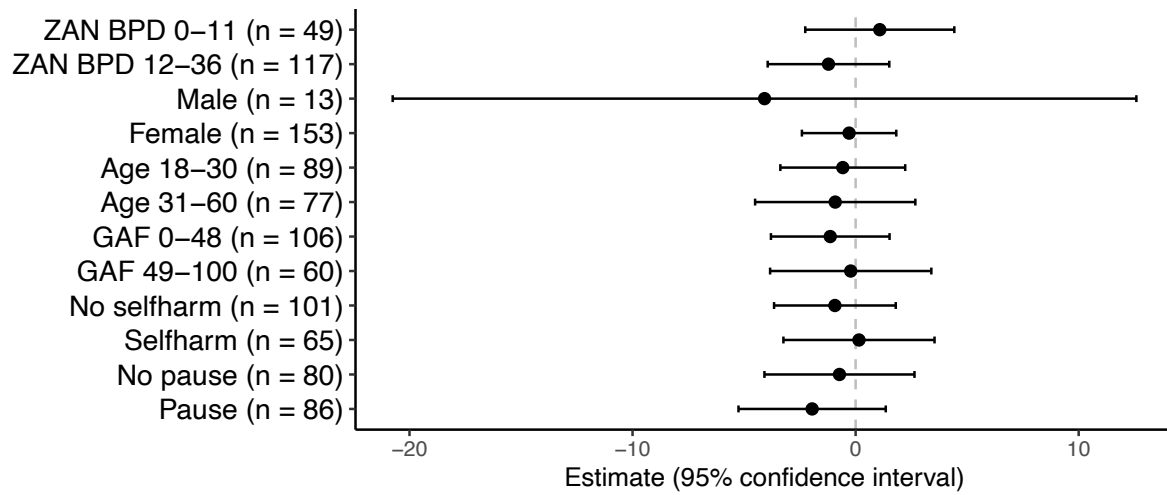
zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	11.00	-0.36 – 22.36	0.057
group [2]	-0.72	-4.09 – 2.64	0.668
sex	0.51	-11.01 – 12.04	0.929
strata zanbpd	-5.09	-8.29 – -1.89	0.002
Observations	60		
R ² / R ² adjusted	0.165		
	/		
	0.120		

Baseline pause ::: {.cell-output-display}

zan_bpd_score_24			
Predictors	Estimates	CI	p
(Intercept)	14.89	10.12 – 19.66	<0.001
group [2]	-1.95	-5.25 – 1.35	0.242
sex	-1.29	-5.80 – 3.22	0.570
strata zanbpd	-3.19	-6.50 – 0.11	0.058
Observations	60		
R ² / R ² adjusted	0.088		
	/		
	0.039		

... ..

Forest plot



Baseline sociodemographic and clinical characteristics – patients **with** SF-36 outcome

data at 24M

Characteristics	Short-term MBT (n=56)	Long-term MBT (n=47)
Sociodemographic characteristics		
Age, mean (SD), years	30,7 (10,9)	32,8 (10,7)
Female sex – no. (%)	53 (94,6%)	43 (91,5%)
Civil status – no. (%)		
Single	32 (57,1%)	26 (55,3%)
In a relationship (living together)	10 (17,9%)	5 (10,6%)
In a relationship (living separately)	7 (12,5%)	11 (23,4%)
Married	7 (12,5%)	5 (10,6%)
Living situation – no. (%)		
With others	31 (55,4%)	29 (61,7%)
Alone	25 (44,6%)	18 (38,3%)
Other	0 (0%)	0 (0,0%)
Education level after high school – no. (%)		
Short education	11 (19,6%)	7 (14,9%)
Medium education	14 (25,0%)	7 (14,9%)
Long education	3 (5,4%)	3 (6,4%)
No higher education	23 (41,1%)	25 (53,2%)
Other education	5 (8,9%)	5 (10,6%)
Employment status – no. (%)		
On social welfare (/unemployed)	23 (41,1%)	28 (59,6%)
Currently studying	17 (30,4%)	11 (23,4%)
Employed	10 (17,9%)	8 (17,0%)
Unemployed/homemaker	2 (3,6%)	0 (0,0%)
Retired	1 (1,8%)	0 (0,0%)
Self-employed	0 (0,0%)	0 (0,0%)
Other	3 (5,4%)	0 (0,0%)
Clinical characteristics		
Psychiatric comorbidity (MINI), no. (%)		
Anxiety disorders	26 (46,4%)	27 (57,4%)
Major depressive disorder	23 (41,1%)	16 (34,0%)
Post-traumatic stress disorder	13 (23,2%)	16 (34,0%)
Eating disorders	7 (12,5%)	6 (12,8%)
Obsessive compulsive disorder	4 (7,1%)	6 (12,8%)
Proportion of participants with subthreshold BPD, no. (%)	6 (10,7%)	2 (4,3%)
No. of BDP criteria in total (SCID-5-PD), mean (SD)	6,2 (1,3)	6,3 (1,1)
No. of PD criteria in total (SCID-5-PD), mean (SD)	18,0 (5,0)	17,7 (5,1)
Personality disorder comorbidity (SCID-5-PD), no. (%)		
Avoidant personality disorder	18 (32,1%)	10 (21,3%)
Obsessive-compulsive personality disorder	9 (16,1%)	13 (27,7%)
Paranoid personality disorder	8 (14,3%)	12 (25,5%)
Dependent personality disorder	5 (8,9%)	4 (8,5%)
Proportion of participants with one or more suicide-attempts the past 8 months, no. (%)	4 (7,1%)	1 (2,1%)
Proportion of participants with one or more acts of severe self-harm the past 8 months, no. (%)	22 (39,3%)	18 (38,3%)
Proportion of participants on antidepressants	23 (41,1%)	19 (40,4%)
Proportion of participants on antipsychotics	6 (10,7%)	5 (10,6%)

Proportion of participants on other psychoactive drugs	15 (26,8%)	13 (27,7%)
Mean (SD) number of days from randomization to assessment time-point		
8 months post-randomization		
16 months post-randomization		

Baseline sociodemographic and clinical characteristics – patients **without** SF-36 data at

24M

Characteristics	Short-term MBT (n=28)	Long-term MBT (n=35)
Sociodemographic characteristics		
Age, mean (SD), years	33,6 (8,2)	31,4 (9,4)
Female sex – no. (%)	24 (85,7%)	33 (94,3%)
Civil status – no. (%)		
Single	13 (46,4%)	15 (42,9%)
In a relationship (living together)	10 (35,7%)	8 (22,9%)
In a relationship (living separately)	4 (14,3%)	8 (22,9%)
Married	1 (3,6%)	4 (11,4%)
Living situation – no. (%)		
With others	17 (60,7%)	24 (68,6%)
Alone	11 (39,3%)	10 (28,6%)
Other	0 (0,0%)	1 (2,9%)
Education level after high school – no. (%)		
Short education	4 (14,3%)	4 (11,4%)
Medium education	6 (21,4%)	10 (28,6%)
Long education	3 (10,7%)	2 (5,7%)
No higher education	10 (35,7%)	13 (37,1%)
Other education	5 (17,9%)	6 (17,1%)
Employment status – no. (%)		
On social welfare (/unemployed)	15 (53,6%)	19 (54,3%)
Currently studying	4 (14,3%)	8 (22,9%)
Employed	8 (28,6%)	6 (17,1%)
Unemployed/homemaker	0 (0,0%)	0 (0,0%)
Retired	0 (0,0%)	0 (0,0%)
Self-employed	1 (3,6%)	1 (2,9%)
Other	0 (0,0%)	1 (2,9%)
Clinical characteristics		
Psychiatric comorbidity (MINI), no. (%)		
Anxiety disorders	15 (53,6%)	26 (74,3%)
Major depressive disorder	8 (28,6%)	12 (34,3%)
Post-traumatic stress disorder	5 (17,9%)	8 (22,9%)
Eating disorders	5 (17,9%)	6 (17,1%)
Obsessive compulsive disorder	3 (10,7%)	4 (11,4%)
Proportion of participants with subthreshold BPD, no. (%)	5 (17,9%)	0 (0,0%)
No. of BDP criteria in total (SCID-5-PD), mean (SD)	6,1 (1,4)	6,8 (1,2)
No. of PD criteria in total (SCID-5-PD), mean (SD)	17,9 (4,8)	18,4 (5,7)
Personality disorder comorbidity (SCID-5-PD), no. (%)		
Avoidant personality disorder	6 (21,4%)	11 (31,4%)
Obsessive-compulsive personality disorder	5 (17,9%)	14 (40,0%)
Paranoid personality disorder	6 (21,4%)	8 (22,9%)
Dependent personality disorder	3 (10,7%)	3 (8,6%)
Proportion of participants with one or more suicide-attempts the past 8 months, no. (%)	1 (3,6%)	5 (14,3%)
Proportion of participants with one or more acts of severe self-harm the past 8 months, no. (%)	10 (35,7%)	15 (42,9%)
Proportion of participants on antidepressants	12 (42,9%)	16 (45,7%)
Proportion of participants on antipsychotics	5 (17,9%)	6 (17,1%)
Proportion of participants on other psychoactive drugs	6 (21,4%)	7 (20,0%)

Mean (SD) number of days from randomization to assessment time-point
8 months post-randomization
16 months post-randomization

Individual serious adverse events at 24 months after randomization

Short-term MBT		Long-term MBT	
Participant	Events	Participant	Events
A1	1. Hospitalisation due to headache and pain in the neck and jaws (obs encephalitis) 2. Hospitalisation due to headache (2 nights)	B1	1. Admission to gynaecology ward (1 night)
A2	1. Hospitalisation* 2. Hospitalisation due to medicine overdose (10 paracetamol)	B2	1. Hospitalisation due to appendicitis
A3	1. Hospitalisation (1 night) due to overdose of sertraline (2000 mg instead of 200 mg), dizziness and diarrhoea	B3	1. Hospitalisation (2 nights) due to medicine overdose (6 oxazepam) 2. Admission to a gastrointestinal surgery department
A4	1. Hospitalisation (21 nights) due to suicidal ideation 2. Hospitalisation (5 nights) due to suicidal ideation 3. Hospitalisation (1 night) due to suicidal ideation, self-harm, and drug abuse	B4	1. Hospitalisation (1 night) due to abdominal pain 2. Hospitalisation (2 nights) due to suicide attempt (overdose of 70x tramadol 50 mg) 3. Admission to a gastrointestinal surgery department 4. Hospitalisation*
A5	1. Hospitalisation (20 nights) due to suicidal ideation and self-harm	B5	1. Hospitalisation (1 night) due to suicide attempt (10-20 ibuprofen 400 mg)
A6	1. Hospitalisation (1 night) due to tick bite and hemiplegia	B6	1. Hospitalisation (1 month) due to suicidal ideation
A7	1. Hospitalisation (1 night) due to suicidal ideation. The patient is drunk and has taken cocaine 2. Hospitalisation (1 night) due to suicide attempt. The patient has taken 2 full packets of quetiapine 25 mg 3. Hospitalisation due to suicidal ideation	B7	1. Suicidal ideation

	4. Hospitalisation due to suicide attempt (60x topiramate 100 mg) 5. Suicide attempt (11 methylphenidate, 5 oxazepam, 1 tramadol).		
A8	1. Hospitalisation (1 night) due to drugs or alcohol	B8	1. Admission to psychiatric ward 2. Admission to psychiatric ward 3. Hospitalisation (1 night) due to suicidal ideation 4. Hospitalisation (1 night) due to suicidal ideation 5. Hospitalisation (1 night) to psychiatric ward
A9	1. Hospitalisation (4 nights) due to suicidal ideation 2. Hospitalisation*	B9	1. Admission to neurological department (1 night)
A10	1. Hospitalisation (1 night) due to overdose of 60x sertraline 50 mg 2. Hospitalisation (10 nights) due to patient swallowing a battery 3. Hospitalisation (2 nights) due to suicidal ideation	B10	1. Hospitalisation (3 nights) due to anxiety and suicidal ideation
A11	1. Hospitalisation (2 nights) due to suicidal ideation	B11	1. Hospitalisation (2 nights) due to suicidal ideation 2. Hospitalisation (1 night) due to suicidal ideation 3. Hospitalisation (3 nights) due to suicidal ideation and severe depression
A12	1. Hospitalisation (1 night) due to seizures	B12	1. Hospitalisation (1 night) due to overdose of 14x sertraline 100 mg, 10x paracetamol, MDMA and cocaine. Drug-induced seizures. 2. Hospitalisation (1 night) due to suicide attempt and alcohol poisoning

			3. Hospitalisation (1 night) due to suicide attempt 4. Hospitalisation (1 night) due to suicidal ideation and possibly psychosis. Under the influence of alcohol and drugs. 5. Hospitalisation (1 night) due to possible overdose of medicine. Unconscious at admission. 6. Hospitalisation (1 night) due to possible overdose of medicine
A13	1. Hospitalisation (1 night) due to suicide attempt (overdose) 2. Hospitalisation (2 nights) due to suicidal ideation	B13	1. Admission to psychiatric ward 2. Admission to psychiatric ward 3. Hospitalisation due to possible psychosis 4. Hospitalisation (1 night) due to possible psychosis 5. Hospitalisation (10 nights) due to possible psychosis
A14	1. Hospitalisation (1 night) due to suicide attempt (overdose) 2. Hospitalisation (1 night) due to suicidal ideation	B14	1. Hospitalisation (1 night) due to abdominal pain 2. Hospitalisation (1 night) due to abdominal pain 3. Hospitalisation (1 night) due to abdominal pain and black stools 4. Hospitalisation (3 nights) due to suicidal ideation and anxiety 5. Hospitalisation (3 nights) due to suicidal ideation 6. Hospitalisation (3 nights) due to suicide attempt (75x promethazine 25 mg and 5x oxazepam 15 mg) 7. Hospitalisation (1 night) due to suicidal ideation

			8. Hospitalisation (1 night) due to self-harm 9. Hospitalisation (1 night) due to suicidal ideation
A15	1. Hospitalisation*	B15	1. Hospitalisation (4 nights) due to depressive symptoms
A16	1. Hospitalisation due to possible overdose 2. Suicide following carbon dioxide poisoning	B16	1. Hospitalisation (1 night) due to severe headache 2. Hospitalisation (2 nights) due to asthma
A17	1. Hospitalisation (1 night) due to anxiety	B17	1. Hospitalisation (3 nights) due to suicide attempt (overdose) 2. Hospitalisation (1 night) due to shortness of breath 3. Hospitalisation (1 night) due to suicidal ideation 4. Hospitalisation (8 nights) due to suicidal ideation
A18	1. Hospitalisation (five weeks) due to severe depression and suicidal ideation	B18	1. Hospitalisation (1 night) due to suicidal ideation
A19	1. Hospitalisation (1 night) to psychiatric ward	B19	1. Hospitalisation (5 nights) due to suicidal ideation and depressive symptoms
A20	1. Hospitalisation (1 night) due to depression and suicidal ideation	B20	1. Hospitalisation (6 nights) due to possible symptoms of bipolar disorder 2. Hospitalisation (1 night) due to absces
A21	1. Hospitalisation (1 night) due to bulimia	B21	1. Electric shock (230v) 2. Electric shock
A22	1. Hospitalisation in gynaecology ward	B22	1. Hospitalisation (1 night) due to chest pain
A23	1. Hospitalisation* (2 nights)	B23	1. Admission to psychiatric ward 2. Admission to cardiological ward 3. Admission to cardiological ward 4. Admission to cardiological ward 5. Hospitalisation* 6. Admission to cardiological ward

		B24	1. Hospitalisation due to suicidal ideation 2. Hospitalisation due to suicidal ideation 3. Hospitalisation due to suicide attempt (overdose of quetiapine)
		B25	1. Hospitalisation (5 nights) due to back pain 2. Hospitalisation (3 nights) due to back pain
		B26	1. Hospitalisation (12 nights) to psychiatric ward
		B27	1. Hospitalisation following abdominal surgery
		B28	1. Hospitalisation* (1 night)
		B29	1. Hospitalisation*
		B30	1. Hospitalisation*
		B31	1. Hospitalised due to traffic accident

*Unclear

Paper 8

Response to Philips et al.: Several Fundamental Misconceptions about Evidence-Based Practice

Sophie Juul^{a, b, c} Janus Christian Jakobsen^{a, d} Emilie Hestbaek^{b, c}
Caroline Barkholt Kamp^{a, d} Markus Harboe Olsen^{a, e} Marie Rishede^f
Frederik Weischer Frandsen^{b, c} Sune Bo^c Stig Poulsen^c Per Sørensen^b
Anthony Bateman^g Sebastian Simonsen^{b, c}

^aCopenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ^bStolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Copenhagen, Denmark; ^cDepartment of Psychology, University of Copenhagen, Copenhagen, Denmark; ^dDepartment of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; ^eDepartment of Neuroanaesthesiology, The Neuroscience Centre, The Neuroscience Centre, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark; ^fDepartment of Psychiatry Mental Health Services East, Region Zealand, Roskilde, Denmark; ^gDepartment of Clinical, Educational, and Health Psychology, University College London, London, UK

Dear Editor,

Philips et al. [1] raise concerns regarding our recent randomised clinical trial (RCT) methodology [2]. In the current reply, we clarify their most important concerns.

First, our goal was to evaluate the relative differences between shorter- versus longer-term MBT durations. We acknowledge that the participants only received individual therapy every second week compared to weekly sessions in the original program [2, 3]. However, this scheduling was consistent for both the short- and long-term participants in our trial, ensuring any effects of the less intensive program would balance out between groups and not influence the relative effect. Our long-term program lasts for 14 months, compared to the original's 18 months [4]. While Philips et al. may argue that MBT is only effective when delivered in the exact 18-month format, or even longer, this belief is theoretical and unsupported by evidence. From a patient's perspective, 14 months of psychotherapy is still long-term.

If long-term psychotherapy is indeed crucial, we should observe a relative difference between our short- and long-term programs.

Second, performing within-group, pre-post analyses within a trial would correspond to treating the data as observational evidence, where it is impossible to separate any genuine changes due to treatment effects from the effect of regression to the mean or confounding by indication [5]. Therefore, we only performed between-group comparisons [6]. Philips et al. refer to an observational study performed by a member of their group [7] showing large pre-post changes on unblinded clinician-rated GAF scores, i.e., ratings that were not independent and where the clinicians obviously knew they rated a patient "post-treatment," which can significantly bias the results [8]. Philips et al. seem to suggest that the RCT is unnecessary and that pre-post changes can be compared between sites, ignoring

problems with confounding. Their claim is highly problematic due to inherent differences in participants, clinicians, assessment time points, co-interventions, access to healthcare, and other factors [5]. Observational evidence will always carry the risk of confounding by indication [9].

Third, Philips et al. [1] seem to believe that to demonstrate no difference between the two interventions, the non-inferiority design must be used, which we believe is a fundamental misconception of the RCT design. When an RCT reaches an estimated sample size based on an anticipated intervention effect and finds no difference, then the anticipated intervention effect can be rejected. For example, we used the minimal important difference of 3.5 points on the Zanarini Rating Scale for Borderline Personality Disorder as the anticipated intervention effect, the sample size of 166 participants was reached, and we found no evidence of a difference. Therefore, we can reject this anticipated intervention effect. Furthermore, a non-inferiority trial is a one-sided test, meaning it can demonstrate that the experimental intervention is “not worse” than the control intervention. We deliberately chose a superiority design, which is a two-sided test, allowing us to conclude that short-term MBT is both “not better” and “not worse” than long-term MBT. The superiority design includes everything (and more) that the non-inferiority design does. We

thank our colleagues, Philips et al. [1], for giving us an opportunity to provide a more elaborate discussion of our trial methodology.

Conflict of Interest Statement

A.B. declares receiving book royalties from Oxford University Press and American Psychiatric Publishing Inc. for publishing books on mentalizing, and payment for training courses in MBT from Anna Freud Centre, London. No other disclosures were reported.

Funding Sources

The MBT-RCT trial was funded by Trygfonden (S.J., Grant No.123488) and the Mental Health Services in the Capital Region of Denmark Research Fund (S.J. and S.S., Grant No. N/A). The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had access to all study data and had final responsibility for the decision to submit for publication.

Author Contributions

S.J. drafted the first version of this manuscript in close collaboration with J.C.J. and S.S. E.H., C.B.K., M.H.O., M.R., F.W.F., S.B., S.P., P.S., and A.B. commented on and approved the final version of the paper.

References

- Philips B, Leichsenring F, Jorret H, Karterud S. Flaws in the design of a randomized controlled trial comparing short-term versus long-term mentalization-based therapy for borderline personality disorder. *Psychother Psychosom*. 2024;1–2. [Accepted for publication]. <https://doi.org/10.1159/000540309>
- Juul S, Jakobsen JC, Hestbaek E, Jørgensen CK, Olsen MH, Rishede M, et al. Short-term versus long-term mentalization-based therapy for borderline personality disorder: a randomized clinical trial (MBT-RCT). *Psychother Psychosom*. 2023;92(5):329–39. <https://doi.org/10.1159/000534289>
- Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019; 20(1):196. <https://doi.org/10.1186/s13063-019-3306-7>
- Bateman A, Fonagy P. *Mentalization-based treatment for personality disorders: a practical guide*. Oxford, England: Oxford University Press; 2016.
- Cuijpers P, Weitz E, Cristea IA, Twisk J. Pre-post effect sizes should be avoided in meta-analyses. *Epidemiol Psychiatr Sci*. 2017;26(4):364–8. <https://doi.org/10.1017/S2045796016000809>
- Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, 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;22(1):497. <https://doi.org/10.1186/s13063-021-05450-y>
- Kvarstein EH, Pedersen G, Urnes Ø, Hummelen B, Wilberg T, Karterud S. Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—Does it make a difference? *Psychol Psychother*. 2015;88(1):71–86. <https://doi.org/10.1111/papt.12036>
- Juul S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med*. 2021; 26(3):109. <https://doi.org/10.1136/bmjebm-2020-111407>
- Collins R, Bowman L, Landray M, Peto R. The Magic of Randomization versus the Myth of Real-World Evidence. *New England Journal of Medicine*. 2020;382(7):674–678. <https://doi.org/10.1056/NEJMs1901642>

Paper 9

Short-term versus long-term psychotherapy for borderline personality disorder. A protocol for an individual patient data meta-analysis

Authors:

Sophie Juul^{1,2,3*}, Shelley McMain^{4,5}, Markus Harboe Olsen^{1,6}, Alexander Chapman^{7,8}, Johanne Pereira Ribeiro^{9,10}, Ole Jakob Storebø^{9,10}, Janice Kuo¹¹, Emilie Hestbaek^{2,3}, Caroline Barkholt Kamp^{1,12}, Marie Rishede¹³, Frederik Weischer Frandsen^{2,3}, Sune Bo³, Stig Poulsen³, Per Sørensen², Anthony Bateman¹⁴, Sebastian Simonsen², Janus Christian Jakobsen^{1,12}

Corresponding author:

Sophie Juul, sophie.juul@ctu.dk

Affiliations:

¹ Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark

² Mental Health Centre Stolpegaard, Copenhagen University Hospital - Mental Health Services CPH, Copenhagen, Denmark

³ Department of Psychology, University of Copenhagen, Copenhagen, Denmark

⁴ Centre for Addiction and Mental Health, Toronto, Ontario, Canada

⁵ Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

⁶ Department of Neuroanaesthesiology, Neuroscience Centre, Copenhagen University Hospital – Rigshospitalet, The Capital Region, Copenhagen, Denmark.

⁷ Department of Psychology, Simon Fraser University, Vancouver, British Columbia, Canada

⁸ DBT Centre of Vancouver, Vancouver, British Columbia, Canada

⁹ Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Region Zealand

¹⁰ Department of Psychology, Faculty of Health Sciences, University of Southern Denmark

¹¹ Department of Psychology, Palo Alto University, Palo Alto, California, USA

¹² Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

¹³ Department of Psychiatry Mental Health Services East, Region Zealand, Roskilde, Denmark

¹⁴ University College London, United Kingdom

Abstract

Background

The evidence for the optimal duration of psychotherapy for borderline personality disorder (BPD) is scarce. Two previous trials have compared different durations of psychotherapy. The first compared 6 months versus 12 months of dialectical behavior therapy for BPD (the FASTER trial). The second compared 5 versus 14 months of mentalization-based therapy for BPD (the MBT-RCT trial). The primary objective of the present study will be to provide an individual patient data meta-analysis (IPD-MA) of the results of the FASTER trial and the MBT-RCT trial by combining the two short-term groups and the two long-term groups, thereby providing greater statistical power than the individual trials. Accordingly, we will evaluate the overall evidence on the effects of short-term versus long-term psychotherapy for BPD and investigate whether certain subgroups might benefit from short-term versus long-term psychotherapy.

Methods

An IPD-MA of the FASTER trial and the MBT-RCT trial will be conducted. The primary outcome will be a composite of the proportion of participants with a suicide, a suicide attempt, or a psychiatric hospitalisation. The secondary outcome will be the proportion of participants with self-harm. Exploratory outcomes will be BPD symptoms, symptom distress, level of functioning, and quality of life. We will primarily assess outcomes at 15-months after randomization for the FASTER trial and at 16 months after randomization for the MBT-RCT trial. Predefined subgroups based on the design variables in the original trials will be tested for interaction with the intervention as follows: trial, sex (male compared to female), age (below or at 30 compared to above 30), and baseline level of functioning (Global Assessment of Functioning baseline score at 0-49 compared to 50-100).

Ethics and dissemination

The statistical analyses will be performed on anonymised trial data that have already been approved by the respective ethical committees that originally assessed the included trials. The final review will be published in a peer-reviewed scientific journal and the results will be presented at national seminars and international conferences.

PROSPERO registration number: CRD42024612840

Strengths and limitations of this study

- This will be the first individual patient data meta-analysis to assess the effects of short-term versus long-term psychotherapy for outpatients with borderline personality disorder.
- This protocol has been developed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Individual Patient Data (PRISMA-IPD) statement.
- Pooling data from two similar trials will allow us to explore potential subgroup analyses with more power than the individual trials.
- Pooling dialectical behavioral therapy and mentalization-based therapy can potentially introduce heterogeneity, which we will explore in the IPD.
- Both trials had missing outcome data, which will be a limitation of this IPD.

Introduction

Borderline personality disorder (BPD) is a serious and debilitating mental health disorder affecting 1 to 6% of the population [1]. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), BPD is characterised by emotional dysregulation, interpersonal problems, identity diffusion, impulsive behavior, chronic emptiness, intense anger, fear of abandonment, and dissociative symptoms, and it is associated with exceedingly high rates of self-harm and suicide [2-5]. Patients with BPD may be highly represented in primary care settings [6, 7] and mental health care settings [8]. In addition, BPD causes a significant socioeconomic burden, not only due to direct health care costs [9], but also due to indirect costs like social and occupational dysfunction [10]. All major guidelines for the treatment of BPD consider psychological interventions as the primary treatment of choice and assign pharmacological interventions an adjunctive role [11-14]. Even though the optimal duration of psychotherapy for BPD is unclear [15, 16], it is typically long-term, lasting no less than a year in duration [17]. Different short-term psychotherapies exist for patients with BPD, including brief dialectical behavior therapy (DBT) skills training for 20 weeks [18], emotion regulation group therapy (ERGT) for 14 weeks [19], systems training for emotional predictability and problem-solving (STEPPS) for 20 weeks [20], mentalization-based therapy (MBT) for 20 weeks [21] and most recently BPD Compass for 18 weeks [22]. However, to assess the optimal duration of psychotherapy, randomised clinical trials must randomise participants to different lengths of psychotherapy.

A noninferiority trial assessing the effects of 6 versus 12 months of DBT for BPD (the FASTER trial) was recently published [23]. In the FASTER trial, 240 participants were randomised to short-term or long-term DBT. Participants were adult outpatients with BPD and a minimum of one suicide attempt or act of self-harm within the past 2 months. The authors found that the noninferiority hypothesis was

supported for the primary outcome, total self-harm, and several secondary outcomes. No evidence of a difference was found for serious or non-serious adverse events.

A similar trial was recently published assessing the effects of 5 versus 14 months of MBT for BPD (the MBT-RCT trial) [21]. In the MBT-RCT trial, 166 participants were randomised to short-term MBT or long-term MBT. Participants were adult outpatients with subthreshold (at least four DSM-5 criteria) or diagnosed BPD. The authors found no evidence of a difference between short-term versus long-term MBT on the primary outcome, borderline symptoms, or any of the secondary outcomes. More participants in the long-term MBT group had a serious adverse event compared with short-term MBT primarily driven by a difference in psychiatric hospitalisations, but the result was not statistically significant.

In both the FASTER trial and the MBT-RCT trial, the short-term and the long-term interventions were relatively similar in lengths, the participants were all adult BPD outpatients, and the trials had similar methodologies which were pre-defined and protocolised [24, 25]. The primary objective of the present study will be to provide an individual patient data meta-analysis of the results of the FASTER trial and the MBT-RCT trial by combining the two short-term groups and the two long-term groups, thereby providing greater statistical power than the individual trials. Accordingly, we will evaluate the overall evidence on the effect of short-term versus long-term psychotherapy for BPD and investigate whether specific subgroups might benefit from short-term versus long-term psychotherapy.

Methods and Materials

This protocol follows the reporting guideline provided by the Preferred Reporting Items for Systematic Reviews and Meta- Analyses-IPD statement [26], and is registered in the PROSPERO database (registration number ID number 612840). The final IPD-MA will be conducted in accordance with this protocol, which was written and registered before analysis, and any deviations will be reported.

Study design

The design and results of the FASTER and MBT-RCT trials were published previously [21, 24, 25, 27, 28]. In short, the population in both trials were adults with BPD. The FASTER trial compared 6 months versus 12 months DBT for adult outpatients with BPD and assessed the effects every 3 months from randomisation to 24 months. The FASTER trial was a multi-centre trial conducted at two sites: the BPD Clinic at the Centre for Addiction and Mental Health (CAMH) in Toronto, Canada, and the Personality and Emotion Research Laboratory in conjunction with the DBT Centre in Vancouver, British Columbia (clinicaltrials.gov registration number: NCT02387736; February 2015) [25]. The MBT-RCT trial compared 5 versus 14 months MBT for adult outpatients with BPD and assessed the effects at 8, 16, and 24 months after randomisation. The MBT-RCT trial was a single-centre trial conducted at the Outpatient Clinic for Personality Disorders at Stolpegaard Psychotherapy Centre in The Capital Region of Denmark (clinicaltrials.gov registration number: NCT03677037; September 2018) [24, 27]. Separate ethics approval for this study was not required, as each trial was approved by ethics committees and authorities in each participating country.

Outcomes

Outcomes for the individual patient data meta-analysis will be defined as primary, secondary, or exploratory.

Primary outcome

- Proportion of participants with either a suicide, a suicide-attempt, or a psychiatric hospitalisation (composite outcome, dichotomous data)

Secondary outcome

- Proportion of participants with deliberate self-harm (dichotomous data). Self-harm will be defined as deliberate acts of self-harm with the conscious intent to cause self-injury or to risk self-injury, e.g. alcohol and/or drug consumption with the intention of self-injury, poison, burning, scratching/cutting, stabbing/puncture, gun shooting, hanging, strangling.

Exploratory outcomes

- BPD symptom severity assessed with the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) or the Borderline Symptom List (BSL-23) (continuous data).
- Quality of life assessed with the Short-form Health Survey (SF-36) mental component or the Euroqol-5D-5 Level (EQ-5D-5 L) (continuous data)
- Level of functioning assessed with the Global Assessment of Functioning (GAF) or the Social Adjustment Scale Self-Report (SAS-SR) (continuous data).
- Symptom distress assessed with the Global Severity Index (GSI) of the Symptom Checklist-90 (SCL-90) (continuous data)
- Each of the components of the primary composite outcome will analysed as exploratory outcomes.

Timepoints of assessment

Our primary time point will be the 15-months after randomization time-point for the FASTER trial and the 16 months after randomization time-point for the MBT-RCT trial. As a secondary analysis we will assess outcomes at 24 months after randomisation in both trials.

Risk of bias and certainty of evidence

Risk of bias will be assessed with the Cochrane Risk of Bias tool – version 2 [29], and certainty of the evidence will be assessed with the Grading of Recommendations, Assessment, Development, and Evaluations [30] guideline by two independent review authors (JPR, OJS) from Center for Evidence-Based Psychiatry, Region Zealand Psychiatry, Denmark, who were not involved in the two trials.

Statistical analysis plan

General analysis principles

All analyses will be conducted according to the intention-to-treat principle. We will both assess if the thresholds for statistical significance and clinical significance are crossed and Bayes factor calculations will be reported [31]. Assessments of clinical significance will be based on the anticipated intervention effects used in the sample size/ power estimations in the MBT trial [27, 31]. Our primary conclusion will be based on the primary outcome, so all tests of statistical significance (including subgroup analyses) will be two-sided with a type I error risk of 5% [31].

It is generally acknowledged that regression analyses ought to be adjusted for the stratification variables used in the randomisation [32-34]. The FASTER trial used ‘site’ as a stratification variable, and the MBT-RCT used sex (male/female) and high versus low baseline scores on the primary

outcome, ZAN-BPD, as the stratification variables. We will adjust all regression analyses for ‘trial site’. We will also assess whether there are significant interactions between trial interventions and ‘site’ (see ‘**Assessments of underlying statistical assumptions**’).

We will also perform the following subgroup analyses for all primary and secondary outcomes: trial, sex (male compared to female), age (below or at 30 compared to above 30) baseline level of functioning (GAF baseline score at 0-49 compared to 50-100). We will present the subgroup results in forest plots. We will additionally include the continuous subgroup variables as continuous variables in regression analyses to assess if the tests of interaction are significant.

Analysis of dichotomous data

Dichotomised outcomes will be presented as proportions of participants in each group with events and risk ratios with 95% confidence intervals. Dichotomous outcomes will be analysed using generalised linear models using a log link function. If the model does not converge, we will try a different optimisation technique or use a logit-link model if possible.

Analysis of continuous data

Continuous outcomes will be presented as means, standard deviations, and 95% confidence levels in each group and analysed using linear regression.

Handling of missing data

All randomised participants with available data will be included in the primary analysis according to the intention-to-treat principle. We anticipate that the proportion of missing values will be less than 5% for the dichotomous outcomes. We expect large proportions of missing data for the continuous

outcomes. Missing data will be handled according to the recommendation by Jakobsen et al [35]. In short, in secondary analyses, we will consider using multiple imputation and/or present best-worst and worst best case scenarios if it is not valid to ignore missing data [35]. Best-worst and worst-best case scenarios assess the potential range of impact of the missing data for the trial results [35]. In the ‘best-worst’ case scenario, it is assumed that all patients lost to follow-up in the short-term group have had a beneficial outcome (have survived, had no poor functional outcome, and so forth), and all those with missing outcomes in the long-term group have had a harmful outcome (have not survived, have had poor functional outcome, and so forth) [35]. Conversely, in the ‘worst-best’ case scenario, it is assumed that all patients who were lost to follow-up in the experimental group have had a harmful outcome, and that all those lost to follow-up in the control group have had a beneficial outcome [35].

Aggregate data meta-analysis

The two trials used two different scales to assess BPD symptom severity, quality of life, and level of functioning, and we will pool them using standardized mean difference meta-analysis statistics. We will undertake the meta-analyses according to the *Cochrane Handbook of Systematic Reviews of Interventions* [36], Keus et al. [37], and the eight-step assessment suggested by Jakobsen et al. [38]. We will use the statistical software Stata version 16 [39] (command: meta) and R to analyse data. We will assess the intervention effect with both random-effects and fixed-effect meta-analysis. We will primarily report the most conservative results (highest p values) of the two and consider the less conservative results as a sensitivity analysis [38]. Our primary conclusion will be based on the primary outcome; hence, we will not adjust the significance level according to multiplicity.

Trial Sequential Analysis

Traditional meta-analysis runs the risk of random errors due to sparse data and repetitive testing of accumulating data when updating reviews. To control the risks of type I errors and type II errors, we will perform Trial Sequential Analysis on the primary and secondary outcomes, in order to calculate the diversity-adjusted required information size (DARIS; that is, the number of participants needed in a meta-analysis to detect or reject a certain intervention effect) and the cumulative Z-curve's breach of relevant trial sequential monitoring boundaries [40-48]. A more detailed description of Trial Sequential Analysis software can be found in the manual [46] and at <http://www.ctu.dk/tsa/>. For the primary and secondary outcomes, which are both dichotomous outcomes, we will estimate the required information size based on the observed proportion of patients with an outcome in the control group (the cumulative proportion of patients with an event in the control groups relative to all patients in the control groups), a relative risk reduction or a relative risk increase of 25%, an alpha of 5% for all our outcomes, a beta of 20%, and the observed diversity as suggested by the trials in the meta-analysis.

Assessments of underlying statistical assumptions

We will systematically assess underlying statistical assumptions for all statistical analyses [49, 50]. For *all* regression analyses, both primary and secondary, we will test for major interactions between 'trial site' and the intervention variable and whether the interactions shows a clinically important effect. If it is concluded that the interaction is significant, we will consider both presenting an analysis separately for each site and an overall analysis including the interaction term in the model [49, 50].

Assessments of underlying statistical assumptions for dichotomous outcomes

We will assess if the deviance by the degrees of freedom is significantly larger than 1 to assess for relevant overdispersion, i.e., the presence of greater variability (statistical dispersion) in a data set

than would be expected based on a given statistical model. If this is the case, we will consider using maximum likelihood estimate of the dispersion parameter [50].

Assessments of underlying statistical assumptions for continuous outcomes

We will visually inspect quantile-quantile plots of the residuals [51, 52] to assess if the residuals are normally distributed and use residuals plotted against covariates and fitted values [51, 52] to assess for homogeneity of variances. If the plots show deviations from the model assumptions, we will consider transforming the outcome e.g. using log transformation or square root and/or use robust standard errors [51, 52].

Statistical reports

Data on all outcomes will be analysed by two independent individuals [50]. Two independent statistical reports will be sent to the chief principal investigator (SJ) and will be shared with the author group. If there are discrepancies between the two primary statistical reports, then possible reasons for that will be identified, and the author group will decide which is the most correct result. A final statistical report will be prepared, and all three statistical reports will be published as supplementary material [50].

Public and patient involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Ethics and dissemination

The statistical analyses will be performed on fully anonymised trial data that have already been approved by the respective ethical committees that originally assessed the included trials in the respective countries. The final individual patient data meta-analysis will be published in a peer-reviewed scientific journal and the results will be presented at Danish national seminars and at international conferences.

Statements

Patients and the Public Statement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research

Conflict of Interest Statement

All authors declare that they have no conflicts of interest

Funding Sources

This work was supported by the Mental Health Services of the Capital Region of Denmark Research Fund. The funder has no role in the design, data collection, data analysis, and reporting of this study.

Author Contributions

SJ wrote up the first manuscript draft in close collaboration with JCJ. All authors read, commented on, and approved the final version of the manuscript.

Ethics statements

Not applicable for this project.

Data statement

Not applicable for this protocol article.

References

1. Torgersen S, Kringlen E, Cramer V. The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry*. 2001;58(6):590-6.
2. Cheng AT, Mann A, Chan K. Personality disorder and suicide: A case-control study. *Br J Psychiatry*. 1997;170(5):441-6.
3. Lesage AD, Boyer R, Grunberg F, Vanier C, Morissette R, Ménard-Buteau C, et al. Suicide and mental disorders: a case-control study of young men. *Am J Psychiatry*. 1994.
4. Black DW, Blum N, Pfohl B, Hale N. Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *J Personal Dis*. 2004;18(3: Special issue):226-39.
5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub; 2013. Report No.: 0890425574.
6. Gross R, Olfson M, Gameroff M, Shea S, Feder A, Fuentes M, et al. Borderline personality disorder in primary care. *Arch Intern Med*. 2002;162(1):53-60.
7. Ellison WD, Rosenstein LK, Morgan TA, Zimmerman M. Community and clinical epidemiology of borderline personality disorder. *Psychiatric Clinics*. 2018;41(4):561-73.
8. Zanarini MC, Frankenburg FR, Hennen J, Silk KR. Mental health service utilization by borderline personality disorder patients and Axis II comparison subjects followed prospectively for 6 years. *J Clin Psychiatry*. 2004.
9. Soeteman DI, Roijen LH-v, Verheul R, Busschbach JJ. The economic burden of personality disorders in mental health care. *J Clin Psychiatry*. 2008;69(2):259.
10. Sansone RA, Sansone LA. Employment in borderline personality disorder. *Innov Clin Neurosci*. 2012;9(9):25.
11. National Institute for Health and Care Excellence (NICE). Borderline personality disorder: recognition and management. Clinical guideline [CG78]. 2009:Available at: <https://www.nice.org.uk/guidance/cg78> [Accessed September 19, 2022].
12. American Psychological Association. Diagnostic and statistical manual of mental disorders - DSM-5. *Am Psychiatric Assoc*. 2013;21(21):591-643.
13. National Health Medical Research Council. Clinical practice guideline for the management of borderline personality disorder. NHMRC Melbourne; 2012.

14. Stoffers-Winterling JM, Ribeiro JP, Kongerslev MT, Völlm BA, Mattivi JT, Faltinsen E, et al. Pharmacological interventions for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*. 2022(11).
15. Juul S, Poulsen S, Lunn S, Sørensen P, Jakobsen JC, Simonsen S. Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Systematic Reviews*. 2019;8(1):169.
16. Juul S, Jakobsen JC, Jørgensen CK, Poulsen S, Sørensen P, Simonsen S. The difference between shorter- versus longer-term psychotherapy for adult mental health disorders: a systematic review with meta-analysis. *BMC Psychiatry*. 2023;23(1):438.
17. Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020(5).
18. McMain SF, Guimond T, Barnhart R, Habinski L, Streiner DL. A randomized trial of brief dialectical behaviour therapy skills training in suicidal patients suffering from borderline disorder. *Acta Psychiatr Scand*. 2017;135(2):138-48.
19. Gratz K, Tull M, Levy R. Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychol Med*. 2014;44(10):2099.
20. Blum N, St. John D, Pfohl B, Stuart S, McCormick B, Allen J, et al. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *Am J Psychiatry*. 2008;165(4):468-78.
21. Juul S, Jakobsen JC, Hestbaek E, Jørgensen CK, Olsen MH, Rishede M, et al. Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder: A Randomized Clinical Trial (MBT-RCT). *Psychotherapy and Psychosomatics*. 2023;92(5):329-39.
22. Sauer-Zavala S, Southward MW, Fruhbauerova M, Semcho SA, Stumpp NE, Hood CO, et al. BPD compass: A randomized controlled trial of a short-term, personality-based treatment for borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment*. 2023.
23. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022:1-16.

24. Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196.
25. McMain SF, Chapman AL, Kuo JR, Guimond T, Streiner DL, Dixon-Gordon KL, et al. The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial protocol. *BMC Psychiatry*. 2018;18:1-16.
26. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. *Jama*. 2015;313(16):1657-65.
27. Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, 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;22(1):497.
28. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022;91(6):382-97.
29. Sterne J, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
30. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64(4):395-400.
31. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC medical research methodology*. 2014;14(34).
32. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Statistics in medicine*. 2011;31:328-40.
33. Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ*. 2012;345:e5840.
34. Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. *BMC Med Res Methodol*. 2003;3:26.

35. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.
36. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019)*. Cochrane, 2019. 2019.
37. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error matrix approach for overviewing available evidence. *BMC medical research methodology*. 2010;10(1):90.
38. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC medical research methodology*. 2014;14(1):120.
39. StataCorp. *Stata Statistical Software: Release 16 2019* [College Station, TX: StataCorp LLC <http://www.stata.com>].
40. Copenhagen Trial Unit. TSA - Trial Sequential Analysis. <http://www.ctu.dk/tsa/> [Web page]. [
41. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of clinical epidemiology*. 2008;61(1):64-75.
42. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *Journal of clinical epidemiology*. 2008;61(8):763-9.
43. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive—trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *International journal of epidemiology*. 2008;38(1):287-98.
44. Thorlund K, Devereaux P, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *International journal of epidemiology*. 2008;38(1):276-86.
45. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC medical research methodology*. 2009;9(1):86.
46. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, C. G. User manual for trial sequential analysis (TSA). http://www.ctu.dk/tsa/files/tsa_manualpdf. 2011.

47. Thorlund K, Anema A, Mills E. Interpreting meta-analysis according to the adequacy of sample size. An example using isoniazid chemoprophylaxis for tuberculosis in purified protein derivative negative HIV-infected individuals. *Clinical Epidemiology*. 2010;2:57.
48. Imberger G, Thorlund K, Gluud C, Wetterslev J. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. *BMJ open*. 2016;6(8):e011890.
49. Nielsen EE, Nørskov AK, Lange T, Thabane L, Wetterslev J, Beyersmann J, et al. Assessing assumptions for statistical analyses in randomised clinical trials. *BMJ Evidence-Based Medicine*. 2019;24(5):185-9.
50. Nørskov AK, Lange T, Nielsen EE, Gluud C, Winkel P, Beyersmann J, et al. Assessment of assumptions of statistical analysis methods in randomised clinical trials: the what and how. *BMJ Evid Based Med*. 2021;26(3):121-6.
51. Frydenberg SJaM. *An Introduction to Stata for Health Researchers*, Fourth Edition. Stata Press. 2014.
52. Oliveira AG. *Biostatistics Decoded*. Wiley. 2013.

Paper 10

Short-term versus long-term psychotherapy for borderline personality disorder: an individual patient data pooled analysis of two randomized clinical trials

Sophie Juul^{1,2,3*}, Shelley McMain^{4,5}, Markus Harboe Olsen^{1,6}, Alexander Chapman^{7,8}, Johanne Pereira Ribeiro^{9,10}, Ole Jakob Storebø^{9,10}, Janice Kuo¹¹, Caroline Barkholt Kamp¹, Emilie Hestbaek^{2,3}, Marie Rishede¹³, Frederik Weischer Frandsen^{2,3}, Sune Bo³, Stig Poulsen³, Per Sørensen², Anthony Bateman¹⁴, Sebastian Simonsen², Janus Christian Jakobsen^{1,12}

Corresponding author:

Sophie Juul, sophie.juul@ctu.dk

Affiliations:

¹ Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark

² Mental Health Centre Stolpegaard, Copenhagen University Hospital - Mental Health Services CPH, Copenhagen, Denmark

³ Department of Psychology, University of Copenhagen, Copenhagen, Denmark

⁴ Centre for Addiction and Mental Health, Toronto, Ontario, Canada

⁵ Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

⁶ Department of Neuroanaesthesiology, Neuroscience Centre, Copenhagen University Hospital – Rigshospitalet, The Capital Region, Copenhagen, Denmark.

⁷ Department of Psychology, Simon Fraser University, Vancouver, British Columbia, Canada

⁸ DBT Centre of Vancouver, Vancouver, British Columbia, Canada

⁹ Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Region Zealand

¹⁰ Department of Psychology, Faculty of Health Sciences, University of Southern Denmark

¹¹ Department of Psychology, Palo Alto University, Palo Alto, California, USA

¹² Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

¹³ Department of Psychiatry Mental Health Services East, Region Zealand, Roskilde, Denmark

¹⁴ University College London, United Kingdom

Abstract

Background: Borderline personality disorder is a severe mental health disorder affecting up to 6% of the general population. Psychotherapy is considered the first-line treatment for borderline personality disorder, but the optimal duration of psychotherapy needs further investigation.

Methods: An individual patient data pooled analysis of two randomized clinical trials comparing short-term versus long-term psychotherapy for borderline personality disorder was conducted using data from two similar trials: the FASTER trial and the MBT-RCT trial. The primary outcome was a composite outcome of either a death by suicide, a suicide attempt, or a psychiatric hospitalization. The secondary outcome was deliberate, non-suicidal, self-harm. Exploratory outcomes were borderline symptoms, quality of life, level of functioning, and symptom distress. Predefined subgroups based on the design variables in the original trials were tested for interaction with the intervention: trial, sex, age, and baseline level of functioning. Outcomes were primarily assessed at 15-16 months after randomization.

Findings: The primary analysis included 406 participants, with 204 assigned to short-term psychotherapy and 202 assigned to long-term psychotherapy. The proportion of participants with either a death by suicide, a suicide attempt, or a psychiatric hospitalization was 11 of 204 participants (5.4%) in the short-term group, and 27 of 202 participants (13.4%) in the long-term group (relative risk [RR], 0.41; 95% confidence interval [CI]: 0.21 to 0.79; $P=0.008$) primarily driven by a difference in psychiatric hospitalizations. We found no evidence of a difference between short-term and long-term psychotherapy on deliberate, non-suicidal, self-harm (RR 0.80; 95% CI: 0.53 to 1.21), or any of the exploratory outcomes. Results were consistent across the predefined subgroup analyses.

Interpretation: In outpatients with borderline personality disorder, long-term psychotherapy increased the risk of psychiatric hospitalizations at 15-16 months compared to short-term psychotherapy.

Funding: Canadian Institute of Health Research (CIHR FRN 133428) funded the FASTER trial. TrygFonden A/S (grant no. 123488), and the Research Fund of the Mental Health Services in the Capital Region of Denmark (grant no. N/A) funded the MBT-RCT trial. The Research Fund of the Mental Health Services in the Capital Region of Denmark (grant no. N/A) funded the present study.

PROSPERO registration: CRD42024612840

Research in context

Evidence before this study

To assess the optimal duration of psychotherapy, clinical trials must randomize participants to different durations of the same psychotherapy type. Two such trials currently exist for borderline personality disorder: The first is a noninferiority trial assessing the effects of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder (the FASTER trial). The second is a similar trial assessing the effects of 5 versus 14 months of mentalization-based therapy for borderline personality disorder (the MBT-RCT trial). Both trials did not find evidence of a difference between short-term and long-term psychotherapy on their respective primary outcomes.

Added value of this study

Pooling individual patient data from the two previous similar trials allowed us to increase the total sample size and introduced the opportunity to conduct more detailed subgroup analyses. In outpatients with borderline personality disorder, long-term psychotherapy increased the risk of psychiatric hospitalizations at 15-16 months compared to short-term psychotherapy. Results were consistent across the predefined subgroup analyses.

Implications of all the available evidence

Our findings indicate that short-term psychotherapy may be a desirable alternative to long-term psychotherapy for outpatients with borderline personality disorder, particularly in mental health settings with scarce treatment resources and limited access to evidence-based care.

Introduction

Borderline personality disorder is a severe mental health disorder affecting up to 6% of the general population [1]. Borderline personality disorder is characterized by symptoms such as emotion dysregulation, interpersonal problems, identity diffusion, impulsive behavior, chronic emptiness, intense anger, fear of abandonment, and dissociative symptoms, and it is associated with exceedingly high rates of self-harm and suicide [2-5]. Patients with borderline personality disorder are highly represented in primary care settings [6, 7] and mental health care settings [8]. In addition, borderline personality disorder causes a significant economic burden, not only because of direct health care costs [9], but also because of indirect costs due to social and occupational dysfunction [10, 11].

Psychotherapy is considered the first-line treatment for borderline personality disorder [12] and is typically regarded as a long-term treatment with a duration of one year or more in clinical practice [13]. The optimal duration of psychotherapy for borderline personality disorder, however, needs further investigation. Two research groups have previously conducted two comparable trials evaluating psychotherapies delivered over two similar durations for borderline personality disorder.

The first is a noninferiority trial assessing the effects of 6 versus 12 months of dialectical behavioral therapy for borderline personality disorder (the FASTER trial) [14]. In the FASTER trial, 240 participants were randomized to short-term or long-term comprehensive dialectical behavioral therapy. Dialectical behavioral therapy, originally developed by Marsha Linehan, is a third-wave cognitive behavioral therapy comprised of a broad range of cognitive, behavioral, and problem-solving therapeutic interventions, including skills training and in-session coaching, chain analysis, contingency management, and exposure [15, 16]. Participants in the FASTER trial were adult outpatients with borderline personality disorder and a minimum of two suicide attempts or episodes

of non-suicidal self-injury in the past five years, with at least one occurring in the past two months. The authors found that the noninferiority hypothesis was supported for the primary outcome, total self-harm, and several secondary outcomes. No evidence of a difference was found for serious or non-serious adverse events.

The second is a similar randomized clinical superiority trial assessing the effects of 5 versus 14 months of mentalization-based therapy for 166 outpatients with borderline personality disorder (the MBT-RCT trial) [17]. Mentalization-based therapy, originally developed by Peter Fonagy and Anthony Bateman, is a psychodynamic therapy rooted in attachment theory [18, 19] focusing on re-establishing mentalization, i.e. the capacity to understand one's own and others' internal mental states such as beliefs and intentions, in times of emotional distress [20-22]. The authors found no evidence of a difference between short-term versus long-term MBT on the primary outcome, borderline symptoms, or any of the secondary outcomes.

In this report, we present an individual patient data pooled analysis of two randomized clinical trials comparing of the results from the FASTER and MBT-RCT trials, combining the two intervention groups and the two comparator groups to provide greater statistical power than the individual trials.

Methods

Study design and participants

This individual patient data pooled analysis of two randomized clinical trials complies with our protocol [23], which was registered before analysis (PROSPERO ID: CRD42024612840; November 22, 2024), and any deviations are reported. We included individual patient data from the FASTER and

MBT-RCT trials. Separate ethics approval for this study was not required, as each trial was approved by ethics committees and authorities in each participating country.

The design and results of the FASTER and MBT-RCT trials have been published previously [17, 24-27]. The populations in both trials were adult outpatients with borderline personality disorder (**Supplementary material p. 2**). The intervention group was 6 months of dialectical behavioral therapy in the FASTER trial and 5 months of mentalization-based therapy in the MBT-RCT trial. The control group was 12 months of dialectical behavioral therapy in the FASTER trial and 14 months of mentalization-based therapy in the MBT-RCT trial.

Outcomes

Outcomes were predefined as primary, secondary, or exploratory [23].

Primary outcome

- Proportion of participants with either a death by suicide, a suicide attempt, or a psychiatric hospitalization (composite outcome, dichotomous data).

Secondary outcome

- Proportion of participants with deliberate self-harm (dichotomous data). Self-harm was defined as deliberate acts of self-harm with the conscious intent to cause self-injury or to risk self-injury, e.g. alcohol and/or drug consumption with the intention of self-injury, poison, burning, scratching/cutting, stabbing/puncture, gun shooting, hanging and/or strangling.

Exploratory outcomes

- Borderline personality disorder symptom severity assessed with the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) or the Borderline Symptom List (BSL-23) (continuous data).
- Quality of life assessed with the Short-form Health Survey (SF-36) mental component or the Euroqol-5D-5 Level (EQ-5D-5 L) (continuous data).
- Level of functioning assessed with the Global Assessment of Functioning (GAF) or the Social Adjustment Scale Self-Report (SAS-SR) (continuous data).
- Symptom distress assessed with the Global Severity Index (GSI) of the Symptom Checklist-90 (SCL-90) (continuous data).
- Each of the components of the primary composite outcome was analyzed as an exploratory outcome.

Timepoints of assessment

Our primary time point was the 15-month time point after randomization for the FASTER trial and the 16-month time point after randomization for the MBT-RCT trial. We chose these time-points because they were the closest to the ‘end of treatment’ for the long-term group, ensuring that all randomized participants had terminated their allocated intervention. As a secondary analysis, we assessed outcomes at 24 months after randomization in both trials.

Statistical analyses

A detailed, pre-defined statistical analysis plan has previously been written [23]. In short, all analyses included the intention-to-treat population from the FASTER and MBT-RCT trial. Dichotomous outcomes (log link) and continuous outcomes (identity link) were analyzed using generalized linear models [28]. All regression analyses were adjusted for ‘site.’ Our primary conclusion was based on our primary outcome, so all tests of statistical significance (including subgroup analyses) were two-

sided with a type I error risk of 5% [29]. We systematically assessed the underlying statistical assumptions for all statistical analyses [30, 31]. We performed the following subgroup analyses for all primary and secondary outcomes: trial, sex (male compared to female), age (below or at 30 compared to above 30) baseline level of functioning (GAF baseline score at 0-49 compared to 50-100). Data on primary and secondary outcomes were analyzed in duplicate and independently (by MHO and JCJ). Missing data was handled according to the recommendations by Jakobsen et al. [32]. Percent of maximum possible (POMP) was used to analyze outcomes where multiple scales were analysed [2]. Thus, mean difference for those analyses is mean difference in POMP, and not in the actual scale values. For SAS-SR and GAF they had reverse interpretation, and SAS-SR was reversed during POMP transformation. Further information about the analyses of the exploratory outcomes can be found in the supplementary material.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Participants and interventions

From February 2015 to June 2017 for the FASTER trial, and from October 2018 to December 2020, for the MBT-RCT, the two trials randomized 406 outpatients with borderline personality disorder at two different sites in Canada and one in Denmark. The primary analysis included a total of 338 participants, of whom 168 were assigned to short-term psychotherapy and 170 were assigned to long-term psychotherapy. The participant characteristics were similar at baseline (**Supplementary**

material p. 3). The results at 24 months after randomization are reported in the statistical report (**Supplementary material p. 8).**

Primary outcome

In total, 11 of 168 participants (6.5%) in the short-term group and 27 of 170 participants (15.9%) in the long-term group died by suicide, had a suicide attempt or a psychiatric hospitalization (relative risk [RR], 0.41; 95% confidence interval [CI], 0.21 to 0.79; $P=0.008$) at 15-16 months after randomization, mainly driven by a difference in psychiatric hospitalizations (RR 0.43; 95% CI, 0.22 to 0.82; $P=0.011$) (**Supplementary material p. 8).** Data on the primary outcome were missing for 68 of 406 participants (16.7%) (**Supplementary material p. 8).** None of the preplanned tests of interaction for subgroups, including trial, sex, age, and baseline level of functioning showed evidence of a difference (**Figure 1).**

--- INSERT FIGURE 1 HERE ---

Figure 1. Forest plot of subgroup analyses for the primary outcome

Secondary outcome

In total, 32 of 152 participants (21.1%) in the short-term group and 39 of 149 participants (26.2%) in the long-term group had at least one episode of deliberate, non-suicidal, self-harm (RR 0.80; 95% CI, 0.53 to 1.21; $P=0.297$) at 15-16 months after randomization (**Supplementary material p. 8).** Data on the secondary outcome were missing for 105 of 406 participants (25.9%) (**Supplementary material p. 8).** None of the preplanned tests of interaction for subgroups, including trial, sex, age, and baseline level of functioning, showed evidence of a difference (**Figure 2).**

--- INSERT FIGURE 2 HERE ---

Figure 2. Forest plot of subgroup analyses for the secondary outcome

Exploratory outcomes

Borderline symptoms

Mixed effects linear regression showed no evidence of a difference between short-term and long-term psychotherapy when assessing borderline symptoms (MD -2.93; 95% CI -7.24 to 1.38; $P = 0.184$). Data were missing for 109 of 406 participants (26.8%) (**Supplementary material p. 8**). Borderline symptoms were assessed using the BSL-23 in the FASTER trial and ZAN-BPD in the MBT-RCT trial.

Quality of life

Mixed effects linear regression showed no evidence of a difference between short-term and long-term psychotherapy when assessing quality of life (MD 0.45; 95% CI -2.21 to 3.10; $P = 0.740$). Data were missing for 122 of 406 participants (30%) (**Supplementary material p. 8**). Quality of life was assessed using the EQ-5D-5L in the FASTER trial and SF-36 mental component in the MBT-RCT trial.

Level of functioning

Mixed effects linear regression showed no evidence of a difference between short-term and long-term psychotherapy when assessing level of functioning (MD -0.15; 95% CI -2.9 to 2.6; $P = 0.916$). Data were missing for 110 of 406 participants (27.1%) (**Supplementary material p. 8**). Level of functioning was assessed using SAR-SR in the FASTER trial and GAF in the MBT-RCT trial.

Symptom distress

Mixed effects linear regression showed no evidence of a difference between short-term and long-term psychotherapy when assessing level of functioning (MD -0.01; 95% CI -16 to 0.13; $P = 0.854$). Data were missing for 122 of 406 participants (30%) (**Supplementary material p. 8**). Symptom distress was assessed using the SCL-90 in both trials.

Outcomes at 24 months after randomization

We found no evidence of a difference between short-term versus long-term psychotherapy when assessing any of the outcomes at 24 months after randomization (**Supplementary material p. 8**).

Discussion

In this individual patient data pooled analysis of the FASTER and the MBT-RCT trials, we found that long-term psychotherapy increased the risk of psychiatric hospitalizations at 15-16 months follow-up compared with short-term psychotherapy. We found no evidence of a difference between short-term and long-term psychotherapy on any of the remaining outcomes. Our analyses showed limited signs of heterogeneity, and the results were consistent across the predefined subgroup analyses.

Our results have important clinical implications for the treatment of borderline personality disorder, as the high prevalence of the disorder, combined with the substantial costs associated with conventional long-term treatments, presents a major barrier to accessing psychotherapies. Hence, reducing the duration of psychotherapy in outpatient settings could improve patient outcomes and availability of care. Our findings challenge the common belief that individuals with borderline

personality disorder necessarily require longer-term psychotherapy and provide empirical support for the use of short-term psychotherapy.

The underlying cause of the increased risk of psychiatric hospitalizations in the long-term group remains unclear. Several theoretical explanations are plausible, but none are currently supported by empirical evidence. Future research should investigate this phenomenon through a mechanistic lens.

Our study has several strengths. First, our methodology was pre-registered and pre-defined in our published protocol prior to analyzing the results [24, 26]. Second, we assessed a ‘hard’ patient-important outcome as the primary outcome, i.e. patient-important binary outcomes that are conclusive regarding the disease progression, demonstrating a patient’s feelings, functionality, or survival [33], which is rare in psychiatric trials [34, 35]. Defining ‘hard’ outcomes as primary outcomes may solve some of the problems known with symptom scales [34, 35] as the patient importance is unquestionable, the results are objective and simple to interpret, and the risk of incomplete outcome data is limited (e.g. when using registries for data collection). Third, the two trials have a high degree of external validity, as they were conducted in real-world outpatient settings, suggesting that short-term psychotherapy for borderline personality disorder could be acceptable for implementation in routine clinical settings with the appropriate training and supervision. Fourth, our analyses showed limited signs of heterogeneity as all subgroup analyses showed no evidence of a difference. Fifth, all participants were assessed at the same time points regardless of their allocated intervention.

Our study also has limitations. We pooled two comparable trials, but they assessed two different types of psychotherapy. Both dialectical behavioral therapy and mentalization-based therapy in longer-term formats (12-18 months) are considered evidence-based for borderline personality disorder [13].

Dialectical behavioral therapy is considered a third-wave cognitive behavioral therapy, while mentalization-based therapy is rooted in psychodynamic theory. While both psychotherapies vary in their therapeutic focus and stance, they also share some common features, including focusing on emotion regulation, increasing distress tolerance, and reducing maladaptive behavior patterns using various therapeutic techniques. Furthermore, the concept of starting each therapy with psychoeducation followed by developing a case formulation is shared by both psychotherapies [36]. We observed no evidence of heterogeneity among the trials, suggesting that, despite theoretical differences between the psychotherapies, the relative effects of short-term and long-term psychotherapy appear to be similar.

Conclusion

In outpatients with borderline personality disorder, long-term psychotherapy increased the risk of psychiatric hospitalizations at 15-16 months compared to short-term psychotherapy. There seems to be no evidence of a difference between short-term and long-term psychotherapy on other outcomes.

Contributors

SJ wrote up the first manuscript draft in close collaboration with JCJ. MHO and JCJ independently performed the statistical analyses. SJ, JCJ, and MHO had full access to all the trial data and take responsibility for the integrity of the data and the accuracy of the data analysis. SM, MHO, AC, JPR, OJS, JK, CBK, EH, MR, FWF, SB, SP, PS, AB, and SS read, commented on, and approved the final version of the manuscript.

Data sharing

Access to deidentified data from the two trials may be obtained by contacting the respective corresponding authors, SJ or SM. All requests for data access must be approved by the IRBs of their institutions in Denmark and Canada and the institution with whom the primary investigator is affiliated. Data access will be governed by a data sharing agreement executed between their institutions and the institution with whom the primary investigator is affiliated.

Declaration of interests

All authors declare that they have no conflicts of interest

References

1. Torgersen S, Kringlen E, Cramer V. The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry*. 2001;58(6):590-6.
2. Cheng AT, Mann A, Chan K. Personality disorder and suicide: A case-control study. *Br J Psychiatry*. 1997;170(5):441-6.
3. Lesage AD, Boyer R, Grunberg F, Vanier C, Morissette R, Ménard-Buteau C, et al. Suicide and mental disorders: a case-control study of young men. *Am J Psychiatry*. 1994.
4. Black DW, Blum N, Pfohl B, Hale N. Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *J Personal Dis*. 2004;18(3: Special issue):226-39.
5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub; 2013. Report No.: 0890425574.
6. Gross R, Olfson M, Gameroff M, Shea S, Feder A, Fuentes M, et al. Borderline personality disorder in primary care. *Arch Intern Med*. 2002;162(1):53-60.
7. Ellison WD, Rosenstein LK, Morgan TA, Zimmerman M. Community and clinical epidemiology of borderline personality disorder. *Psychiatric Clinics*. 2018;41(4):561-73.
8. Zanarini MC, Frankenburg FR, Hennen J, Silk KR. Mental health service utilization by borderline personality disorder patients and Axis II comparison subjects followed prospectively for 6 years. *J Clin Psychiatry*. 2004.
9. Soeteman DI, Roijen LH-v, Verheul R, Busschbach JJ. The economic burden of personality disorders in mental health care. *J Clin Psychiatry*. 2008;69(2):259.
10. Sansone RA, Sansone LA. Employment in borderline personality disorder. *Innov Clin Neurosci*. 2012;9(9):25.
11. Hastrup LH, Jennum P, Ibsen R, Kjellberg J, Simonsen E. Societal costs of Borderline Personality Disorders: A matched-controlled nationwide study of patients and spouses. *Acta Psychiatrica Scandinavica*. 2019;140(5):458-67.
12. National Institute for Health and Care Excellence. Borderline personality disorder: recognition and management [CG78]. 2009:Available at: <https://www.nice.org.uk/guidance/cg78/resources/2018-surveillance-of-personality-disorders-nice-guidelines-cg77-and-cg78-4906490080/chapter/Surveillance-decision?tab=evidence>.

13. Storebø OJ, Stoffers-Winterling JM, Völlm BA, Kongerslev MT, Mattivi JT, Jørgensen MS, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020(5).
14. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022:1-16.
15. Ellison WD. Psychotherapy for Borderline Personality Disorder: Does the Type of Treatment Make a Difference? *Current Treatment Options in Psychiatry*. 2020;7(3):416-28.
16. Linehan M. *Cognitive-behavioral treatment of borderline personality disorder*: Guilford press; 1993.
17. Juul S, Jakobsen JC, Hestbaek E, Jørgensen CK, Olsen MH, Rishede M, et al. Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder: A Randomized Clinical Trial (MBT-RCT). *Psychotherapy and Psychosomatics*. 2023;92(5):329-39.
18. Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166(12):1355-64.
19. Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563-9.
20. Bateman, Fonagy P. *Mentalization-based treatment for personality disorders: a practical guide*: Oxford University Press; 2016.
21. Bateman A, Bales D, MBT. HJAQMf. A Quality Manual for MBT. 2020:Available at: <https://www.annafreud.org/media/14278/adherence-scale-august-2020.pdf> [Accessed January 10, 2023].
22. Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385(9969):735-43.
23. Juul S, McMain SF, Olsen MH, Chapman AL, Ribeiro JP, Storebø OJ, et al. Short-term versus long-term psychotherapy for borderline personality disorder. A protocol for an individual patient data meta-analysis. In peer-review at *BMJ Open*. 2025.
24. Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, 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;22(1):497.

25. McMain SF, Chapman AL, Kuo JR, Guimond T, Streiner DL, Dixon-Gordon KL, et al. The effectiveness of 6 versus 12-months of dialectical behaviour therapy for borderline personality disorder: the feasibility of a shorter treatment and evaluating responses (FASTER) trial protocol. *BMC Psychiatry*. 2018;18:1-16.

26. Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen JC, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. 2019;20(1):196.

27. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022;91(6):382-97.

28. Moeller J. A word on standardization in longitudinal studies: don't. *Frontiers in Psychology*. 2015;6:1389.

29. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC medical research methodology*. 2014;14(34).

30. Nielsen EE, Nørskov AK, Lange T, Thabane L, Wetterslev J, Beyersmann J, et al. Assessing assumptions for statistical analyses in randomised clinical trials. *BMJ Evidence-Based Medicine*. 2019;24(5):185-9.

31. Nørskov AK, Lange T, Nielsen EE, Gluud C, Winkel P, Beyersmann J, et al. Assessment of assumptions of statistical analysis methods in randomised clinical trials: the what and how. *BMJ Evid Based Med*. 2021;26(3):121-6.

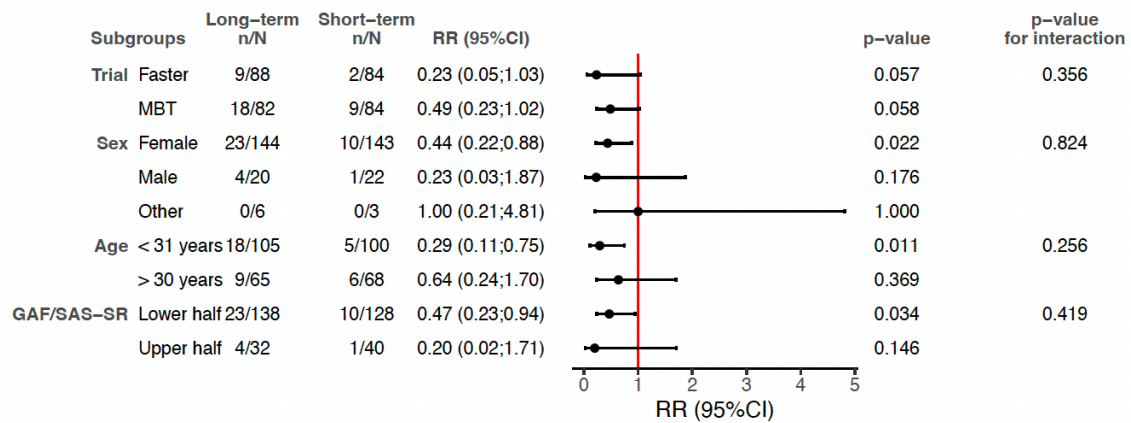
32. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.

33. Grey P, Grey A, Bolland MJ. Outcomes, interventions and funding in randomised research published in high-impact journals. *Trials*. 2018;19(1):592.

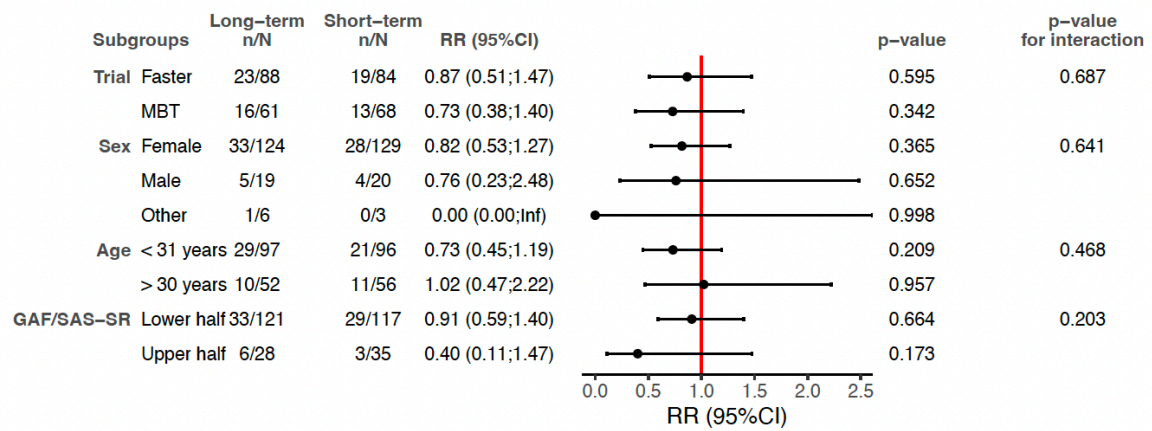
34. Juul S, Siddiqui F, Faltermeier P, Petersen JJ, Kamp CB, Jakobsen RH, et al. Less focus on symptom scales in psychiatric trials: it is time to ensure research equality between psychiatry and other medical specialities. *Lancet Reg Health – Eur*. 2024;43.

35. Juul S, Faltermeier P, Siddiqui F, Petersen JJ, Kamp CB, Jakobsen RH, et al. Challenges in the selection and measurement of outcomes in psychiatric trials. *BMJ Evidence-Based Medicine*. 2025.
36. Kramer U. Case formulation for personality disorders: Tailoring psychotherapy to the individual client: Academic Press; 2019.

Primary outcome



Secondary outcome



Supplementary material

“Short-term versus long-term psychotherapy for borderline personality disorder: An individual patient data pooled analysis of two randomized clinical trials”

Content

Eligibility criteria of the two trials	2
<i>FASTER trial.....</i>	<i>2</i>
<i>MBT trial.....</i>	<i>2</i>
Participant characteristics of the two trials	3
<i>FASTER trial.....</i>	<i>3</i>
<i>MBT trial.....</i>	<i>4</i>
Risk of bias assessments.....	6
Trial Sequential Analysis results	6
<i>Primary outcome – fixed effect</i>	<i>6</i>
<i>Primary outcome – random effects</i>	<i>7</i>
<i>Secondary outcome – fixed effect.....</i>	<i>7</i>
<i>Secondary outcome – random effects</i>	<i>8</i>
Deviations from the protocol	8
Statistical report	8
References:	9

Eligibility criteria of the two trials

FASTER trial

“Inclusion criteria were as follows: a Diagnostic and Statistical Manual Version IV (DSM-IV) diagnosis of BPD; age 18–60 years; at least two episodes of suicide attempts or nonsuicidal self-injury (NSSI) in the past 5 years, including one or more within the past 2 months; proficiency in English; and valid provincial health insurance in Ontario or British Columbia during the past year. Exclusion criteria included a DSM-IV diagnosis of bipolar I disorder, dementia, and/or a psychotic disorder; a serious physical health issue with anticipated hospitalization in the next year; IQ of 70 or less; at least 8 weeks of DBT in the last year; and plans to move out of the province during the trial.” [1]

MBT trial

“Eligible participants were adults (≥ 18 years) with subthreshold or diagnosed BPD according to the diagnostic criteria provided in the Diagnostic and Statistical Manual for Mental Disorders, 5th edition (DSM-5) [1] who provided written informed consent. Exclusion criteria were reduced intellectual functioning (IQ < 75), a full diagnosis of schizotypal personality disorder or antisocial personality disorder according to DSM-5, a comorbid mental health disorder requiring specialist treatment, current (past 2 months) substance dependence assessed with the Mini International Neuropsychiatric Interview (MINI), concurrent psychotherapeutic treatment, inability to understand Danish, and lack of informed consent.” [2]

Participant characteristics of the two trials

FASTER trial

Characteristic	Site		Length of treatment		Overall (N = 168)
	CAMH (n = 115)	SFU (n = 63)	12 months (n = 78)	6 months (n = 90)	
Mean (SD) age	28.13 [7.96]	28.58 [9.99]	27.65 [8.38]	28.81 [8.84]	28.27 [8.62]
Female	91 [79]	41 [77]	60 [77]	72 [80]	132 [79]
Marital status					
Married	16 [14]	10 [19]	16 [21]	10 [11]	26 [15]
Separated, divorced, widowed	9 [8]	5 [9]	4 [5]	10 [11]	14 [8]
Never married	90 [78]	38 [72]	58 [74]	70 [78]	128 [76]
Education					
High school or less	25 [22]	12 [23]	17 [22]	20 [22]	37 [22]
Some post-secondary	38 [33]	13 [25]	24 [31]	27 [30]	51 [30]
Post-secondary	52 [45]	28 [53]	37 [47]	43 [48]	80 [48]
Employed	42 [37]	24 [45]	24 [31]	42 [47]	66 [39]
Income ^a					
<15,000 USD	66 [57]	25 [47]	46 [59]	45 [50]	91 [54]
Between 15,000 USD and 29,000 USD	30 [26]	25 [47]	23 [29]	32 [36]	55 [33]
Between 30,000 USD and 49,000 USD	14 [12]	2 [4]	5 [6]	11 [12]	16 [10]
Mdn (IQR) lifetime suicide attempts	1 [5]	2 [5]	1 [4]	2 [6]	1 [5]
Mean (SD) GAF	49.28 [6.79]	51.49 [8.61]	49.56 [7.75]	50.33 [7.22]	49.98 [7.46]
Lifetime comorbid axis I disorders					
Major depressive disorder	95 [83]	41 [77]	62 [79]	74 [82]	136 [81]
Panic disorder	35 [30]	18 [34]	25 [32]	28 [31]	53 [32]
Post-traumatic stress disorder	57 [50]	29 [55]	41 [53]	45 [50]	86 [51]
Any anxiety disorder	96 [83]	47 [89]	64 [82]	79 [88]	143 [85]
Any substance use disorder	93 [81]	39 [74]	65 [83]	67 [74]	132 [79]
Any eating disorder ^b	61 [53]	19 [36]	38 [49]	42 [47]	80 [48]
Current comorbid axis I disorders					
Major depressive disorder	44 [38]	22 [42]	29 [37]	37 [41]	66 [39]
Panic disorder	27 [23]	16 [30]	20 [26]	23 [26]	43 [26]
Post-traumatic stress disorder	34 [30]	24 [45]	25 [32]	33 [37]	58 [35]
Any anxiety disorder	90 [78]	45 [85]	61 [78]	74 [82]	135 [80]
Any substance use disorder ^c	52 [45]	11 [21]	31 [40]	32 [36]	63 [38]
Any eating disorder	26 [23]	9 [17]	11 [14]	24 [27]	35 [21]
Axis II cluster A diagnosis ^d	1 [1]	12 [23]	6 [8]	7 [8]	13 [8]
Axis II cluster B diagnosis (excl. BPD) ^e	3 [3]	8 [15]	3 [4]	8 [9]	11 [7]
Axis II cluster C diagnosis ^{e,f}	27 [23]	22 [42]	16 [21]	33 [37]	49 [29]
Mean (SD) axis I current disorders	2.75 [1.50]	3.02 [1.92]	2.82 [1.64]	2.84 [1.65]	2.83 [1.64]
Mean (SD) axis I lifetime disorders	4.88 [2.22]	5.12 [2.60]	5.28 [2.33]	4.67 [2.32]	4.95 [2.34]
Mean (SD) axis II disorders (excl. BPD) ^f	0.28 [0.47]	1.00 [1.21]	0.38 [0.74]	0.61 [0.92]	0.51 [0.85]

Data reported are based on per protocol sample (N = 168). Values reported are n (%) unless otherwise noted. χ^2 and Fisher exact tests (and T tests for variables reporting a mean) revealed no significant differences between sites or conditions, except as noted. BPD, borderline personality disorder; CAMH, Center for Addiction and Mental Health, Toronto, ON, Canada; GAF, Global Assessment of Function; SFU, Simon Fraser University, Burnaby, BC, Canada. ^aThe proportion of participants with income of 15–29K was significantly higher at SFU ($p = 0.043$). There were no significant differences between conditions. ^bThe proportion of participants with any lifetime eating disorder was significantly higher at CAMH ($p = 0.045$). There were no significant differences between conditions. ^cThe proportion of participants currently diagnosed with any substance use disorder was significantly higher at CAMH ($p = 0.005$). There were no significant differences between conditions. ^dThe proportion of participants diagnosed with a comorbid cluster A, B, or C axis II disorder was significantly higher at SFU ($p < 0.001$, $p = 0.005$, $p = 0.027$, respectively). There were no significant between-condition differences for both comorbid cluster A and cluster B axis II disorders. ^eThe proportion of participants with any cluster C axis II disorder was significantly higher in the 6-month treatment condition. ^fParticipants at SFU were diagnosed with significantly more comorbid axis II disorders ($p < 0.001$). Participants assigned to the 6-month treatment were also diagnosed with significantly more comorbid axis II disorders ($p = 0.034$).

(Table published in the original trial report [1])

MBT trial

Characteristics	Short-term MBT (<i>n</i> = 84)	Long-term MBT (<i>n</i> = 82)
Sociodemographic characteristics		
Age, mean (SD), years	31.7 (10.1)	32.2 (10.1)
Female sex, <i>n</i> (%)	75 (89.1)	76 (92.7)
Civil status, <i>n</i> (%)		
Single	45 (53.6)	41 (50)
In a relationship (living together)	20 (23.8)	13 (15.9)
In a relationship (living separately)	11 (13.1)	19 (23.2)
Married	8 (9.5)	9 (10.9)
Living situation, <i>n</i> (%)		
With others	48 (57.1)	53 (64.6)
Alone	36 (42.9)	28 (34.2)
Other	0 (0)	1 (1.2)
Education level after high school, <i>n</i> (%)		
Short education	15 (17.9)	11 (13.4)
Medium education	20 (23.8)	17 (20.7)
Long education	6 (7.1)	5 (6.1)
No higher education	33 (39.3)	38 (46.3)
Other education	10 (11.9)	11 (13.4)
Employment status, <i>n</i> (%)		
On social welfare (/unemployed)	38 (45.2)	47 (57.3)
Currently studying	21 (25)	19 (23.2)
Employed	18 (21.4)	14 (17.1)
Unemployed/homemaker	2 (2.4)	0 (0)
Retired	1 (1.2)	0 (0)
Self-employed	1 (1.2)	1 (1.2)
Other	3 (3.6)	1 (1.2)
Clinical characteristics		
Psychiatric comorbidity (MINI), <i>n</i> (%)		
Anxiety disorders	42 (50)	53 (64.6)

Characteristics	Short-term MBT (<i>n</i> = 84)	Long-term MBT (<i>n</i> = 82)
Major depressive disorder	30 (35.7)	28 (34.1)
Post-traumatic stress disorder	17 (20.2)	24 (29.3)
Eating disorders	12 (14.3)	12 (14.6)
Obsessive compulsive disorder	10 (11.9)	7 (8.5)
Proportion of participants with subthreshold BPD, <i>n</i> (%)	20 (23.8)	13 (15.9)
No. of BDP criteria in total (SCID-5-PD), mean (SD)	6.2 (1.4)	6.6 (1.1)
No. of PD criteria in total (SCID-5-PD), mean (SD)	18.1 (4.9)	17.9 (5.5)
Personality disorder comorbidity (SCID-5-PD), <i>n</i> (%)		
Avoidant personality disorder	25 (29.8)	21 (25.6)
Obsessive-compulsive personality disorder	14 (16.7)	27 (32.9)
Paranoid personality disorder	15 (17.9)	20 (24.4)
Dependent personality disorder	8 (9.5)	7 (8.5)
Proportion of participants with one or more suicide-attempts the past 8 months, <i>n</i> (%)	5 (5.9)	6 (7.3)
Proportion of participants with one or more acts of severe self-harm the past 8 months, <i>n</i> (%)	32 (38.1)	33 (40.2)
Proportion of participants on antidepressants, <i>n</i> (%)	33 (39.3)	35 (42.7)
Proportion of participants on antipsychotics, <i>n</i> (%)	10 (11.9)	11 (13.3)
Proportion of participants on other psychoactive drugs, <i>n</i> (%)	21 (25.0)	19 (23.2)
Mean (SD) number of days from randomization to assessment time point		
8 months post-randomization	213 (12.4)	219 (13.8)
16 months post-randomization	389 (15.2)	354 (16.4)

(Table published in the original trial report [2])

Risk of bias assessments

Both trials were assessed as at low risk of bias on all bias risk domains except domain two caused by lack of blinding of participants and clinicians.

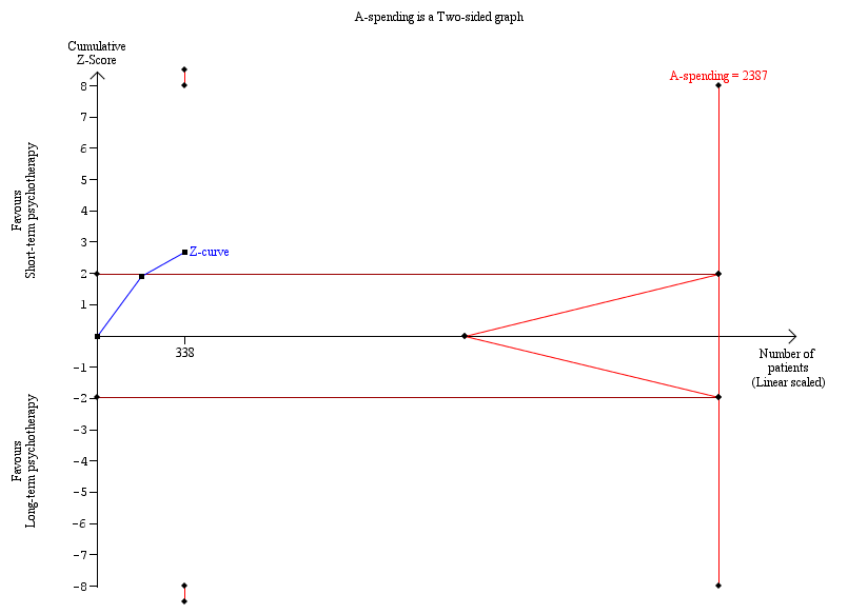
		Risk of bias domains				
		D1	D2	D3	D4	D5
Study	FASTER					
	MBT-RCT					

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.

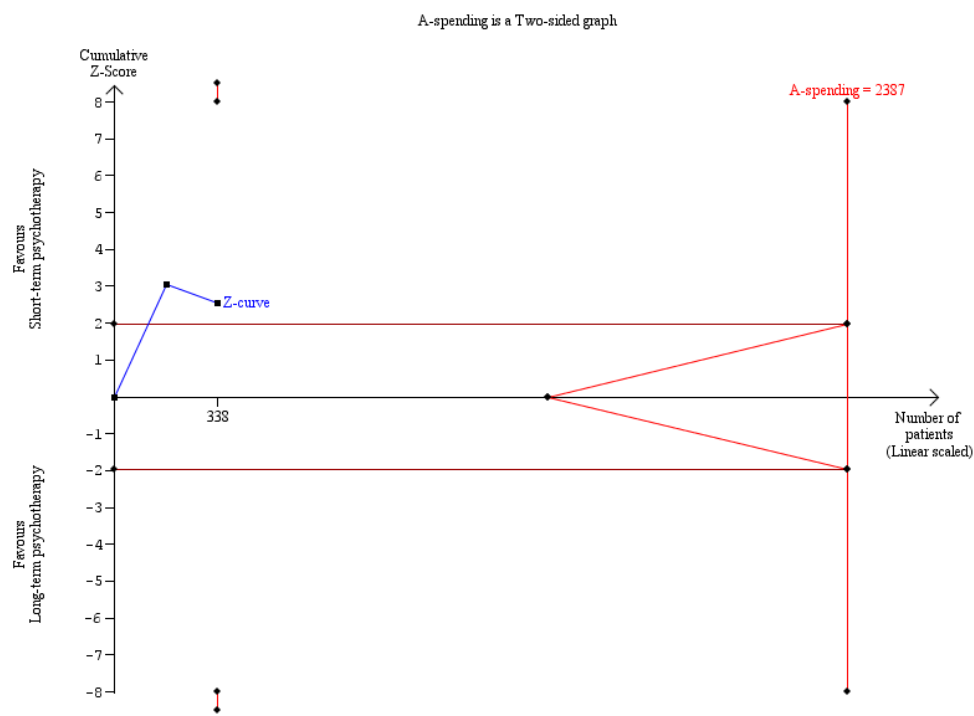
Judgement
 Some concerns
 Low

Trial Sequential Analysis results

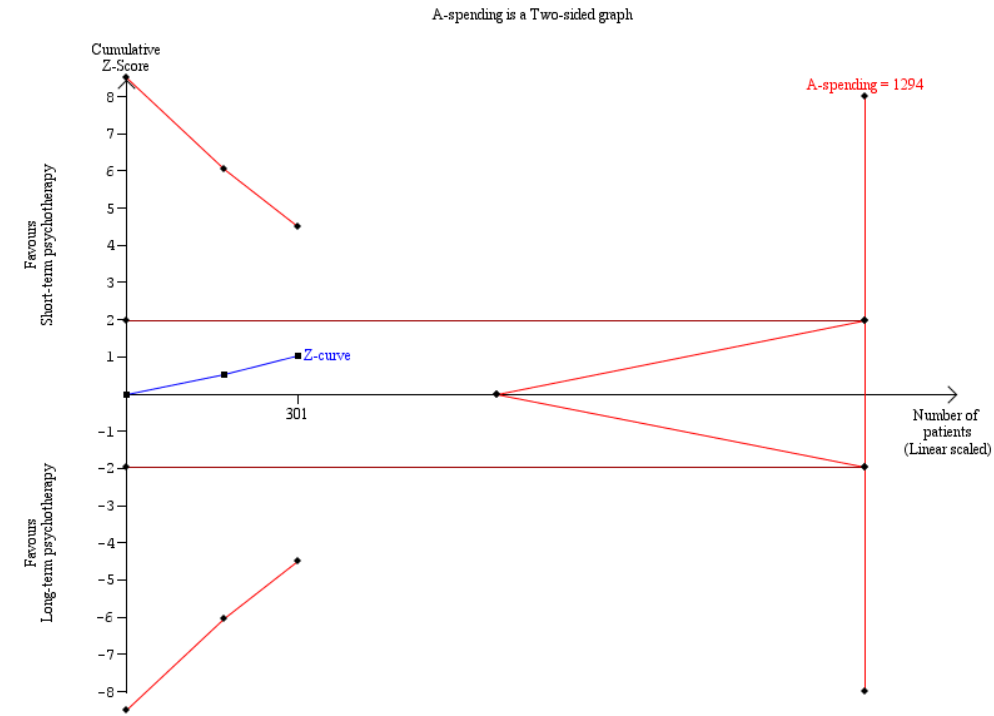
Primary outcome – fixed effect



Primary outcome – random effects

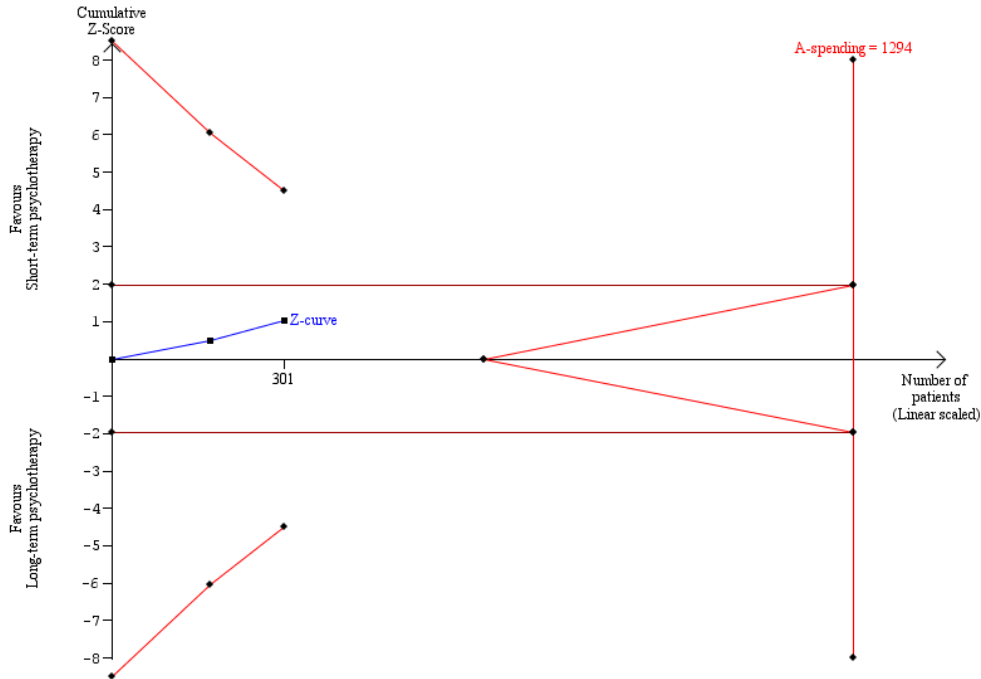


Secondary outcome – fixed effect



Secondary outcome – random effects

A-spending is a Two-sided graph



Deviations from the protocol

The protocol stated that we would use meta-analytical statistics using standardized mean differences for continuous outcomes; however, standardization for longitudinal data is not preferred and *percent of maximum possible (POMP)* is recommended instead [3]. Hence, we analyzed continuous outcomes using mixed effects linear regression instead [3].

Statistical report

The statistical report can be found below.

References:

1. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The effectiveness of 6 versus 12 months of dialectical behavior therapy for borderline personality disorder: a noninferiority randomized clinical trial. *Psychother Psychosom*. 2022;91(6):382-97.
2. Juul S, Jakobsen JC, Hestbaek E, Jørgensen CK, Olsen MH, Rishede M, et al. Short-Term versus Long-Term Mentalization-Based Therapy for Borderline Personality Disorder: A Randomized Clinical Trial (MBT-RCT). *Psychotherapy and Psychosomatics*. 2023;92(5):329-39.
3. Moeller J. A word on standardization in longitudinal studies: don't. *Frontiers in Psychology*. 2015;6:1389.

Statistical report for an individual patient data meta-analysis on Short-term versus long-term psychotherapy for borderline personality disorder

Markus Harboe Olsen

2025-03-27

Table of contents

Primary outcome	2
Secondary outcome	2
Exploratory outcomes	3
ZAN-BPD or BSL-23	3
SF-36 MCS or EQ-5D-5L	4
GAF or SAS-SR	5
GSI	6
Suicide or suicide attempts	6
Psychiatric hospitalization	6
Subgroup analysis	7
Primary outcome	7
Secondary outcome	7
Bayes factor	8

Primary outcome

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects logistic regression <i>RR</i> (95% <i>CI</i>)	<i>p</i>
Primary outcome (short)	27 (13.4%)	11 (5.4%)	0.41 (0.21;0.79)	0.008
missing	32 (15.8%)	36 (17.6%)		
Primary outcome (long)	32 (15.8%)	24 (11.8%)	0.77 (0.48;1.25)	0.296
missing	42 (20.8%)	49 (24.0%)		

Without missingness

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects logistic regression <i>RR</i> (95% <i>CI</i>)	<i>p</i>
Primary outcome (short)	27 (15.9%)	11 (6.5%)	0.41 (0.21;0.79)	0.008
Primary outcome (long)	32 (20.0%)	24 (15.5%)	0.77 (0.48;1.25)	0.296

Secondary outcome

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects logistic regression <i>RR</i> (95% <i>CI</i>)	<i>p</i>
Deliberate self-ham (short)	39 (19.3%)	32 (15.7%)	0.80 (0.53;1.21)	0.297
missing	53 (26.2%)	52 (25.5%)		
Deliberate self-ham (long)	35 (17.3%)	31 (15.2%)	0.91 (0.60;1.38)	0.644
missing	67 (33.2%)	70 (34.3%)		

Without missingness

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects logistic regression <i>RR</i> (95% <i>CI</i>)	<i>p</i>
Deliberate self-ham (short)	39 (26.2%)	32 (21.1%)	0.80 (0.53;1.21)	0.297
Deliberate self-ham (long)	35 (25.9%)	31 (23.1%)	0.91 (0.60;1.38)	0.644

Exploratory outcomes

ZAN-BPD or BSL-23

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204
ZAN-BPD (baseline) - mean (95%CI)	13.80 (12.98;14.63)	13.90 (13.10;14.71)
missing	120 (59.4%)	120 (58.8%)
ZAN-BPD (short) - mean (95%CI)	12.28 (10.78;13.78)	11.34 (9.86;12.83)
missing	141 (69.8%)	134 (65.7%)
ZAN-BPD (long) - mean (95%CI)	10.23 (8.77;11.69)	10.87 (9.22;12.53)
missing	145 (71.8%)	141 (69.1%)
BSL-23 (baseline) - mean (95%CI)	2.32 (2.18;2.47)	2.31 (2.18;2.44)
missing	82 (40.6%)	84 (41.2%)
BSL-23 (short) - mean (95%CI)	1.60 (1.40;1.81)	1.43 (1.24;1.62)
missing	115 (56.9%)	125 (61.3%)
BSL-23 (long) - mean (95%CI)	1.57 (1.34;1.81)	1.62 (1.38;1.86)
missing	125 (61.9%)	135 (66.2%)

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects linear regression <i>estimate</i> (95%CI)	<i>p</i>
ZAN-BPD or BSL-23 (short) - mean (95%CI)	37.63 (34.17;41.08)	33.78 (30.59;36.97)	-2.93 (-7.24;1.38)	0.184
missing	54 (26.7%)	55 (27%)		
ZAN-BPD or BSL-23 (long) - mean (95%CI)	34.65 (30.80;38.51)	35.59 (31.72;39.47)	1.91 (-3.07;6.90)	0.454
missing	68 (33.7%)	72 (35.3%)		

SF-36 MCS or EQ-5D-5L

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204
SF-36 MCS (baseline) - mean (95%CI)	49.21 (47.00;51.43)	50.26 (48.01;52.51)
missing	120 (59.4%)	120 (58.8%)
SF-36 MCS (short) - mean (95%CI)	52.91 (50.18;55.63)	53.87 (51.37;56.38)
missing	146 (72.3%)	141 (69.1%)
SF-36 MCS (long) - mean (95%CI)	54.44 (51.41;57.46)	49.05 (46.62;51.48)
missing	155 (76.7%)	148 (72.5%)
EQ-5D-5L (baseline) - mean (95%CI)	0.67 (0.63;0.71)	0.65 (0.61;0.68)
missing	82 (40.6%)	84 (41.2%)
EQ-5D-5L (short) - mean (95%CI)	0.73 (0.68;0.77)	0.74 (0.70;0.79)
missing	116 (57.4%)	125 (61.3%)
EQ-5D-5L (long) - mean (95%CI)	0.71 (0.66;0.76)	0.71 (0.66;0.77)
missing	126 (62.4%)	137 (67.2%)

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects linear regression <i>estimate</i> (95%CI)	<i>p</i>
SF-36 MCS or EQ-5D-5L (short) - mean (95%CI)	71.09 (67.93;74.24)	70.60 (67.50;73.71)	0.45 (-2.21;3.10)	0.740
missing	60 (29.7%)	62 (30.4%)		
SF-36 MCS or EQ-5D-5L (long) - mean (95%CI)	71.43 (68.11;74.74)	66.99 (63.34;70.64)	-2.72 (-5.82;0.34)	0.084
missing	79 (39.1%)	81 (39.7%)		

GAF or SAS-SR

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204		
GAF (baseline) - mean (95%CI)	46.28 (44.88;47.68)	46.55 (44.78;48.31)		
missing	120 (59.4%)	120 (58.8%)		
GAF (short) - mean (95%CI)	49.74 (46.73;52.74)	52.23 (48.65;55.81)		
missing	141 (69.8%)	134 (65.7%)		
GAF (long) - mean (95%CI)	50.70 (47.10;54.29)	48.40 (45.58;51.22)		
missing	146 (72.3%)	141 (69.1%)		
SAS-SR (baseline) - mean (95%CI)	2.84 (2.75;2.94)	2.78 (2.69;2.86)		
missing	82 (40.6%)	84 (41.2%)		
SAS-SR (short) - mean (95%CI)	2.45 (2.34;2.57)	2.48 (2.35;2.62)		
missing	116 (57.4%)	125 (61.3%)		
SAS-SR (long) - mean (95%CI)	2.44 (2.30;2.58)	2.50 (2.35;2.65)		
missing	126 (62.4%)	137 (67.2%)		

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects linear regression <i>estimate</i> (95%CI)	<i>p</i>
GAF or SAS-SR (short) - mean (95%CI)	47.23 (45.26;49.21)	48.28 (45.89;50.67)	-0.15 (-2.90;2.60)	0.916
missing	55 (27.2%)	55 (27%)		
GAF or SAS-SR (long) - mean (95%CI)	47.83 (45.49;50.17)	46.36 (44.18;48.54)	-2.33 (-5.24;0.58)	0.118
missing	70 (34.7%)	74 (36.3%)		

GSI

	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects linear regression <i>estimate (95%CI)</i>	<i>p</i>
GSI (baseline) - mean (95%CI)	1.76 (1.67;1.84)	1.70 (1.61;1.78)		
GSI (short) - mean (95%CI)	1.36 (1.23;1.48)	1.26 (1.14;1.38)	-0.01 (-0.16;0.13)	0.854
missing	59 (29.2%)	63 (30.9%)		
GSI (long) - mean (95%CI)	1.27 (1.13;1.42)	1.32 (1.17;1.46)	0.13 (-0.04;0.30)	0.141
missing	78 (38.6%)	79 (38.7%)		

Suicide or suicide attempts

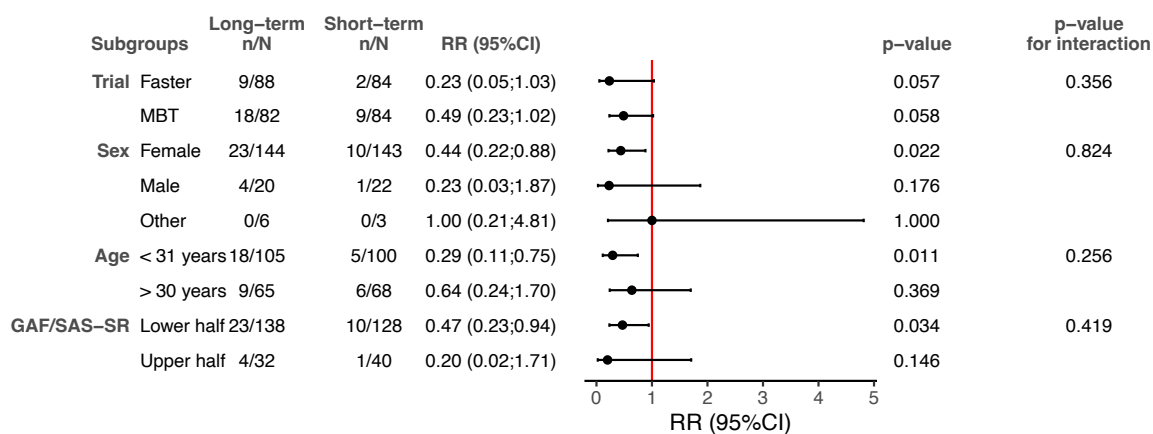
	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects logistic regression <i>RR (95%CI)</i>	<i>p</i>
Suicide or suicide attempt (short)	9 (4.5%)	6 (2.9%)	0.66 (0.25;1.78)	0.415
missing	32 (15.8%)	36 (17.6%)		
Suicide or suicide attempt (long)	13 (6.4%)	10 (4.9%)	0.79 (0.36;1.76)	0.569
missing	42 (20.8%)	49 (24.0%)		

Psychiatric hospitalization

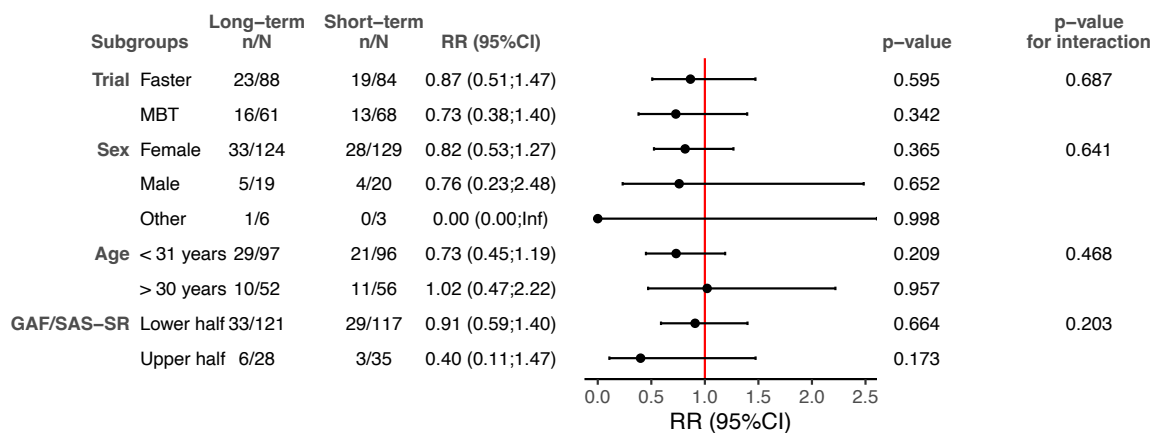
	Long-term <i>n</i> = 202	Short-term <i>n</i> = 204	Mixed effects logistic regression <i>RR (95%CI)</i>	<i>p</i>
Psychiatric hospitalisation (short)	26 (12.9%)	11 (5.4%)	0.43 (0.22;0.82)	0.011
missing	32 (15.8%)	36 (17.6%)		
Psychiatric hospitalisation (long)	28 (13.9%)	23 (11.3%)	0.83 (0.50;1.36)	0.457
missing	42 (20.8%)	48 (23.5%)		

Subgroup analysis

Primary outcome



Secondary outcome



Bayes factor

Primary outcome (short) - Bayes factor: 0.06473219

Deliberate self-ham (short) - Bayes factor: 1.678633

Paper 11

Review

Pharmacological interventions for co-occurring psychopathology in people with borderline personality disorder: secondary analysis of the Cochrane systematic review with meta-analyses

Johanne Pereira Ribeiro, Sophie Juul, Mickey T. Kongerslev, Mie Sedoc Jørgensen, Birgit A Völlm, Henriette Edemann-Callesen, Christian Sales, Julie P. Schaug, Klaus Lieb, Erik Simonsen, Jutta M. Stoffers-Winterling[†] and Ole Jakob Storebø[†]

Background

Medications are commonly used to treat co-occurring psychopathology in persons with borderline personality disorder (BPD)

Aims

To systematically review and integrate the evidence of medications for treatment of co-occurring psychopathology in people with BPD, and explore the role of comorbidities.

Method

Building on the current Cochrane review of medications in BPD, an update literature search was done in March 2024. We followed the methods of this Cochrane review, but scrutinised all identified placebo-controlled trials *post hoc* for reporting of non BPD-specific ('co-occurring') psychopathology, and explored treatment effects in subgroups of samples with and without defined co-occurring disorders. GRADE ratings were done to assess the evidence certainty.

Results

Twenty-two trials were available for quantitative analyses. For antipsychotics, we found very-low-certainty evidence (VLCE) of an effect on depressive symptoms (standardised mean difference (SMD) -0.22 , $P = 0.04$), and low-certainty evidence (LCE) of an effect on psychotic–dissociative symptoms (SMD -0.28 , $P = 0.007$). There was evidence of effects of anticonvulsants on depressive (SMD -0.44 , $P = 0.02$; LCE) and anxious

symptoms (SMD -1.11 , $P < 0.00001$; VLCE). For antidepressants, no significant findings were observed (VLCE). Exploratory subgroup analyses indicated a greater effect of antipsychotics in samples including participants with co-occurring substance use disorders on psychotic–dissociative symptoms ($P = 0.001$).

Conclusions

Our findings, based on VLCE and LCE only, do not support the use of pharmacological interventions in people with BPD to target co-occurring psychopathology. Overall, the current evidence does not support differential treatment effects in persons with versus without defined comorbidities. Medications should be used cautiously to target co-occurring psychopathology.

Keywords

Comorbidity; pharmaceutical drug trial; clinical trials; mental health services; personality disorders.

Copyright and usage

© The Author(s), 2024. Published by Cambridge University Press on behalf of Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Borderline personality disorder (BPD) commonly co-occurs with other mental health disorders. Large epidemiological studies report prevalence rates of around 80% for confirmed diagnoses of mood disorders and substance use disorders (SUDs).¹ Other disorders are also highly prevalent, such as post-traumatic stress disorder (PTSD), eating disorders and attention-deficit hyperactivity disorder.^{2,3} However, methodologically high-standing clinical trials, such as randomised controlled trials (RCTs), often use highly selected samples that explicitly exclude these co-occurring disorders.^{4,5} Although psychotherapy research has advanced in terms of developing and evaluating adapted treatments designed for patients who have defined comorbidities such as PTSD, SUDs or eating disorders, to our knowledge, no placebo-controlled RCTs have tested the effects of major pharmacological intervention classes (i.e. antipsychotics, antidepressants or anticonvulsants) used in BPD samples to treat co-occurring psychopathology.

The disconnection between clinical guidelines and practice

Despite international guidelines consistently advising against the use of pharmacological interventions for continued BPD treatment,^{6–9} impulsivity, affective instability and dissociative symptoms are still considered targets of psychopharmacological treatments in clinical practice.^{10,11} Consequently, medication rates are high in clinical settings, and although a decrease in overall medication use can be observed, co-occurring psychopathology remain the main factor for the prescription of pharmacological interventions, as well as for polypharmacy.¹⁰ Although co-existing conditions such as major depressive disorder (MDD) may require pharmacological intervention, and guidelines advise to apply to the clinical guidelines for these co-occurring disorders,^{7,8} the evidence regarding these interventions is very limited. To our knowledge, there are no placebo-controlled RCTs testing the effects of pharmacological interventions in samples of people with a diagnosis of BPD plus another confirmed co-occurring psychopathology, despite the high rates of co-occurring psychopathology in this population. Instead, the generalisability of clinical trials in the field of BPD has been questioned by a study that scrutinised participants of a large epidemiological study,

[†] Joint last authors.

regarding their eligibility to enter pharmacological clinical trials, and found that as many as 76% of community-dwelling people with BPD would have been excluded from such trials.⁴

Recent findings: insights from a Cochrane systematic review

We recently published an update of a Cochrane systematic review on pharmacological interventions for people with BPD.¹² It included 46 RCTs that comprised 2769 participants with a confirmed diagnosis of BPD. No restrictions were applied concerning co-occurring psychopathology. The main intervention categories were antipsychotics, antidepressants and anticonvulsants. Primary outcomes comprised BPD symptom severity, self-harm, suicide-related outcomes and psychosocial functioning. Secondary outcomes comprised specific BPD symptoms as defined by the individual nine DSM-5 criteria,¹³ attrition and adverse events.

Following up on the Cochrane review, we assessed the following clinical questions in this publication: (a) What are the effects of pharmacological interventions compared with placebo in people with BPD on common co-occurring psychopathology not specific to BPD (e.g. depressive, anxious, dissociative, substance use or eating disorder symptoms)? and (b) Is there any evidence of differential outcomes depending on the type of co-occurring psychopathology in samples including participants with defined co-occurring psychopathology (e.g. MDD, SUD, PTSD) versus samples excluding them?

Method

This systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (see PRISMA checklist, Supplementary Appendix 1 available at <https://doi.org/10.1192/bjp.2024.172>) and is associated with our Cochrane systematic review 'Pharmacological interventions for people with borderline personality disorder'.¹² The protocol was preregistered in PROSPERO (identifier CRD42018091044) and published in the Cochrane Library in 2018.¹⁴

This paper has two main aims. The first aim was to explore the effects of medications on co-occurring psychopathology in persons with BPD. As co-occurring psychopathology, we defined psychopathology outcomes that are not defined as BPD criteria by the DSM-5.¹⁵ In the 2022 Cochrane review, such outcomes were limited to depressive and psychotic symptoms. Therefore, all included placebo-controlled trials were re-assessed for any additional outcomes regarding symptoms of co-occurring disorders at the end of treatment (such as symptoms of PTSD, anxiety, psychosis, depression, substance misuse, eating disorder, etc.), and analysed *ad hoc*.

The second aim was to explore if there is reason to assume that medication effects on co-occurring psychopathology might differ depending on the co-occurrence of manifest psychiatric comorbidities. Therefore, the placebo-controlled effects of medications on co-occurring psychopathology in selected versus non-selected subgroups were compared in exploratory subgroup analyses *post hoc*, using the individual trial eligibility criteria for specific co-occurring psychiatric disorders (i.e. whether trials excluded specific co-occurring psychiatric disorders or not).

We applied these *post hoc* exploratory subgroup analyses to all outcomes of co-occurring psychopathology. As co-occurring psychopathology outcomes, we used those outcomes that were predefined in the Cochrane 2022 review (i.e. depressive and dissociative symptoms), as well as those outcomes of co-occurring psychopathology that we identified *ad hoc* in the primary studies. We deemed a minimum of three effect sizes per comparison and outcome necessary for combining them meta-analytically.

Search strategy

The full methodology of this systematic review is predefined and described elsewhere.¹² In short, we included RCTs that investigated any pharmacological interventions compared with placebo for people with a BPD diagnosis according to any available version of the DSM and ICD. We included trials if at least 70% of participants had a BPD diagnosis. If a smaller proportion of trial participants had a BPD diagnosis, we contacted the study authors and requested data for the BPD subsample. Trials with or without co-occurring psychopathology were included.

For the Cochrane review, we searched 21 databases and trial registries until 21 February 2022. This search was updated on 6 March 2024 in Medline, Cochrane Central, EMBASE, PsycINFO, CINAHL, Web of Science, ERIC, LILACS, ProQuest Dissertations and Sociological Abstracts. In addition, we traced all studies categorised as 'ongoing' in the full review,¹² for meantime end-point data publications. Two independent review authors conducted title and abstract screening and full-text screening. A third review author settled any discrepancies. We additionally traced cross-references from relevant literature and contacted trial authors and sponsors regarding any unpublished data.

Data extraction and critical assessment of included trials and evidence summaries

Data extraction was carried out by two independent review authors (J.P.R. and S.J.), using a predefined data extraction template for consistency. Data extraction was done starting 1 June 2022 for the full Cochrane review, and starting 15 April 2024 for the update search. Additionally, we specifically searched for any end-of-treatment data on co-occurring psychopathology. In each included trial, we extracted all available information on inclusion and exclusion criteria, as well as the baseline data on co-occurring psychopathology.

We assessed the risk of bias in individual trials by using the Cochrane Collaboration Risk of Bias tool (RoB). Since assessments took place in the context of the full Cochrane review, which was started before publication of the current Cochrane RoB-2 tool, all assessments are based on the first version of the RoB tool.¹⁶ All risk-of-bias assessments were conducted independently by two authors, and discrepancies were settled by a third review author.

In accordance with Cochrane standards, the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) tool^{17–20} was used to assess the quality, i.e. certainty, of the statistically integrated evidence. Using GRADE, the body of evidence for each comparison and outcome (i.e. the pooled effect estimates) is categorised as high, moderate, low or very low certainty. We considered the following factors as suggested by GRADE: risk of bias in primary studies (based on risk-of-bias ratings previously conducted by use of the Cochrane RoB tool), inconsistency, indirectness and imprecision. Moreover, funnel plots were drawn to examine publication bias if at least ten effect estimates were available.¹⁶ Two authors (J.P.R. and J.M.S.-W.) independently conducted all GRADE assessments and solved any discrepancies by discussion.^{17–20}

Data analysis

We included data on intention-to-treat samples whenever possible. If diverse measurement scales were used to assess the same outcome in primary studies, we calculated standardised mean differences (SMDs) and corresponding 95% confidence intervals. Estimates for continuous outcomes were generated with inverse variance. SMD summary estimates were generated with random-effects models, as we expected substantial heterogeneity among trials. Heterogeneity was assessed by use of the I^2 score.²¹ As inconsistency depends also on additional factors such as the magnitude and direction of

effects, or χ^2 *P*-values, we considered both the I^2 statistic and visual inspection of the forest plot and exclusion of outliers to assess heterogeneity, instead of solely relying on the I^2 statistic. According to the Cochrane handbook, heterogeneity was therefore considered moderate when between 30 and 60%, substantial when between 50 and 90%, and considerable when between 75 and 100%.²² For predefined methods not used in this review, see Supplementary Appendix 2.

We conducted sensitivity analyses in terms of excluding outliers, defined as individual study estimates for which the 95% confidence interval boundaries did not overlap with the 95% confidence interval of the pooled effect estimates.²³ As in the main review, we also performed sensitivity analyses in terms of excluding (primary analyses) versus including (secondary analyses) data from cross-over trials for which no separate data for the period before first cross-over were available, using end-of-period data.^{24,25}

To examine the role of comorbid psychopathology in pharmacotherapy effects, we conducted subgroup analyses to compare samples excluding versus not excluding defined psychiatric comorbidities. Such subgroup analyses were only done if, for each subgroup, a minimum of three primary studies was available.

Analyses along with forest plots and risk-of-bias tables were generated with RevMan5 version 5.4.1 for Windows (The Cochrane Collaboration, London, UK; see <https://training.cochrane.org/online-learning/core-software/revman>).²⁶

Results

General description of all included trials

Of the 46 trials included in the Cochrane review, 38 had a placebo group and were considered for quantitative analyses here. In

March 2024, an update search plus tracking of trials included as ongoing in the 2022 Cochrane review yielded three more eligible, placebo-controlled trials with published end-point data.^{27–29} Therefore, 41 placebo-controlled trials were considered for inclusion into quantitative analyses (Fig. 1).

Nineteen of the 41 trials were not included in the quantitative analyses for the following reasons: eight trials provided no outcomes relevant to this review,^{27,30–36} another eight trials^{29,36–40} were not included because of a lack of a sufficient number of trials testing the same kind of substance (minimum requirement: three poolable effect estimates per analysis), and three trials reported data in a way that was not suitable for effect size calculations.^{41–43} Study characteristics of these trials are detailed in Supplementary Appendix 3. After exclusion of these trials, 22 trials were included in the quantitative analyses (Fig. 1, Supplementary Appendix 4).

Characteristics of the included trials are detailed in Table 1. Altogether, the 22 trials comprised 1690 participants with BPD. The mean age ranged from 21.7⁶³ to 40.7 years.⁶⁴ Five trials included females only;^{46,49,58,60,62} all remaining trials included both genders, but with predominantly females.

The trials were published between 1986 and 2024. Two were cross-over trials,^{46,55} and the rest had a parallel design. The smallest trial consisted of 16 participants,⁴⁶ and the largest provided data on 314 participants.⁶⁵ Four trials included more than 100 participants.^{56,65–67} Twelve trials were conducted in the USA, and eight were carried out in Europe. The remaining two multicentre trials included study sites mainly in North America and Europe.^{65,66}

As within the 2022 Cochrane review, the placebo-controlled effect estimates available for meta-analyses were categorised into the three broad therapeutic classes: antipsychotics (including trials on aripiprazole, brexpiprazole, clozapine, haloperidol, olanzapine,

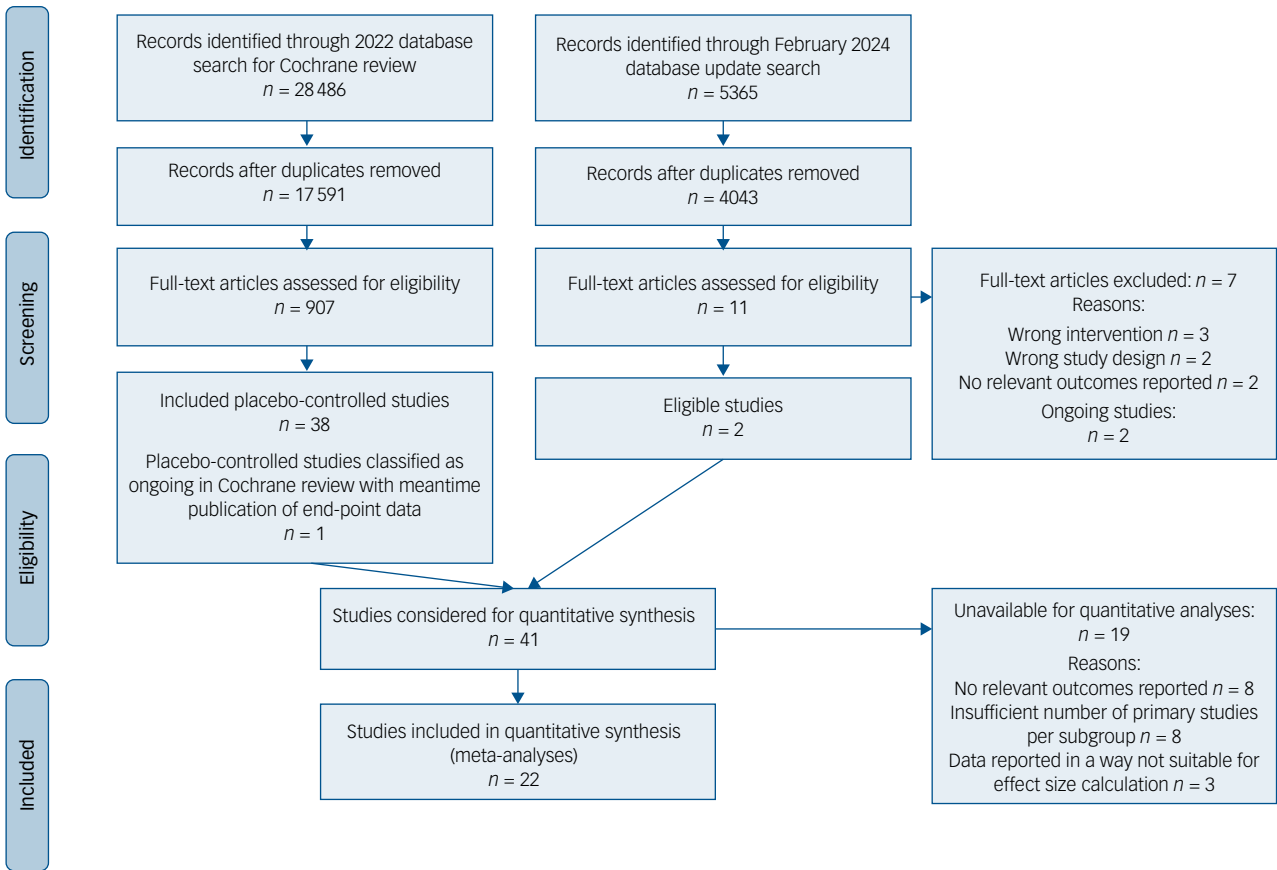


Fig. 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart of study selection.

Table 1 Characteristics of trials available for quantitative analyses (22 randomised controlled trials, reporting on 26 placebo-controlled comparisons)

Study	Country	Gender (% female)	Age (years), mean	Sample population comorbidity	Concomitant medication allowed	Setting	Design	Duration	Sample size	Intervention	Comparator	Outcomes relevant to this review
Antipsychotics versus placebo Black et al ⁴⁴	USA	70.53	29.5	Mood disorders, anxiety disorders, substance misuse and other Axis 1 disorders	Benzodiazepines and anticholinergics were allowed	Out-patient	Parallel	8 weeks	95	Quetiapine	Placebo	Depressive symptoms
Bogenschutz and Nurnberg ⁴⁵	USA	62.50	32.6	Mean of 2.9 SCID-II personality disorders (including BPD) and a mean of 2.2 Axis 1 diagnoses	Medication for stable, chronic medical conditions such as hypertension	Out-patient	Parallel	12 weeks	40	Olanzapine (2.5–20 mg/d, most 5–10 mg/d, mean at end-point 6.9 mg/d (s.d. 3.2))	Placebo	Anxiety symptoms Psychotic symptoms
Cowdry and Gardner ⁴⁶	USA	100	31.6	Other Axis 2 personality disorders (dependent, avoidant, histrionic, schizotypal, narcissistic and paranoid), and Axis 1 disorders (atypical bipolar disorder, major depressive episode with and without melancholia, dysthymia)	Unclear, no information	Out-patient	Cross-over, multi-arm	6 weeks	16 ^c	Trifluoperazine hydrochloride	Placebo	Depressive symptoms Anxiety symptoms
Crawford et al ⁴⁸	UK	76%	30.0	High rates of personality disorder comorbidity (90% were rated as having a moderate or severe personality disorder (SASPD criteria))	No restrictions on the use of other treatments	In-patient (forensic setting)	Parallel	12 weeks	29	Clonidine	Placebo	Dissociation/ psychotic-like symptoms
Goldberg et al ⁴⁷	USA	58	32	Schizotypal personality disorder ^a	Unclear, no information	Out-patient	Parallel	12 weeks	40	Thiothixene	Placebo	Depressive symptoms Anxiety symptoms Psychotic symptoms
Grant et al ⁴⁸	USA	53.6	39.7	Anxiety disorder, mood disorder, eating disorder, ADHD	Allowed as prescribed (antidepressants, mood stabilisers and stimulants)	Out-patient	Parallel	12 weeks	80	Brexiprazole	Placebo	Depressive symptoms Anxiety symptoms
Linehan et al ⁴⁹	USA	100	36.8	Unclear, no information	Unclear, no information	Out-patient	Parallel	24 weeks	24	Olanzapine (2.5–15 mg/d, mean final dose 4.46, s.d. 1.16) plus DBT	Placebo plus DBT	Depressive symptoms
Nickel et al ⁵⁰	Germany	82.69	21.65	Depressive disorders, anxiety disorders, obsessive-compulsive disorders and somatoform disorders	Not allowed	Out-patient	Parallel	8 weeks	52	Aripiprazole	Placebo	Depressive symptoms Anxiety symptoms Psychotic symptoms
Pascual et al ⁵¹	Spain	81.67	29.2	Unclear, there is no information on population comorbidity. However, many psychiatric disorders were excluded	It was allowed to continue with benzodiazepine (maximum of 40 mg/d), antidepressants and mood stabilisers if initiated before inclusion; doses could not be modified	Out-patient	Parallel	12 weeks	60	Ziprasidone	Placebo	Depressive symptoms Anxiety symptoms Psychotic symptoms

(Continued)

Table 1 (Continued)

Study	Country	Gender (% female)	Age (years), mean	Sample population comorbidity	Concomitant medication allowed	Setting	Design	Duration	Sample size	Intervention	Comparator	Outcomes relevant to this review
Schulz et al ⁶⁵	Belgium, France, Germany, Norway, Portugal, Spain, Sweden, UK and USA	71.02	31.81	No apparent comorbidity (based on exclusion criteria)	No medications with primarily CNS activity (except for protocol-specified benzodiazepines and hypnotics)	Out-patient	Parallel	12 weeks	314	Olanzapine (2.5–20 mg/d, mean final dose 7.09 mg/d, s.d. 5.11)	Placebo	Depressive symptoms Psychotic symptoms
Soler et al ⁵²	Spain	86.57	40.74	No apparent comorbidity, (e.g. patients with unstable comorbid Axis 1 disorders were excluded)	Treatment with benzodiazepines, antidepressants and mood stabilisers could be continued, but doses could not be modified during the trial	Out-patient	Parallel	12 weeks	60	Olanzapine plus DBT	Placebo plus DBT	Depressive symptoms Anxiety symptoms
Soloff et al ⁵³	USA	75.56	25.1	Comorbid schizotypal personality disorder ^a	Piperiden hydrochloride (2 mg)	In-patient	Parallel, multi-arm	5 weeks	90	Haloperidol (amitriptyline)	Placebo	Anxiety Depression Dissociation/psychosis
Cornelius et al ⁵⁴	USA	75.93	26.7	Major depressive disorder	Not allowed	In- and out-patients	Parallel, multi-arm	5 weeks	108	Haloperidol (phenelzine sulphate)	Placebo	Depression Anxiety Dissociation/psychosis
Zanarini et al ⁶⁶	Argentina, Chile, Italy, Peru, Poland, Romania, Turkey, USA and Venezuela	73.61	32.98	Mood disorders, substance use disorders, anxiety disorders, eating disorders and Axis 2 disorders besides BPD	Benzodiazepines, hypnotics, lorazepam and episodic use of anticholinergics, benzotropine mesylate, piperiden or trihexyphenidyl allowed during trial at specific doses; however, prophylactic use of anticholinergics was not allowed	Out-patient	Parallel	12 weeks	301 ^b	Olanzapine (5–10 mg/d, mean dose 6.66 mg/d)	Placebo	Depression Dissociation/psychosis
Ziegenhorn et al ^{55c}	Germany	94.44	32	88% had comorbid disorders: PTSD, eating disorders and/or substance misuse	Stable psychotropic medication regimen of up to three psychotropic drugs allowed	In-patient	Cross-over	18 weeks	18	Clonidine	Placebo	Depression Dissociation/psychosis
Anticonvulsants versus placebo Cowdry and Gardner ⁴⁶	USA	100	31.6	Other Axis 2 personality disorders (dependent, avoidant, histrionic schizotypal, narcissistic and paranoid), and Axis 1 disorders (atypical bipolar disorder, major depressive episode with and without melancholia, dysthymia)	Unclear, no information	Out-patient	Cross-over, multi-arm	6 weeks	16 ^c	Carbamazepine	Placebo	Depressive symptoms Anxiety symptoms

(Continued)

Table 1 (Continued)

Study	Country	Gender (% female)	Age (years), mean	Sample population comorbidity	Concomitant medication allowed	Setting	Design	Sample size	Intervention	Comparator	Outcomes relevant to this review
Crawford et al ⁵⁶	UK	75.36	36.1	Unclear, there is no information on population comorbidity. However, many psychiatric disorders were exclusion criteria	Mood stabilisers were not allowed, no further information	In- and out-patients	Parallel	276	Lamotrigine plus usual care	Placebo plus usual care	Depressive symptoms
De la Fuente and Lotstra ⁵⁷	Belgium	70	32.7	Unclear, no information	Not allowed	In-patient	Parallel	20	Carbamazepine	Placebo	Depressive symptoms Anxiety symptoms Psychotic symptoms
Frankenburg and Zanarini ⁵⁸	USA	100	26.85	Bipolar disorder type 2	Psychotropics not allowed	Out-patient	Parallel	30	Valproate semisodium	Placebo	Depressive symptoms
Hollander et al ⁵⁹	USA	52.38	38.6	Unclear, no information	Unclear, no information	Out-patient	Parallel	21 ^a	Valproate semisodium	Placebo	Depressive symptoms
Loew et al ⁶⁰	Germany	100	25.25	Unclear, no information	Psychotropics not allowed	Out-patient	Parallel	56	Topiramate	Placebo	Depressive symptoms Anxiety symptoms Psychotic symptoms
Antidepressants versus placebo Cowdry and Gardner ⁴⁶	USA	100	31.6	Other Axis 2 personality disorders (dependent, avoidant, histrionic schizotypal, narcissistic and paranoid), and Axis 1 disorders (atypical bipolar disorder, major depressive episode with and without melancholia, dysthymia)	Unclear, no information	Out-patient	Cross-over, multi-arm	16 ^c	Tranylcypromine sulfate	Placebo	Depressive symptoms Anxiety symptoms
Salzman et al ⁶¹	USA	63.64 ^d	36.3	Unclear, no information	Psychotropics not allowed	Out-patient	Parallel	22	Fluoxetine	Placebo	Depressive symptoms
Simpson et al ⁶²	USA	100	36.26	Comorbid Axis I pathology. No further information	Trazodone 50–100 mg/d for insomnia	In- and out-patient	Parallel	25	Fluoxetine plus DBT	Placebo plus DBT	Anxiety symptoms Psychotic symptoms
Soloff et al ⁶³	USA	75.56	25.1	Comorbid schizotypal personality disorder ^{a,^}	Biperiden hydrochloride (2 mg)	In-patient	Parallel, multi-arm	90	Amitriptyline (haloperidol)	Placebo	Anxiety Depression Dissociation/psychosis
Cornelius et al ⁶⁴	USA	75.93	26.7	Major depressive disorder	Not allowed	In- and out-patients	Parallel, multi-arm	108	Phenelzine sulphate (haloperidol)	Placebo	Depression Anxiety Dissociation/psychosis

SOLD-II, Structured Clinical Interview for DSM personality disorders; BPD, borderline personality disorder; SASPD, Standardised Assessment of Severity of Personality Disorder; ADHD, attention-deficit hyperactivity disorder; DBT, dialectical behavioural therapy; CNS, central nervous system; PTSD, post-traumatic stress disorder; ^, inclusion criteria to enter the trial.

a. Initially, 21 participants entered the trial, and only 16 were randomised to a treatment group. No explanation for this was reported.

b. Trial had three arms with an overall sample size of 451, the sample size reported in this table is for the olanzapine 5–10 mg/d arm and placebo arm only.

c. Only included in secondary sensitivity analyses (cross-over trial with only end-of-period data available).

d. Completers only.

quetiapine, trifluoperazine and ziprasidone),^{28,44–49,51,53,55,63–67} anticonvulsants (including carbamazepine, lamotrigine, topiramate, valproate)^{46,57–60,68} and antidepressants (amitriptyline, fluoxetine, phenelzine, tranylcypromine).^{46,53,61,62,67} Three trials^{49,62,64} had a concomitant dialectical behavioural therapy intervention in both arms, and one⁶⁸ had concomitant treatment as usual in both arms (see main characteristics of trials that went into quantitative analyses, detailed in Table 1).

Assessment of co-occurring psychopathology in the primary studies

In general, end-point data on symptoms of co-occurring psychopathology were sparsely investigated (Table 1). The outcome most reported on by the primary studies was depression. Thirteen trials of antipsychotics^{44,46–49,51,53,55,63–67} reported on this outcome, plus all included six trials of anticonvulsants^{57–60,68,69} and all five trials of antidepressants.^{46,53,61,62,67} Psychotic/dissociative symptoms were reported on by nine trials testing antipsychotics,^{28,45,51,53,55,63,66,67,70} three trials of anticonvulsants^{57,60,68} and three trials of antidepressants.^{53,62,67} The third outcome for which analyses could be done was anxiety, which was reported on by six antipsychotic trials,^{46,51,53,63,64,67} three anticonvulsant studies^{46,57,60} and four antidepressant trials.^{46,53,62,67} An overview of the measurement scales used to assess outcomes in the primary studies is available in Supplementary Appendix 5.

Although other outcomes of co-occurring psychopathology have been assessed by eligible primary studies, i.e. obsessive-compulsive disorder (OCD) symptoms^{53,55,63,70} substance use,^{45,68} schizotypal symptoms,^{53,67} bipolar disorder symptoms⁴⁴ and PTSD,⁵⁵ meta-analyses could not be conducted for any of them because of sparse data per intervention and outcome that would not have allowed for pooling at least three individual effect estimates.

Risk of bias in included trials

None of trials included in the quantitative analysis had an overall low risk of bias. Fourteen trials^{28,44–46,48,49,51,58,59,62,64–66,68} had a high risk of bias in at least one domain, and the remaining had an unclear risk of bias in at least three domains.^{28,53,57,60,61,63,67,70}

Random sequence generation was assessed to be of low risk of bias in eight trials,^{28,48,49,51,58,65,66,68} whereas the remaining trials did not describe their randomisation procedure sufficiently to permit a judgement of low or high risk, and for this reason, they were classed as of unclear risk of bias. Four trials^{28,48,58,68} were assessed to be of low risk of bias regarding allocation concealment. The remaining trials were of unclear risk of bias in this domain because insufficient reporting did not allow for a judgement of low or high risk of bias. Blinding of participants and personnel were assessed to be of low risk of bias in nine trials,^{44–46,48,58,63,67,68,70} high in one trial²⁸ and the remaining trials in this domain were of unclear risk of bias. Blinding of outcome assessment was of low risk of bias in nine trials,^{49,53,55,57,59,61,62,67,68} high in one trial²⁸ and all remaining trials in this domain were of unclear risk of bias. Incomplete outcome data was assessed as low risk of bias in five trials,^{28,53,63,67,70} and high risk of bias in eight trials.^{44,46,48,51,55,59,62,68} The remaining trials in this domain were assessed as unclear risk of bias. Similarly, for selective reporting, four trials were assessed as having a high risk of bias^{28,44,65,71} whereas the majority of the remaining trials had an unclear risk of bias. Eleven of the trials were assessed to be at high risk of bias because of vested interests (i.e. author affiliated with or funding of trials by pharmaceutical companies,^{44,45,48,49,51,52,58,59,62,65,66} and the majority of the remaining trials were assessed to be of unclear risk of bias in this regard. There seemed to be no indication of any other potential source of bias for the majority trials. One trial⁴⁶ was rated as unclear in this domain because of an obvious carry-over effect between medication phases, and one trial²⁸ was rated as having a high risk of bias because of the impact the COVID-19 pandemic had on the trial (early termination owing to severe problems recruiting participants). A risk-of-bias summary and graph are available in Supplementary Appendix 6.

General treatment effects of major psychopharmacological classes on co-occurring psychopathology

Pharmacotherapy effects on co-occurring depressive symptoms
Twelve trials^{44,46,48,49,51,53,63,64–67,70} comparing antipsychotics and placebo had end-point data for depressive symptoms. There was evidence of a difference in depressive symptoms at the end

Table 2 Effects of medications on co-occurring depressive, anxiety and dissociative-psychotic symptoms – summary of findings							
Outcomes and interventions	Trials	n	Standardised mean difference ^a	95% CI	P-value	I ²	Certainty of evidence (GRADE)
Depressive symptoms							
Antipsychotics	12 ^{44,46,48,49,51,53,63–65,67,70}	1138	−0.22	−0.42 to −0.01	0.04	59%	⊕⊕⊕⊕ Very low ^{b,c}
Anticonvulsants	6 ^{46,56–60}	344	−0.44	−0.80 to −0.08	0.02	46%	⊕⊕⊕⊕ Low ^{b,d}
Antidepressants	5 ^{46,53,61,62,67}	187	−0.37	.082 to 0.08	0.11	52%	⊕⊕⊕⊕ Very low ^{d,e,f}
Dissociative-psychotic symptoms							
Antipsychotics	9 ^{28,45,51,53,63,65–67,70}	936	−0.28	−0.49 to −0.08	0.007	49%	⊕⊕⊕⊕ Low ^b
Anticonvulsants	3 ^{56,57,60}	270	−0.23	−0.66 to 0.20	0.30	51%	⊕⊕⊕⊕ Very low ^{b,f,d}
Antidepressants	3 ^{53,62,67}	139	−0.22	−0.62 to 0.18	0.29	25%	⊕⊕⊕⊕ Very low ^{b,f,d}
Anxiety symptoms							
Antipsychotics	6 ^{46,51,53,63,64,67}	309	−0.35	−0.72 to 0.02	0.06	61%	⊕⊕⊕⊕ Low ^{b,d}
Anticonvulsants	3 ^{46,57,63}	104	−1.11	−1.60 to −0.62	<0.00001	24%	⊕⊕⊕⊕ Very low ^{b,d,f}
Antidepressants	4 ^{46,53,62,67}	164	−0.23	−0.58 to 0.12	0.20	17%	⊕⊕⊕⊕ Very low ^{b,d,f}
Grading of Recommendations, Assessment, Development and Evaluations (GRADE) Working Group grades of evidence: ‘high certainty’, further research is very unlikely to change our confidence in the estimate of effect; moderate certainty’, further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; ‘low certainty’, further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; ‘very low certainty’, we are very uncertain about the estimate.							
a. Pooled standardised mean differences; negative values indicate beneficial effects of the experimental treatment. The inverse variance method and random effect model were used for pooling of effect estimates.							
b. Downgraded two levels because of the inclusion of studies with a high risk of bias (i.e. several domains with a high risk of bias or most domains with unclear risk of bias).							
c. Downgraded one level because of strongly suspected publication bias (cf. Fig. 2).							
d. Publication bias undetected; unable to draw a funnel plot because of too few primary studies (≤10).							
e. Downgraded two levels because of the inclusion of studies with a high risk of bias (i.e. several domains with a high risk of bias or most domains with unclear risk of bias).							
f. Downgraded two levels because of a small sample size of <50% of optimal information size (assumed as n ≥ 400).							

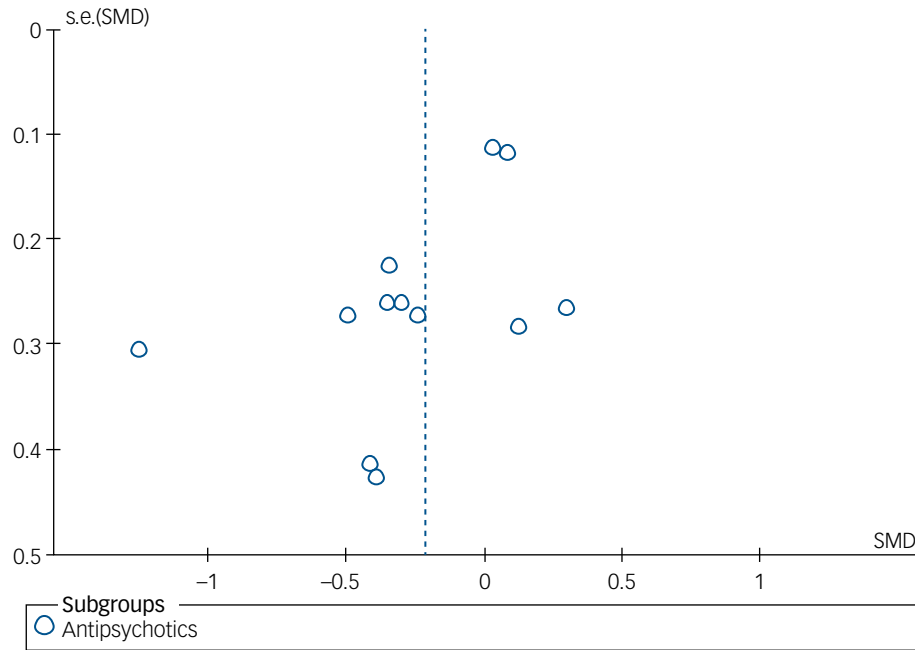


Fig. 2 Funnel plot of included trials with antipsychotic intervention on depression at the end of treatment. SMD, standardised mean difference.

of treatment (SMD -0.22 , 95% CI -0.42 to -0.01 , $P = 0.04$, 12 trials, $I^2 = 59\%$, very low certainty), favouring antipsychotics (Table 2). Upon visual inspection of the funnel plot, publication bias was strongly suspected in terms of a lack of studies reporting unfavourable effects (Fig. 2). The exclusion of the outlying study effect estimate of Nickel et al⁶³ (SMD -1.25 , 95% CI -1.85 to -0.65 , 52 participants) resulted in a considerably smaller pooled effect estimate and less statistical heterogeneity (SMD -0.09 , 95% CI -0.24 to 0.05 , $P = 0.26$, $I^2 = 19\%$, 1066 participants).

Five trials^{46,53,61,62,67} comparing antidepressants with placebo had end-point data for depressive symptoms. There was no evidence of a difference in depressive symptoms at the end of treatment between antidepressants and placebo (SMD -0.37 , 95% CI -0.82 to 0.08 , $P = 0.11$, five trials, $I^2 = 52\%$, very low certainty) (Table 2).

Six trials^{46,57–60,68} comparing anticonvulsants and placebo had end-point data for depressive symptoms. There was evidence of a difference in depression at the end of treatment, favouring anticonvulsants (SMD -0.44 , 95% CI -0.80 to -0.08 , $P = 0.02$, six trials, $I^2 = 46\%$, low certainty) (Table 2).

Pharmacotherapy effects on co-occurring anxious symptoms (post hoc)

Ten trials comparing antipsychotics and placebo had end-point data for anxious symptoms.^{41,45,46,48,51,53,63,64,67,70} However, only six of them^{46,51,53,63,64,67} reported usable data for generating summary estimates. Overall, there was no evidence of a difference between antipsychotics and placebo on anxious symptoms at the end of treatment (SMD -0.35 , 95% CI -0.72 to 0.02 , $I^2 = 61\%$, $P = 0.06$, six trials, 309 participants, very low certainty).

Six trials comparing antidepressants and placebo had end-point data for anxious symptoms.^{42,43,46,53,62,67} Unfortunately, two of them^{42,43} did not report usable data for generating summary estimates.

There was no evidence of a difference between antidepressants and placebo on anxiety (SMD -0.23 , 95% CI -0.58 to 0.12 , $I^2 = 17\%$, $P = 0.20$, four trials, 164 participants, very low certainty) (Table 2).

Three trials^{46,57,60} comparing anticonvulsants and placebo had end-point data for anxious symptoms. The summary estimate

indicates evidence of a difference between anticonvulsants and placebo on anxiety, favouring anticonvulsants (SMD -1.11 , 95% CI -1.60 to -0.62 , $I^2 = 24\%$, $P < 0.00001$, three trials, 104 participants, very low certainty) (Table 2).

Pharmacotherapy effects on co-occurring dissociative symptoms

Nine trials^{28,45,51,53,63,65–67,70} comparing antipsychotics with placebo had end-point data for dissociative symptoms. There was evidence of a difference between groups at the end of treatment, favouring antipsychotics (SMD -0.28 , 95% CI -0.49 to -0.08 , $I^2 = 49\%$, nine trials, $P = 0.007$, 936 participants, low certainty).

Three trials^{53,62,67} with antidepressant interventions reported end-of-treatment data for dissociative symptoms. There was no evidence of a difference between antidepressants and placebo on dissociative symptoms at the end of treatment (SMD -0.22 , 95% CI -0.62 to 0.18 , $I^2 = 25\%$, $P = 0.29$, three trials, 139 participants, very low certainty) (Table 2). Only three trials^{57,60,68} reported dissociative symptoms for anticonvulsants at the end of treatment, and there was no evidence of a difference between groups (SMD -0.23 , 95% CI -0.66 to 0.20 , $I^2 = 51\%$, $P = 0.30$, three trials, 270 participants, very low certainty) (Table 2).

Treatment effects of therapeutic classes: the role of co-occurring psychopathology

Treatment effects on depressive symptoms in trials including or excluding individual co-occurring psychopathology

Subgroup analyses were done for trials excluding versus not excluding PTSD, OCD, depression, bipolar disorder or psychotic disorders (Table 3, Supplementary Appendix 7). No statistically significant subgroup differences were observed for any of these analyses. However, data indicated a borderline significant subgroup differences ($P = 0.05$) between subgroups of trials that excluded participants with PTSD and those that did not: a numerically larger treatment effect was observed in the trials that excluded PTSD (SMD -0.32 , 95% CI -0.60 to -0.03) versus those that did not exclude PTSD (SMD 0.02 , 95% CI -0.15 to 0.18). The same was the case for subgroups of trials that excluded participants

Table 3 Differential effects of medications on co-occurring depressive, anxiety and dissociative–psychotic symptoms in relation to psychiatric exclusion criteria – summary of findings								
Outcomes and interventions	Trials	<i>n</i>	Standardised mean difference ^a	95% CI	<i>P</i> -value	<i>I</i> ²	Certainty of evidence (GRADE)	χ^2 -Test for subgroup differences (<i>p</i>) ^b
Antipsychotics: depression								
Substance use excluded	6 ^{45,51,65–67,70}	799	–0.16	–0.30 to –0.02	0.02	0%	⊕⊕⊕⊕ Low ^{c,d}	0.001
Substance use not excluded	3 ^{28,53,63}	137	–0.79	–1.14 to –0.44	<0.0001	0%	⊕⊕⊕⊕ Very low ^{d,e,f}	
Grading of Recommendations, Assessment, Development and Evaluations (GRADE) Working Group grades of evidence: ‘high certainty’, further research is very unlikely to change our confidence in the estimate of effect; ‘moderate certainty’, further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; ‘low certainty’, further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; ‘very low certainty’, we are very uncertain about the estimate.								
a. Pooled standardised mean differences; negative values indicate beneficial effects of the experimental treatment. The inverse variance method and random effect model were used for pooling of effect estimates.								
b. When including study data from the cross-over study of Ziegenhorn et al ⁵⁵ in a predefined secondary sensitivity analyses (see Stoffers-Winterling et al ¹²), the χ^2 -test for subgroup differences reached statistical significance (<i>P</i> = 0.04) for the outcome of depression, if trials excluding (four trials, <i>N</i> = 473, SMD –0.32, 95% CI –0.60 to –0.03, <i>P</i> = 0.03, <i>I</i> ² = 55%) versus not excluding obsessive–compulsive disorder (four trials, <i>n</i> = 683, SMD 0.02, 95% CI –0.13 to 0.17, <i>P</i> = 0.84, <i>I</i> ² = 0%) were compared. No substantial changes were observed if including Ziegenhorn et al with any of the other relevant comparisons in terms of pooled effect estimates changing the direction of effect or crossing boundaries of confidence intervals.								
c. Downgraded two levels because of the inclusion of studies with a high risk of bias (i.e. several domains with a high risk of bias or most domains with unclear risk of bias).								
d. Publication bias undetected; unable to draw a funnel plot because of too few primary studies (≤ 10).								
e. Downgraded two levels because of a small sample size <50% of optimal information size (assumed as <i>n</i> ≥ 400).								
f. Downgraded one level because of the inclusion of studies with moderate risk of bias (i.e. high risk of bias for a maximum of one domain and/or unclear risk of bias for several domains).								

with OCD versus those that did not (test for subgroup differences: *P* = 0.05): numerically larger effects were reported for antipsychotics in trials that excluded OCD (SMD –0.32, 95% CI –0.60 to –0.03) versus those that did not exclude OCD (SMD 0.02, 95% CI –0.15 to 0.18).

Excluding the outlying effect estimate of Nickel et al from the subgroup analyses comparing trials with and without PTSD or OCD exclusion resulted in substantially lower heterogeneity within the individual comparison groups (PTSD/OCD excluded; *I*² = 1% instead of *I*² = 55%), a smaller pooled effect estimate (SMD –0.20 instead of SMD –0.32) and diminishing of the borderline significant subgroup effect (*P* = 0.10 instead of *P* = 0.05). Similarly, excluding the outlying effect estimate of Nickel et al from the subgroup analysis regarding psychotic disorders resulted in less heterogeneity within study estimates from trials excluding psychotic disorders (*I*² = 8% instead of *I*² = 63%), and a smaller pooled effect estimate (SMD –0.09 instead of SMD 0.28). The test for subgroup differences remained not significant (*P* = 0.80 instead of *P* = 0.35).

For antidepressants, no subgroup analyses could be done, as the minimum number of effect estimates per analysis was not reached.

For anticonvulsants compared with placebo, one subgroup analysis on trials excluding or including participants with depression was possible. There was no evidence of a difference in effect on depression at the end of treatment between trials that included participants with depression and trials that specifically excluded participants with depression (Table 3).

Treatment effects on anxious symptoms in trials including or excluding individual co-occurring psychopathology

For antipsychotics compared with placebo, it was not possible to conduct subgroup analyses, since the minimum number of three studies per subgroups was not reached for any psychopathology category (Table 3).

It was also not possible to do subgroup analyses for antidepressants or anticonvulsants, as the minimum number of effect estimates per subgroup was not reached.

Treatment effects on dissociative symptoms in trials including or excluding individual co-occurring psychopathology

For antipsychotics compared with placebo, it was possible to conduct subgroup analyses by trials excluding versus not excluding substance use and depression, respectively. There was evidence of a difference in dissociative symptoms at the end of treatment between

trials that excluded participants with substance use (SMD –0.16, 95% CI –0.30 to –0.02) and those that did not (SMD –0.79, 95% CI –1.14 to –0.44), with a much larger effect size in the trials that did not exclude participants with substance use (test for subgroup differences: *P* = 0.001) (Table 3). There was no evidence of a difference in dissociative symptoms at the end of treatment between studies excluding and not excluding depression (Table 3).

For antidepressants and anticonvulsants, no subgroup analyses could be done, as the minimum number of effect estimates per subgroup was not reached.

Discussion

This systematic review aimed to assess the effects of pharmacological interventions on non-BPD-specific psychopathology, and whether people diagnosed with BPD and co-occurring mental health disorders might respond differently to medications.

First, we assessed the evidence of pharmacological interventions on symptoms of co-occurring disorders in people with BPD. We were able to investigate the effects of co-occurring depressive, anxious and psychotic symptoms. We found evidence of small effects of antipsychotics on depressive (very low certainty) and dissociative symptoms (low certainty). However, the effect of antidepressants was not robust to the exclusion of an outlying study effect estimate. Moderate to large effects were found for anticonvulsants regarding depressive and anxiety symptoms (very low certainty). No evidence of any effects of antidepressants was observed for any outcome (very low certainty). Our previous Cochrane review already failed to identify any effects of antidepressants on BPD pathology (including impulsivity, for which selective serotonin reuptake inhibitors are frequently used), and this review complements the findings with regard to non-BPD-specific psychopathology. Given antidepressants are still among the most prevalently administered medications for patients with BPD in clinical settings,^{72,73} the results of this review are of high relevance to clinical practice.

Second, we explored if different effects could be observed in persons with BPD plus distinct psychiatric comorbidities. Comparing effect estimates from samples excluding versus including different types of co-occurring psychiatric disorders, we did not find any differences in effect regarding PTSD, OCD, MDD, bipolar disorders or psychotic disorders. The only subgroup difference found related to SUDs: trials including participants with SUDs reported a significantly greater reduction in dissociative symptoms by antipsychotics

than those samples excluding people with a co-existing SUD. In other terms, antipsychotics had better outcomes in samples including participants with dual diagnoses. Although only a tentative finding, it might be of interest, as this patient group is conventionally regarded as having an unfavourable prognosis and worse treatment outcomes than persons with either condition alone.⁷⁴ This assumption, however, has already been challenged by the availability of tailored psychotherapies (e.g. dialectical behaviour therapy for substance use disorders) and long-term observations.⁷⁵

Strengths and limitations

To our knowledge, this is the first meta-analysis that systematically investigates the effects of pharmacological interventions with BPD plus co-occurring disorders. It also highlights a lack of evidence on co-occurring psychopathology, since studies usually focus on few outcomes. Guideline developers and clinicians should be aware of these evidence gaps. Robust methods were applied as within the main Cochrane review, including comprehensive literature searches, which were updated for this review in April 2024. Predefined methods were applied, and any *post hoc* modifications made clear (see also Supplementary Appendix 2). The exploratory finding of greater effects by antipsychotics in samples including persons with a diagnosis of SUD may stimulate further research.

This review has some limitations. Ideally, co-occurring psychopathology in participants of trial samples should be investigated in the confirmed presence of a given disorder, and not based on exclusion criteria of trials. Given that such trials are lacking, and the question of intervention effects of pharmacotherapies in people with BPD is of high clinical relevance, we used exclusion criteria of primary studies as a proxy of the co-occurrence of disorders. A more ideal way of investigating the co-occurrence of disorders in our BPD samples would be to conduct meta-analyses of individual participant data, in which there is a certainty of co-occurring disorders in each participant. Such analyses are planned and underway for psychotherapeutic interventions,⁷⁶ but to our knowledge, none are planned for pharmacological interventions.

Overall, the certainty of the evidence was low to very low. The risk of bias in included trials was overall moderate to high. Most of the included trials had small sample sizes and, as a result, the information size in most of the meta-analyses was low. This is especially true for the subgroup analyses, and as such, results from these should be considered exploratory. Given the low overall information size and small samples included in the individual trials, the power to detect existing differences was very likely too small (type 2 error); on the other hand, a random inflation of type 1 errors (i.e. falsely detecting a difference where there is none) cannot be ruled out. It was only possible to investigate publication bias by funnel plot inspection in one analysis (Fig. 2), therefore publication bias in the remaining analyses cannot be ruled out. Moreover, some estimates had significant heterogeneity. Most analyses in this review were conducted *post hoc*. We included multiple outcome measures and did not conduct tests for multiplicity, so the result from this article should be considered exploratory and for hypothesis generation. The absence of reported 95% confidence intervals alongside our calculated I^2 values should be acknowledged as a limitation. The interpretation of I^2 estimates warrants caution, especially in meta-analyses comprising a limited number of events or trials. Nonetheless, we addressed some of the observed heterogeneity by following the recommendation outlined in the Cochrane Handbook, specifically by excluding outliers. This approach is consistent with established methodologies.²²

In conclusion, although our Cochrane review on medications for treatment of BPD failed to demonstrate substantial effects, medications are nonetheless prevalently prescribed to people with BPD,

oftentimes with the rationale to target co-occurring psychopathology. Indeed, persons with BPD usually present with psychiatric symptoms that are not specific to BPD (e.g. depressive or anxious symptoms), which may or may not necessarily be explained by another co-occurring psychiatric disorder. However, a systematic review of medication effects on such co-occurring psychopathology, and the role of manifest comorbidities for treatment effects, has, to our knowledge, not yet been performed. The findings of this review aim to enable clinicians and persons affected by BPD to make informed decisions on whether to use medications to treat co-occurring psychopathology.

Medication effects on co-occurring psychopathology (i.e. depressive, dissociative–psychotic and anxiety symptoms) were small (antipsychotics) to moderate (anticonvulsants) to moderate. Tentative evidence suggests that persons with dual diagnoses (i.e. having BPD plus any SUD) might benefit better from antipsychotics regarding dissociative–psychotic symptoms than persons with BPD only.

It is important not to take the absence of robust, supporting evidence of medication effects, or limited findings of differential treatment effects linked to comorbidities, as ‘proof’ that no such effects or differences existed. Our findings, however, based on the currently available evidence – which is of low to very low certainty only – do not support the standard use of pharmacological interventions in people with BPD to target co-occurring psychopathology.

Research is urgently needed to shed light on medication effects in persons with defined comorbidities, given the vast majority of persons with BPD is affected by psychic comorbidities,⁷⁷ and these comorbidities are usually accounted for as the major reason for using medications.

Johanne Pereira Ribeiro , Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Slagelse, Denmark; and Department of Psychology, University of Southern Denmark, Odense, Denmark; **Sophie Juul** , Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region of Denmark, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark; and Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Copenhagen, Denmark; **Mickey T. Kongerslev** , Department of Psychology, University of Southern Denmark, Odense, Denmark; and Mental Health Services West, Psychiatry Region Zealand, Slagelse, Denmark; **Mie Sedoc Jørgensen** , Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Slagelse, Denmark; **Birgit A Völlm** , Department for Forensic Psychiatry, University Medicine Rostock, Rostock, Germany; **Henriette Edemann-Callesen** , Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Slagelse, Denmark; **Christian Sales** , Nottinghamshire Healthcare NHS Foundation Trust, Nottingham, UK; **Julie P. Schaag** , Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Slagelse, Denmark; **Klaus Lieb** , Department of Psychiatry and Psychotherapy, University Medical Center Mainz, Mainz, Germany; **Erik Simonsen** , Mental Health Services East, Copenhagen University Hospital, Psychiatry Region Zealand, Copenhagen, Denmark; and Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; **Jutta M. Stoffers-Winterling** , Department of Psychiatry and Psychotherapy, University Medical Center Mainz, Mainz, Germany; **Ole Jakob Storebø** , Center for Evidence-Based Psychiatry, Psychiatric Research Unit, Psychiatry Region Zealand, Slagelse, Denmark; and Department of Psychology, University of Southern Denmark, Odense, Denmark

Correspondence: Jutta M. Stoffers-Winterling. Email: j.stoffers-winterling@uni-mainz.de

First received 30 Dec 2023, final revision 5 Jul 2024, accepted 9 Jul 2024

Supplementary material

Supplementary material is available online at <https://doi.org/10.1192/bjp.2024.172>

Data availability

The data supporting the findings of this study are available within the article and/or its Supplementary materials. More data for reproducing results or replicating procedures are available from the original Cochrane review and/or on request from the corresponding author, O.J.S.

Author contributions

The study was conceptualised by J.P.R., J.M.S.-W., S.J. and O.J.S. J.P.R. took the lead in drafting the initial manuscript, which was subsequently critically reviewed by all authors. The statistical

analyses GRADE assessments were collectively conducted by J.P.R. and J.M.S.-W. The interpretation of the findings involved contributions from all authors. Consensus was reached among all authors regarding the final version of the manuscript, and unanimous agreement was reached for its submission.

Funding

This study received no specific grant from any funding agency, commercial or not-for-profit sectors.

Declaration of interest

None of the authors have any competing interests to declare.

References

1 Tomko RL, Trull TJ, Wood PK, Sher KJ. Characteristics of borderline personality disorder in a community sample: comorbidity, treatment utilization, and general functioning. *J Pers Disord* 2014; **28**: 734–50.

2 Bohus M, Stoffers-Winterling J, Sharp C, Krause-Utz A, Schmahl C, Lieb K. Borderline personality disorder. *Lancet* 2021; **398**: 1528–40.

3 Shah R, Zanarini MC. Comorbidity of borderline personality disorder: current status and future directions. *Psychiatr Clin North Am* 2018; **41**: 583–93.

4 Hoertel N, López S, Wang S, González-Pinto A, Limosin F, Blanco C. Generalizability of pharmacological and psychotherapy clinical trial results for borderline personality disorder to community samples. *Personal Disord* 2015; **6**: 81–7.

5 Stoffers-Winterling J, Kongerslev M, Simonsen E, Lieb K, Storebø OJ. Clinical trials in personality disorders: recommendations for study design and conduct. *Personal Disord* 2023; **14**: 127–36.

6 National Health and Medical Research Council. *Clinical Practice Guideline for the Management of Borderline Personality Disorder*. National Health and Medical Research Council, 2013.

7 National Institute for Health and Care Excellence (NICE). *Borderline Personality Disorder: Recognition and Management*. Clinical Guideline [CG78]. NICE, 2009 (www.nice.org.uk/guidance/cg78).

8 Stoffers-Winterling J, Taubner S, Renneberg B, Lieb K. S3-Leitlinie borderline-persönlichkeitsstörung. [S3 guidelines for borderline personality disorder.] *Nervenarzt* 2023; **94**: 1062–74.

9 Simonsen S, Bateman A, Bohus M, Dalewijk HJ, Doering S, Kaera A, et al. European guidelines for personality disorders: past, present and future. *Borderline Personal Disord Emot Dysreg* 2019; **6**: 9.

10 Pascual JC, Martín-Blanco A, Soler J. Twenty-year trends in the psychopharmacological treatment of outpatients with borderline personality disorder: a cross-sectional naturalistic study in Spain. *CNS Drugs* 2021; **35**: 1023–32.

11 Riffer F, Farkas M, Streibl L, Kaiser E, Sprung M. Psychopharmacological treatment of patients with borderline personality disorder: comparing data from routine clinical care with recommended guidelines. *Int J Psychiatry Clin Pract* 2019; **23**: 178–88.

12 Stoffers-Winterling JM, Storebø OJ, Pereira Ribeiro J, Kongerslev MT, Völm BA, Mattivi JT, et al. Pharmacological interventions for people with borderline personality disorder. *Cochrane Database Syst Rev* 2022; **11**: CD012956.

13 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* (5th edn). American Psychiatric Publishing, 2013.

14 Stoffers-Winterling JM, Storebø OJ, Völm BA, Mattivi JT, Nielsen SS, Kielsholm ML, et al. Pharmacological interventions for people with borderline personality disorder. *Cochrane Database Syst Rev* 2018; **2**: CD012956 (<https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012956/full>).

15 Fisher G, Roget N. Diagnostic and statistical manual of mental disorders. In *Encyclopedia of Substance Abuse Prevention, Treatment, & Recovery* (eds GL Fisher, NA Roget): 283–6. SAGE Publications, 2014.

16 Higgins JPT, Altman DG, S J. Chapter 8: assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (eds JPT Higgins, S Green): Chapter 8. The Cochrane Collaboration, 2011.

17 Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. *Chinese J Evidence-Based Med* 2011; **11**: 1459–63.

18 Guyatt G, Zhao Y, Mayer M, Briel M, Mustafa R, Izcovich A, et al. GRADE guidance 36: updates to GRADE’s approach to addressing inconsistency. *J Clin Epidemiol* 2023; **158**: 70–83.

19 Zeng L, Brignardello-Petersen R, Hultcrantz M, Mustafa RA, Murad MH, Iorio A, et al. GRADE guidance 34: update on rating imprecision using a minimally contextualized approach. *J Clin Epidemiol* 2022; **150**: 216–24.

20 Page M, Higgins J, Sterne J, et al. Chapter 13: assessing risk of bias due to missing results in a synthesis. In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (Updated August 2023)* (eds J Higgins, J Thomas, J Chandler, M Cumpston, T Li, M Page, et al): Chapter 13. The Cochrane Collaboration, 2023.

21 Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–60.

22 Deeks J, Higgins J, Altman D. Chapter 10: analysing data and undertaking meta-analyses. In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (6.4.)* (eds J Higgins, J Thomas, J Chandler, M Cumpston, T Li, M Page, et al): Chapter 10. The Cochrane Collaboration, 2023.

23 Viechtbauer W, Cheung MW-L. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods* 2010; **1**: 112–25.

24 Curtin F, Elbourne D, Altman DG. Meta-analysis combining parallel and cross-over clinical trials. III: the issue of carry-over. *Stat Med* 2002; **21**: 2161–73.

25 Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2002; **31**: 140–9.

26 The Cochrane Collaboration. *Review Manager 5 (RevMan 5)*, version 5.4. The Cochrane Collaboration, 2020 (<https://training.cochrane.org/online-learning/core-software/revman>).

27 Boehringer Ingelheim. *A Study to Test Different Doses of BI 1358894 and Find Out Whether They Reduce Symptoms in People With Borderline Personality Disorder*. Clinicaltrials.gov, 2024 (<https://clinicaltrials.gov/study/NCT04566601?a=14>).

28 Crawford MJ, Leeson VC, Evans R, Barrett B, McQuaid A, Cheshire J, et al. The clinical effectiveness and cost effectiveness of clozapine for inpatients with severe borderline personality disorder (CALMED study): a randomised placebo-controlled trial. *Ther Adv Psychopharmacol* 2022; **12**: 204512532210908.

29 Maoz H, Grossman-Giron A, Sedoff O, Nitzan U, Kashua H, Yarmishin M, et al. Intranasal oxytocin as an adjunct treatment among patients with severe major depression with and without comorbid borderline personality disorder. *J Affect Disord* 2024; **347**: 39–44.

30 Moen R, Freitag M, Miller M, Lee S, Romine A, Song S. Efficacy of extended-release divalproex combined with ‘condensed’ dialectical behavior therapy for individuals with borderline personality disorder. *Ann Clin Psychiatry* 2012; **24**: 255–60.

31 Montgomery SA, Montgomery D. Pharmacological prevention of suicidal behaviour. *J Affect Disord* 1982; **4**: 291–8.

32 New York State Psychiatric Institute. *Treating Suicidal Behavior and Self-Mutilation in People With Borderline Personality Disorder*. Clinicaltrials.gov, 2007 (<https://clinicaltrials.gov/show/NCT00533117>).

33 Nickel MK, Nickel C, Kaplan P, Lahmann C, Mühlbacher M, Tritt K, et al. Treatment of aggression with topiramate in male borderline patients: a double-blind, placebo-controlled study. *Biol Psychiatry* 2005; **57**: 495–9.

34 Reich DB, Zanarini MC, Bieri KA. A preliminary study of lamotrigine in the treatment of affective instability in borderline personality disorder. *Int Clin Psychopharmacol* 2009; **24**: 270–5.

35 Tritt K, Nickel C, Lahmann C, Leiberich PK, Rother WK, Loew TH, et al. Lamotrigine treatment of aggression in female borderline-patients: a randomized, double-blind, placebo-controlled study. *J Psychopharmacol* 2005; **19**: 287–91.

36 Kulkarni J, Thomas N, Hudaib AR, Gavrilidis E, Grigg J, Tan R, et al. Effect of the glutamate NMDA receptor antagonist memantine as adjunctive treatment in borderline personality disorder: an exploratory, randomised, double-blind, placebo-controlled trial. *CNS Drugs* 2018; **32**: 179–87.

37 Amminger GP, Chanan AM, Ohmann S, Klier CM, Mossaheb N, Bechdolf A, et al. Omega-3 fatty acid supplementation in adolescents with borderline personality disorder and ultra-high risk criteria for psychosis: a post hoc subgroup analysis of a double-blind, randomised controlled trial. *Can J Psychiatry* 2013; **58**: 402–8.

38 Hallahan B, Hibbeln JR, Davis JM, Garland MR. Omega-3 fatty acid supplementation in patients with recurrent self-harm: single-centre double-blind randomised controlled trial. *Br J Psychiatry* 2007; **190**: 118–22.

39 Schmahl C, Kleindienst N, Limberger M, Ludäscher P, Mauchnik J, Deibler P, et al. Evaluation of naltrexone for dissociative symptoms in borderline personality disorder. *Int Clin Psychopharmacol* 2012; **27**: 61–8.

40 Zanarini MC, Ed D, Frankenburg FR. Omega-3 fatty acid treatment of women with borderline personality disorder. *Am J Psychiatry* 2003; **160**: 167–9.

41 Zanarini MC, Frankenburg FR. Olanzapine treatment of female borderline personality disorder patients: a double-blind, placebo-controlled pilot study. *J Clin Psychiatry* 2001; **62**: 849–54.

- 42 Rinne T, Van den Brink W, Wouters L, Van Dyck R. SSRI treatment of borderline personality disorder: a randomized, placebo-controlled clinical trial for female patients with borderline personality disorder. *Am J Psychiatry* 2002; **159**: 2048–54.
- 43 Markovitz P. Pharmacotherapy of impulsivity, aggression, and related disorders. In *Impulsivity and Aggression* (eds Stein DJ, Hollander E): 263–87. Wiley, 1995.
- 44 Black DW, Zanarini MC, Romine A, Shaw M, Allen J, Schulz SCC. Comparison of low and moderate dosages of extended-release quetiapine in borderline personality disorder: a randomized, double-blind, placebo-controlled trial. *Am J Psychiatry* 2014; **171**: 1174–82.
- 45 Bogenschutz MP, Nurnberg HG. Olanzapine versus placebo in the treatment of borderline personality disorder. *J Clin Psychiatry* 2004; **65**: 104–9.
- 46 Cowdry RW, Gardner DL. Pharmacotherapy of borderline personality disorder: alprazolam, carbamazepine, trifluoperazine, and tranylcypromine. *Arch Gen Psychiatry* 1988; **45**: 111–9.
- 47 Goldberg SC, Schulz SC, Schulz PM, Resnick RJ, Hamer RM, Friedel RO. Borderline and schizotypal personality disorders treated with low-dose thiothixene vs placebo. *Arch Gen Psychiatry* 1986; **43**: 680–6.
- 48 Grant JE, Valle S, Chesivoir E, Ehsan D, Chamberlain SR. A double-blind placebo-controlled study of brexpiprazole for the treatment of borderline personality disorder. *Br J Psychiatry* 2022; **220**: 58–63.
- 49 Linehan MM, McDavid JP, Brown MZ, Sayrs JHR, Gallop RJ. Olanzapine plus dialectical behavior therapy for women with high irritability who meet criteria for borderline personality disorder: a double-blind, placebo-controlled pilot study. *J Clin Psychiatry* 2008; **69**: 999–1005.
- 50 Nickel MK, Muehlbacher M, Nickel C, Kettler C, Gil FP, Bachler E, et al. Aripiprazole in the treatment of patients with borderline placebo-controlled study. *Am J Psychiatry* 2006; **163**: 833–8.
- 51 Pascual JC, Soler J, Puigdemont D, Pérez-Egea R, Tiana T, Alvarez E, et al. Ziprasidone in the treatment of borderline personality disorder: a double-blind, placebo-controlled, randomized study. *J Clin Psychiatry* 2008; **69**: 603–8.
- 52 Soler J, Pascual JC, Campins J, Barrachina J, Puigdemont D, Alvarez E, et al. Double-blind, placebo-controlled study of dialectical behavior therapy plus olanzapine for borderline personality disorder. *Am J Psychiatry* 2005; **162**: 1221–4.
- 53 Soloff PH, George A, Nathan RS, Schulz PM, Cornelius JR, Herring J, et al. Amitriptyline versus haloperidol in borderlines: final outcomes and predictors of response. *J Clin Psychopharmacol* 1989; **9**: 238–46.
- 54 Cornelius JR, Soloff PH, Perel JM, Ulrich RF. Continuation pharmacotherapy of borderline personality disorder with haloperidol and phenelzine. *Am J Psychiatry* 1993; **150**: 1843–8.
- 55 Ziegenhorn AA, Roepke S, Schommer NC, Merkl A, Danker-Hopfe H, Perschel FH, 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; **29**: 170–3.
- 56 Crawford MJ, Sanatinia R, Barrett B, Cunningham G, Dale O, Ganguli P, et al. Lamotrigine for people with borderline personality disorder: a RCT. *Health Technol Assess* 2018; **22**: 1–68.
- 57 De la Fuente J, Lotstra F. A trial of carbamazepine in borderline personality disorder. *Eur Neuropsychopharmacol* 1994; **4**: 479–86.
- 58 Frankenburg FR, Zanarini MC. Divalproex sodium treatment of women with borderline personality disorder and bipolar II disorder: a double-blind placebo-controlled pilot study. *J Clin Psychiatry* 2002; **63**: 442–6.
- 59 Hollander E, Allen A, Lopez RP, Bienstock CA, Grossman R, Siever LJ, et al. A preliminary double-blind, placebo-controlled trial of divalproex sodium in borderline personality disorder. *J Clin Psychiatry* 2001; **62**: 199–203.
- 60 Loew TH, Nickel MK, Muehlbacher M, Kaplan P, Nickel C, Kettler C, et al. Topiramate treatment for women with borderline personality disorder: a double-blind, placebo-controlled study. *J Clin Psychopharmacol* 2006; **26**: 61–6.
- 61 Salzman C, Wolfson AN, Schatzberg A, Looper J, Henke R, Albanese M, et al. Effect of fluoxetine on anger in symptomatic volunteers with borderline personality disorder. *J Clin Psychopharmacol* 1995; **15**: 23–9.
- 62 Simpson EB, Yen S, Costello E, Rosen K, Begin A, Pistorello J, et al. Combined dialectical behavior therapy and fluoxetine in the treatment of borderline personality disorder. *J Clin Psychiatry* 2004; **65**: 379–85.
- 63 Nickel MK, Muehlbacher M, Nickel C, Kettler C, Gil FP, Bachler E, et al. Aripiprazole in the treatment of patients with borderline personality disorder: a double-blind, placebo-controlled study. *Am J Psychiatry* 2006; **163**: 833–8.
- 64 Soler J, Pascual JC, Campins J, Barrachina J, Puigdemont D, Alvarez E, et al. Double-blind, placebo-controlled study of dialectical behavior therapy plus olanzapine for borderline personality disorder. *Am J Psychiatry* 2005; **162**: 1221–4.
- 65 Schulz SC, Zanarini MC, Bateman A, Bohus M, Detke HC, Trzaskoma Q, et al. Olanzapine for the treatment of borderline personality disorder: variable dose 12-week randomised double-blind placebo-controlled study. *Br J Psychiatry* 2008; **193**: 485–92.
- 66 Zanarini MC, Schulz SC, Detke HC, Tanaka Y, Zhao F, Lin D, et al. A dose comparison of olanzapine for the treatment of borderline personality disorder: a 12-week randomized double-blind placebo-controlled study. *Eur Psychiatry* 2007; **22**: S172–3.
- 67 Soloff PH, Cornelius J, George A, Nathan S, Perel JM, Ulrich RF. Efficacy of phenelzine and haloperidol in borderline personality disorder. *Arch Gen Psychiatry* 1993; **50**: 377–85.
- 68 Crawford MJ, Sanatinia R, Barrett B, Cunningham G, Dale O, Ganguli P, et al. The clinical effectiveness and cost-effectiveness of lamotrigine in borderline personality disorder: a randomized placebo-controlled trial. *Am J Psychiatry* 2018; **175**: 756–64.
- 69 Gardner DL, Cowdry RW. Development of melancholia during carbamazepine treatment in borderline personality disorder. *J Clin Psychopharmacol* 1986; **6**: 236–9.
- 70 Goldberg CS, Schulz SC, Schulz MP, Resnick JR, Hamer MR, Friedel OR, et al. Borderline and schizotypal personality disorders treated with low-dose thiothixene vs placebo. *Arch Gen Psychiatry* 1986; **43**: 680–6.
- 71 Gardner DL, Cowdry RW. Positive effects of carbamazepine on behavioral dyscontrol in borderline personality disorder. *Am J Psychiatry* 1986; **143**: 519–22.
- 72 Paton C, Crawford MJ, Bhatti SF, Patel MX, Barnes TRE. The use of psychotropic medication in patients with emotionally unstable personality disorder under the care of UK mental health services. *J Clin Psychiatry* 2015; **76**: e512–8.
- 73 Martín-Blanco A, Ancochea A, Soler J, Elices M, Carmona C, Pascual JC. Changes over the last 15 years in the psychopharmacological management of persons with borderline personality disorder. *Acta Psychiatr Scand* 2017; **136**: 323–31.
- 74 Stetsiv K, McNamara IA, Nance M, Carpenter RW. The co-occurrence of personality disorders and substance use disorders. *Curr Psychiatry Rep* 2023; **25**: 545–54.
- 75 Newton-Howes GM, Foulds JA, Guy NH, Boden JM, Mulder RT. Personality disorder and alcohol treatment outcome: systematic review and meta-analysis. *Br J Psychiatry* 2017; **211**: 22–30.
- 76 Storebø OJ, Ribeiro JP, Kongerslev MT, Stoffers-Winterling J, Sedoc Jørgensen M, Lieb K, et al. Individual participant data systematic reviews with meta-analyses of psychotherapies for borderline personality disorder. *BMJ Open* 2021; **11**: e047416.
- 77 Zanarini MC, Athanasiadi A, Temes CM, Magni LR, Hein KE, Fitzmaurice GM, et al. Symptomatic disorders in adults and adolescents with borderline personality disorder. *J Pers Disord* 2021; **35**: 48–55.



Supplement S1: PRISMA 2020 Checklist



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	p. 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	p. 3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	p. 3
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	p. 5
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	p. 5
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	p. 5
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	p.5; Stoffers-Winterling et al., 2022 ¹
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	pp. 5-6, Stoffers-Winterling et al., 2022 ¹
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	pp. 5-6, Stoffers-Winterling et al., 2022 ¹
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	5-6, Stoffers-Winterling et al., 2022 ¹

Section and Topic	Item #	Checklist item	Location where item is reported
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	5-6, Stoffers-Winterling et al., 2022 ¹
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	pp.5-6
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	p. 6
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	p. 6
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Online supplement 2
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	p. 6
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	p. 6
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	p. 6
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	p. 6
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	p. 6
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	p. 6
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	p. 6; Fig. 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	p. 6; online supplement S4
Study characteristics	17	Cite each included study and present its characteristics.	Tab. 1
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	online supplement S6

Section and Topic	Item #	Checklist item	Location where item is reported
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	pp. 6-10, Tab. 2, Tab. 3
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	pp. 6-10, Tab.2, Tab. 3
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	pp. 6-10, Tab. 2, Tab. 3
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	p.8; p. 9.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	p.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	p. 8
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	pp. 6-10, Tab.2, Tab. 3
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	p. 10
	23b	Discuss any limitations of the evidence included in the review.	p. 11
	23c	Discuss any limitations of the review processes used.	p. 11
	23d	Discuss implications of the results for practice, policy, and future research.	pp. 11-12
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	p. 5
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	p. 5
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	pp. 4-6, online supplement S2
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	p. 12
Competing interests	26	Declare any competing interests of review authors.	p. 12
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	s. online Supplements S5; remaining

Section and Topic	Item #	Checklist item	Location where item is reported
			materials are available upon request

- ¹ Stoffers-Winterling, J. M., Storebø, O. J., Pereira Ribeiro, J., Kongerslev, M. T., Völlm, B. A., Mattivi, J. T., Faltinsen, E., Todorovac, A., Jørgensen, M. S., Callesen, H. E., Sales, C. P., Schaug, J. P., Simonsen, E., & Lieb, K. (2022). Pharmacological interventions for people with borderline personality disorder. *The Cochrane database of systematic reviews*, 11(11), CD012956. <https://doi.org/10.1002/14651858.CD012956.pub2>

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71

Supplement S2: Pre-defined methods not used in this review:

- If all trials included in a meta-analysis of continuous data would have used the same measurement scale for assessing the outcome, we would have calculated mean differences (MDs) and corresponding 95% confidence intervals (95% CI).
 - We did not calculate any MDs as outcomes had been assessed by use of diverse outcome scales at all analyses.
- For dichotomous outcomes, we would have calculated the risk ratio and corresponding 95% CI with Mantel-Haenszel.
 - However, none of the primary studies of this review reported relevant dichotomous data.”
- In case of symmetrical funnel plots, we would subsequently have done an Egger’s test to investigate further on publication bias
 - The only funnel plot we were able to draw (because of a sufficient number of individual study effect estimates, i.e. ten or more) already indicated publication bias by asymmetry, therefore, an Egger’s test was not necessary.
- If the number of studies available would have permitted, we would have done subgroup analyses with studies of high/low risk of bias (minimum 3 effect estimates per subgroup).
 - We did not perform such subgroup analyses as there were too few studies to build up such subgroups.

Supplement S3: Characteristics of eligible trials that were not available for quantitative analyses, with reasons

Study ID	Country	Sex (% female)	Age (years) Mean	Sample population Comorbidity	Concomitant medication allowed	Setting	Design	Duration	Sample size	Intervention	Comparator	reason for exclusion from quantitative analyses
Amminger 2013 ¹	Austria	93.33	16.2	Unclear. No information.	Antidepressants and benzodiazepines were allowed	Outpatient	Parallel	12 weeks	15	Omega-3 fatty acids	Placebo	Insufficient number of poolable effect estimates per analyses
Hallahan 2007 ²	Ireland	65.31	30.6	Unclear, no information.	Psychotropics as prescribed (if prescribed more than 6 weeks before screening)	Outpatient	Parallel	12 weeks	49	Omega-3 fatty acids	Placebo	Insufficient number of poolable effect estimates per analyses
Maoz 2024 ³	Israel	88.57	29.94	MDD ^a	Yes	Inpatient	Parallel	4 weeks	35	Intranasal oxytocin	Placebo	Insufficient number of poolable effect estimates per analyses
Markovitz 1995a ⁴	US	-	-	Each patient had on average 3 current Axis I diagnoses. A large proportion had comorbid personality disorders besides borderline: 82% self-defeating, 82% paranoid, 71% compulsive, 65% avoidant, 65% dependent, 59% histrionic, 59% passive-aggressive, 53% schizotypal, 35% narcissistic, 35% antisocial.	Unclear, no information.	Inpatient	Parallel	14 weeks	17	Fluoxetine	Placebo	Data reported but unusable for effect size calculation
Moen 2012 ⁵	US	80	35.5	Major depression	Not allowed	Inpatient	Parallel	12 weeks	15	Valproate semisodium	Placebo	No relevant outcomes reported
Montgomery 1982a ⁶	UK	70	35.05	Unclear, no information	Unclear, no information	Outpatient	Parallel	6 months	30	Flupentixol	Placebo	No relevant outcomes reported
Montgomery 1982b ⁶	UK	68.42	35.65	Unclear, no information	Unclear, no information	Outpatient	Parallel	6 months	28	Mianserin	Placebo	No relevant outcomes reported

NCT00533117 ^{7, b}	US	77.3	30.2	Unclear, no information	Benzodiazepines were permitted for sleep	Inpatient	Parallel	12 months	37 ^d	Fluoxetine + DBT	Placebo + DBT	No relevant outcomes reported
NCT04566601 ⁸	international (North and South Americas, Europe, Australia)	86.2	30.2	Current paranoid, schizoid, schizotypal or antisocial PD; lifetime schizophrenia, schizoaffective disorder, schizophreniform disorder, bipolar I disorder, or delusional disorder	Psychotropics not allowed	Outpatient	Parallel	12 weeks	390	BI 1358894	Placebo	Insufficient number of poolable effect estimates per analyses
Nickel 2004 ⁹	Germany	100	26.05	Unclear, no information	Not allowed	Outpatient	Parallel	8 weeks	31	Topiramate	Placebo	No relevant outcomes reported
Nickel 2005 ¹⁰	Germany	0	29.1	mood disorders, somatoform disorders, anxiety disorders and eating disorders	Psychotropics not allowed	Outpatient	Parallel	8 weeks	44	Topiramate	Placebo	No relevant outcomes reported
Kulkarni 2018 ¹¹⁴	Australia	85.29	34.4	Bipolar II disorder	Antipsychotics, Antidepressants, and mood stabilisers allowed (TAU)	Outpatient	Parallel	8 weeks	34	Memantine hydrochloride + TAU	Placebo +TAU	Insufficient number of poolable effect estimates per analyses
Reich 2009 ¹²⁾	US	88.89	31.2	major depression, PTSD, OCD, GAD, panic disorder, social phobia and specific phobia.	Patients could be taking one type of antidepressant but had to have been on a stable dose of that medication for 1 month.	Outpatient	Parallel	12 weeks	27	Lamotrigine	Placebo	No relevant outcomes reported
Rinne 2002 ¹³	The Netherlands	100	29.2	Depression, dysthymia, GAD, PTSD	Not allowed	Outpatient	Parallel	6 weeks	38	Fluvoxamine	Placebo	Data reported but unusable for effect size calculation
Schmahl 2012a ¹⁴	Germany	100	29.2	Unclear, no information	Psychotropics not allowed	In- and outpatient	Cross-over	6 weeks	16	Naltrexone 50mg	Placebo	Insufficient number of poolable effect estimates per analyses
Schmahl 2012b ¹⁴	Germany	100	29.2	pre-existing substance misuse	Psychotropics not allowed	In- and outpatient	Cross-over	6 weeks	16	Naltrexone 200 mg	Placebo	Insufficient number of poolable effect estimates per analyses
Tritt 2005 ¹⁵	Germany	100	29.15	Unclear, no information	Psychotropics not allowed	Outpatient	Parallel	8 weeks	27	Lamotrigine	Placebo	No relevant outcomes reported
Zanarini 2001 ¹⁶	US	100	26.7	Unclear, no information	Psychotropics not allowed	Outpatient	Parallel	24 weeks	28	Olanzapine	Placebo	Data reported but unusable for effect size calculation

Zanarini 2003 ¹⁷	US	100	26.3	Unclear, no information	Psychotropic medication or taking E-EPA supplements	Outpatient	Parallel	8 weeks	30	Omega-3 fatty acids	Placebo (mineral oil)	Insufficient number of poolable effect estimates per analyses
-----------------------------	----	-----	------	-------------------------	---	------------	----------	---------	----	---------------------	-----------------------	---

DBT: Dialectical Behavioural Therapy; DSM: Diagnostic and Statistical Manual of Mental Disorders; E-EPA: ethyl-eicosapentaenoic acid; GAD: Generalised anxiety disorder; OCD: Obsessive-compulsive disorder; PD: Personality disorder; PTSD: Post-traumatic stress disorder

^a Defined comorbidities according to inclusion criteria

^b Trial had 4 arms with an overall sample size of 86. The sample size reported in this table is for the fluoxetine+DBT arm and Placebo+DBT arm only.

^x Major depressive disorder without psychosis was a defined comorbidity for inclusion

References

- 1 Amminger GP, Schäfer MR, Papageorgiou K, Klier CM, Cotton SM, Harrigan M SM, *et al.* Long-chain ω -3 fatty acids for indicated prevention of psychotic disorders: A randomized, placebo-controlled trial. *Arch Gen Psychiatry* 2010; **67**: 146–54.
- 2 Hallahan B, Hibbeln JR, Davis JM, Garland MR. Omega-3 fatty acid supplementation in patients with recurrent self-harm: Single-centre double-blind randomised controlled trial. *Br J Psychiatry* 2007; **190**: 118–22.
- 3 Maoz H, Grossman-Giron A, Sedoff O, Nitzan U, Kashua H, Yarmishin M, *et al.* Intranasal oxytocin as an adjunct treatment among patients with severe major depression with and without comorbid borderline personality disorder. *J Affect Disord* 2024; **347**: 39–44.
- 4 Markovitz P. Pharmacotherapy of impulsivity, aggression, and related disorders. In *Impulsivity and aggression* (ed SDJ Hollander E): 263–87. Wiley, 1995.
- 5 Moen R, Freitag M, Miller M, Lee S, Romine A, Song S. Efficacy of extended-release divalproex combined with ‘condensed’ dialectical behavior therapy for individuals with borderline personality disorder. *Ann Clin Psychiatry* 2012; **24**: 255–60.
- 6 Montgomery SA, Montgomery D. Pharmacological prevention of suicidal behaviour. *J Affect Disord* 1982; **4**: 291–8.
- 7 NCT00533117. Treating Suicidal Behavior and Self-Mutilation in People With Borderline Personality Disorder. <https://clinicaltrials.gov/show/NCT00533117> 2007. (clinicaltrials.gov/ct2/show/results/NCT00533117?sect=X043716&view=results#part).
- 8 Boehringer Ingelheim. A Study to Test Different Doses of BI 1358894 and Find Out Whether They Reduce Symptoms in People With Borderline Personality Disorder. clinicaltrials.gov. 2024. (<https://clinicaltrials.gov/study/NCT04566601?a=14>).
- 9 Nickel MK, Nickel C, Mitterlehner FO, Tritt K, Lahmann C, Leiberich PK, *et al.* Topiramate treatment of aggression in female borderline personality disorder patients: A double-blind, placebo-controlled study. *J Clin Psychiatry* 2004; **65**: 1515–9.
- 10 Nickel MK, Nickel C, Kaplan P, Lahmann C, Mühlbacher M, Tritt K, *et al.* Treatment of aggression with topiramate in male borderline patients: A double-blind, placebo-controlled study. *Biol Psychiatry* 2005; **57**: 495–9.

- 11 Kulkarni J, Thomas N, Hudaib AR, Gavrilidis E, Grigg J, Tan R, *et al.* Effect of the glutamate NMDA receptor antagonist memantine as adjunctive treatment in borderline personality disorder: an exploratory, randomised, double-blind, placebo-controlled trial. *Cent Nerv Syst Drugs* 2018; **32**: 179–87.
- 12 Reich DB, Zanarini MC, Bieri KA. A preliminary study of lamotrigine in the treatment of affective instability in borderline personality disorder. *Int Clin Psychopharmacol* 2009; **24**: 270–5.
- 13 Rinne T, Van den Brink W, Wouters L, Van Dyck R. SSRI treatment of borderline personality disorder: A randomized, placebo-controlled clinical trial for female patients with borderline personality disorder. *Am J Psychiatry* 2002; **159**: 2048–54.
- 14 Schmahl C, Kleindienst N, Limberger M, Ludäscher P, Mauchnik J, Deibler P, *et al.* Evaluation of naltrexone for dissociative symptoms in borderline personality disorder. *Int Clin Psychopharmacol* 2012; **27**: 61–8.
- 15 Tritt K, Nickel C, Lahmann C, Leiberich PK, Rother WK, Loew TH, *et al.* Lamotrigine treatment of aggression in female borderline-patients: A randomized, double-blind, placebo-controlled study. *J Psychopharmacol* 2005; **19**: 287–91.
- 16 Zanarini MC, Frankenburg FR, Parachini EA. A preliminary, randomized trial of fluoxetine, olanzapine, and the olanzapine-fluoxetine combination in women with borderline personality disorder. *J Clin Psychiatry* 2004; **65**: 903–7.
- 17 Zanarini MC, Ed D, Frankenburg FR. Omega-3 Fatty Acid Treatment of Women With Borderline Personality Disorder : *Am J Psychiatry* 2003; **160**: 167–9.

Supplement S4: Reference list of trials included into quantitative analyses

- Black, D. W., Zanarini, M. C., Romine, A., Shaw, M., Allen, J., & Schulz, S. C. C. (2014). Comparison of low and moderate dosages of extended-release quetiapine in borderline personality disorder: a randomized, double-blind, placebo-controlled trial. *American Journal of Psychiatry*, 171(11), 1174–1182. <https://doi.org/10.1176/appi.ajp.2014.13101348>
- Bogenschutz, M. P., & Nurnberg, H. G. (2004). Olanzapine versus placebo in the treatment of borderline personality disorder. *Journal of Clinical Psychiatry*, 65(1), 104–109. <https://doi.org/10.4088/JCP.v65n0118>
- Cowdry, R. W., & Gardner, D. L. (1988). Pharmacotherapy of Borderline Personality Disorder: Alprazolam, Carbamazepine, Trifluoperazine, and Tranylcypromine. *Archives of General Psychiatry*, 45(2), 111–119. <https://doi.org/10.1001/archpsyc.1988.01800260015002>
- Crawford, M. J., Leeson, V. C., Evans, R., Barrett, B., McQuaid, A., Cheshire, J., Sanatinia, R., Lamph, G., Sen, P., Anagnostakis, K., Millard, L., Qurashi, I., Larkin, F., Husain, N., Moran, P., Barnes, T. R. E., Paton, C., Hoare, Z., Picchioni, M., & Gibbon, S. (2022). The clinical effectiveness and cost effectiveness of clozapine for inpatients with severe borderline personality disorder (CALMED study): a randomised placebo-controlled trial. *Therapeutic Advances in Psychopharmacology*, 12, 204512532210908. <https://doi.org/10.1177/20451253221090832>
- Crawford, M. J., Sanatinia, R., Barrett, B., Cunningham, G., Dale, O., Ganguli, P., Lawrence-Smith, G., Leeson, V., Lemonsky, F., Lykomitrou, G., Montgomery, A. A., Morris, R., Munjiza, J., Paton, C., Skorodzien, I., Singh, V., Tan, W., Tyrer, P., & Reilly, J. G. (2018). The clinical effectiveness and cost-effectiveness of lamotrigine in borderline personality disorder: a randomized placebo-controlled trial. *American Journal of Psychiatry*, 175(8), 756–764. <https://doi.org/10.1176/appi.ajp.2018.17091006>
- De la Fuente, J., & Lotstra, F. (1994). A trial of carbamazepine in borderline personality disorder. *European Neuropsychopharmacology*, 4(4), 479–486. [https://doi.org/10.1016/0924-977X\(94\)90296-8](https://doi.org/10.1016/0924-977X(94)90296-8)
- Frankenburg, F. R., & Zanarini, M. C. (2002). Divalproex sodium treatment of women with borderline personality disorder and bipolar II disorder: A double-blind placebo-controlled pilot study. *Journal of Clinical Psychiatry*, 63(5), 442–446. <https://doi.org/10.4088/JCP.v63n0511>
- Goldberg, C. S., Schulz, S. C., Schulz, M. P., Resnick, J. R., Hamer, M. R., Friedel, O. R., Goldberg, S. C., Schulz, S. C., Schulz, P. M., Resnick, R. J., Hamer, R. M., & Friedel, R. O. (1986). Borderline and schizotypal personality disorders treated with low-dose thiothixene vs placebo. *Archives of General Psychiatry*, 43(7), 680–686.
- Grant, J. E., Valle, S., Chesivoir, E., Ehsan, D., & Chamberlain, S. R. (2022). A double-blind placebo-controlled study of brexpiprazole for the treatment of borderline personality disorder. *British Journal of Psychiatry*, 220(2), 58–63. <https://doi.org/10.1192/bjp.2021.159>
- Hollander, E., Allen, A., Lopez, R. P., Bienstock, C. A., Grossman, R., Siever, L. J., Merkat, L., Stein, D. J., & et al. (2001). A preliminary double-blind, placebo-controlled trial of divalproex sodium in borderline personality disorder. *Journal of Clinical Psychiatry*, 62(3), 199–203.

<https://doi.org/10.4088/JCP.v62n0311>

- Linehan, M. M., McDavid, J. P., Brown, M. Z., Sayrs, J. H. R., & Gallop, R. J. (2008). Olanzapine plus dialectical behavior therapy for women with high irritability who Meet criteria for borderline personality disorder: A double-blind, placebo-controlled pilot study. *Journal of Clinical Psychiatry*, 69(6), 999–1005. <https://doi.org/10.4088/JCP.v69n0617>
- Loew, T. H., Nickel, M. K., Muehlbacher, M., Kaplan, P., Nickel, C., Kettler, C., Fartacek, R., Lahmann, C., Buschmann, W., Tritt, K., Bachler, E., Mitterlehner, F., Gil, F. P., Leiberich, P., Rother, W. K., & Egger, C. (2006). Topiramate treatment for women with borderline personality disorder: A double-blind, placebo-controlled study. *Journal of Clinical Psychopharmacology*, 26(1), 61–66. <https://doi.org/10.1097/01.jcp.0000195113.61291.48>
- Nickel, M. K., Muehlbacher, M., Nickel, C., Kettler, C., Gil, F. P., Bachler, E., Buschmann, W., Rother, N., Fartacek, R., Egger, C., Anvar, J., Rother, W. K., Loew, T. H., & Kaplan, P. (2006). Aripiprazole in the treatment of patients with borderline personality disorder: A double-blind, placebo-controlled study. *American Journal of Psychiatry*, 163(5), 833–838. <https://doi.org/10.1176/ajp.2006.163.5.833>
- Pascual, J. C., Soler, J., Puigdemont, D., Pérez-Egea, R., Tiana, T., Alvarez, E., & Pérez, V. (2008). Ziprasidone in the treatment of borderline personality disorder: A double-blind, placebo-controlled, randomized study. *Journal of Clinical Psychiatry*, 69(4), 603–608. <https://doi.org/10.4088/JCP.v69n0412>
- Salzman, C., Wolfson, A. N., Schatzberg, A., Looper, J., Henke, R., Albanese, M., Schwartz, J., & Miyawaki, E. (1995). Effect of fluoxetine on anger in symptomatic volunteers with borderline personality disorder. *Journal of Clinical Psychopharmacology*, 15(1), 23–29. <https://doi.org/10.1097/00004714-199502000-00005>
- Schulz, S. C., Zanarini, M. C., Bateman, A., Bohus, M., Detke, H. C., Trzaskoma, Q., Tanaka, Y., Lin, D., Deberdt, W., & Corya, S. (2008). Olanzapine for the treatment of borderline personality disorder: Variable dose 12-week randomised double-blind placebo-controlled study. *British Journal of Psychiatry*, 193(6), 485–492. <https://doi.org/10.1192/bjp.bp.107.037903>
- Simpson, E. B., Yen, S., Costello, E., Rosen, K., Begin, A., Pistorello, J., & Pearlstein, T. (2004). Combined dialectical behavior therapy and fluoxetine in the treatment of borderline personality disorder. *Journal of Clinical Psychiatry*, 65(3), 379–385. <https://doi.org/10.4088/JCP.v65n0314>
- Soler, J., Pascual, J. C., Campins, J., Barrachina, J., Puigdemont, D., Alvarez, E., & Pérez, V. (2005). Double-blind, placebo-controlled study of dialectical behavior therapy plus olanzapine for borderline personality disorder. *American Journal of Psychiatry*, 162(6), 1221–1224. <https://doi.org/10.1176/appi.ajp.162.6.1221>
- Soloff, P. H., Cornelius, J., George, A., Nathan, S., Perel, J. M., & Ulrich, R. F. (1993). Efficacy of Phenelzine and Haloperidol in Borderline Personality Disorder. *Archives of General Psychiatry*, 50(5), 377–385. <https://doi.org/10.1001/archpsyc.1993.01820170055007>
- Soloff, P. H., George, A., Nathan, R. S., Schulz, P. M., Cornelius, J. R., Herring, J., & Perel, J. M. (1989). Amitriptyline versus haloperidol in borderlines: Final outcomes and predictors of response. *Journal of Clinical Psychopharmacology*, 9(4), 238–246. <https://doi.org/10.1097/00004714-198908000-00002>

- Zanarini, M.C., Schulz, S. C., Detke, H. C., Tanaka, Y., Zhao, F., Lin, D., DeBerdt, W., & Corya, S. (2007). A dose comparison of olanzapine for the treatment of borderline personality disorder: A 12-week randomized double-blind placebo-controlled study. *European Psychiatry*, 22(Suppl 1), S172–S173.
<https://doi.org/10.1016/j.eurpsy.2007.01.565>
- Ziegenhorn AA, Roepke S, Schommer NC, Merkl A, Danker-Hopfe H, Perschel FH, *et al.* (2009). Clonidine improves hyperarousal in borderline personality disorder with or without comorbid posttraumatic stress disorder: a randomized, double-blind, placebo-controlled trial. *J Clin Psychopharmacol*; 29, 170–3.

Supplement S5: Measurement scales of primary studies included into quantitative analyses

Study ID	Measurement scale		
	Depression	Anxiety	Dissociation/psychosis
Black 2014	Montgomery-Åsberg Depression Rating Scale	-	-
Bogenschutz 2004	-	Hamilton Anxiety Rating Scale [‡]	CGI-BPD (transient paranoia or dissociation)
Cowdry 1988	CGI or modified Bunney Hamburg scale*	Modified Bunney-Hamburg Scale Clinician	-
Crawford 2018	Beck Depression Inventory	-	ZAN-BPD-Cognitive disturbance
Crawford 2022	-	-	BPRS
De la Fuente 1994	Hamilton Depression Rating Scale (24 items)	SCL-90 anxiety subscale	SCL-90 R psychoticism
Frankenburg 2002	SCL-90-R Depression	-	-
Goldberg 1986	Hopkins Symptoms check List - Depression	Hopkins Symptom Checklist 90 [‡]	SIB-Suspicious/paranoid
Grant 2022	Hamilton Depression Rating Scale	Hamilton Anxiety Rating Scale [‡]	-
Hollander 2001	Beck Depression Inventory	-	-
Linehan 2008	Hamilton Depression Rating Scale	-	-
Loew 2006	SCL-90-R Depression	SCL-90 anxiety subscale	SCL-90 R psychoticism
Nickel 2006	Hamilton Depression Rating Scale	Hamilton Anxiety Rating Scale	SCL-90 R psychoticism
Pascual 2008	Hamilton Depression Rating Scale (17 items)	Hamilton Anxiety Rating Scale	CGI-BPD (transient paranoia or dissociation)
Salzman 1995	Hamilton Depression Rating Scale	-	-
Schulz 2007	Montgomery-Åsberg Depression Rating Scale	-	ZAN-BPD-paranoid ideation
Simpson 2004	Beck Depression Inventory	State-Trait Anxiety Inventory	Dissociative Experiences Scale
Soler 2005	Hamilton Depression Rating Scale	Hamilton Anxiety Rating Scale	-
Soloff 1989	Beck Depression Inventory	SCL-90 anxiety subscale	SCL-90 R psychoticism
Soloff 1993	Beck Depression Inventory	SCL-90 anxiety subscale	SCL-90 R psychoticism
Zanarini 2007	Montgomery-Åsberg Depression Rating Scale	-	ZAN-BPD-paranoid ideation
Ziegenhorn 2009	Beck Depression Inventory	-	-

BPRS: Brief Psychiatric Rating Scale; CGI-BPD: Clinical Global Impression scale for Borderline Personality Disorder; SCL-90-R: Symptom Check List-90 – Revised; SIB: Schedule of Interviewing Schizotypal Personalities - Borderline score; ZAN-BPD: Zanarini rating scale for borderline personality disorder.

* "At the end of each trial, we obtained ratings of clinical change on seven-point scales similar in concept to the Clinical Global Improvement scale, on which the patient or physician was asked to rate change on each scale "compared to a usual month prior to the start of the study." (Cowdry 1988, p. 11).

‡ Unable to generate summary estimates of reported data.

Supplement S6: Risk of bias graph and summary

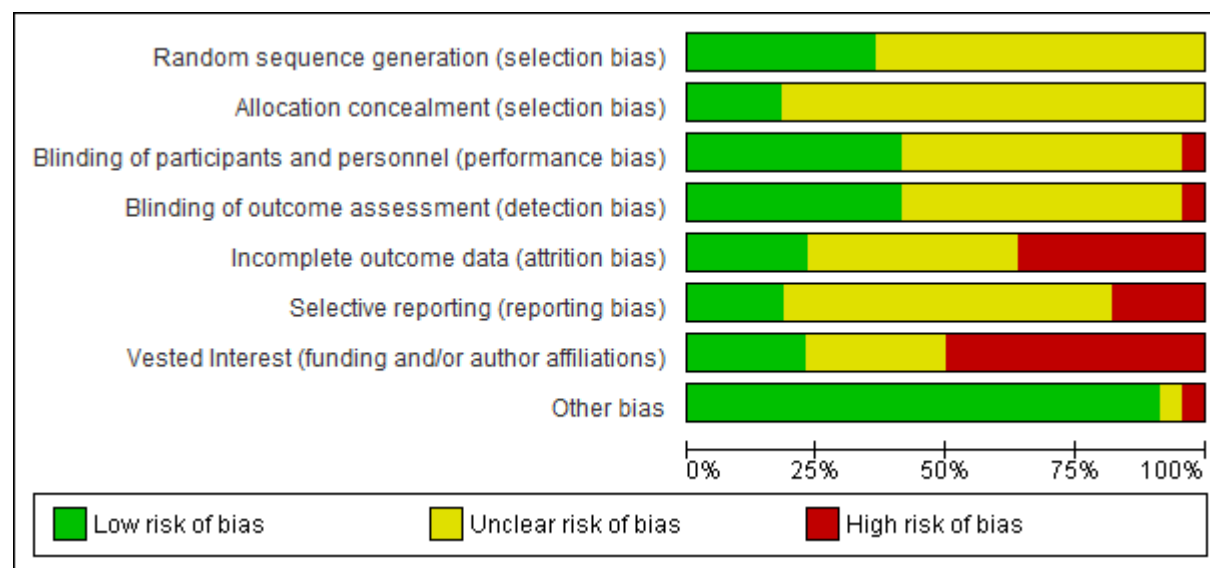


Figure S6.1

Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Vested interest (funding and/or author affiliations)	Other bias
Black 2014	?	?	+	?	-	-	-	+
Bogenschutz 2004	?	?	+	?	?	?	-	+
Cowdry 1988	?	?	+	?	-	-	?	?
Crawford 2018	+	+	+	+	-	+	+	+
Crawford 2022	+	+	-	-	+	-	+	-
De la Fuente 1994	?	?	?	+	?	?	?	+
Frankenburg 2002	+	+	+	?	?	?	-	+
Goldberg 1986	?	?	+	?	+	?	?	+
Grant 2022	+	+	+	?	-	?	-	+
Hollander 2001	?	?	?	+	-	?	-	+
Linehan 2008	+	?	?	+	?	?	-	+
Loew 2006	?	?	?	?	?	?	?	+
Nickel 2006	?	?	+	?	+	?	?	+
Pascual 2008	+	?	?	?	-	+	-	+
Salzman 1995	?	?	?	+	?	?	?	+
Schulz 2007	+	?	?	?	?	-	-	+
Simpson 2004	?	?	?	+	-	?	-	+
Soler 2005	?	?	?	?	?	+	-	+
Soloff 1989	?	?	?	+	+	?	+	+
Soloff 1993	?	?	+	+	+	?	+	+
Zanarini 2007	+	?	?	?	?	+	-	+
Ziegenhorn 2009	?	?	?	+	-	?	+	+

Figure S6.2

Risk of bias summary: review authors' judgements about each risk of bias item for each included trial

Supplement S7a: Effects of antipsychotics on co-occurring psychopathology overall and in subgroups of exclusion criteria for included trials.

Outcomes and subgroups	Trials	N	Statistical method	Effect size*	95% CI	p	I ²	Certainty of evidence (GRADE)	Chi ² Test for subgroup differences (p)
Depressive symptoms	12 ¹⁻¹²	1138	IV, random	SMD -0.22	-0.42 to -0.01	0.04	59%	⊕⊕⊕⊕ Very low ^{e,h}	0.05
PTSD excluded	9 ^{1,2,5-11}	473	IV, random	SMD -0.32	-0.60 to -0.03	0.03	55%	⊕⊕⊕⊕ Low ^e	
PTSD not excluded	3 ^{3,4,12}	665	IV, random	SMD 0.02	-0.15 to 0.18	0.85	13%	⊕⊕⊕⊕ Very low ^{a,e}	
OCD excluded	9 ^{1,2,5-11}	473	IV, random	SMD -0.32	-0.60 to -0.03	0.03	55%	⊕⊕⊕⊕ Low ^e	0.05 [#]
OCD not excluded	3 ^{3,4,12}	665	IV, random	SMD 0.02	-0.15 to 0.18	0.85	13%	⊕⊕⊕⊕ Very low ^{a,e}	
Depression excluded	6 ^{2,6-8,10,11,13}	268	IV, random	SMD -0.40	-0.84 to 0.04	0.07	67%	⊕⊕⊕⊕ Very low ^{b,e,g}	0.14
Depression not excluded	6 ^{1,3-5,9,12}	870	IV, random	SMD -0.05	-0.21 to 0.11	0.54	18%	⊕⊕⊕⊕ Very low ^{a,e}	
Bipolar disorder excluded	9 ^{2-7,9,10,12}	933	IV, random	SMD -0.08	-0.23 to 0.06	0.27	11%	⊕⊕⊕⊕ Low ^e	0.42
Bipolar disorder not excluded	3 ^{1,8,10}	205	IV, random	SMD -0.42	-1.22 to 0.39	0.31	86%	⊕⊕⊕⊕ Very low ^{d,e,g}	
Psychotic disorders excluded	9 ^{1-4,6-9,12}	974	IV, random	SMD -0.28	-0.52 to -0.04	0.02	63%	⊕⊕⊕⊕ Very low ^{e,g}	0.35
Psychotic disorders not excluded	3 ^{5,10,11}	164	IV, random	SMD -0.02	-0.50, 0.45	0.09	58%	⊕⊕⊕⊕ Very low ^{a,f}	
Anxiety symptoms	6 ^{2,3,8-11}	309	IV, random	SMD -0.35	-0.72 to 0.02	0.06	61%	⊕⊕⊕⊕ Very low ^e	0.001
Dissociative symptoms	9 ^{4,5,8-12,14,15}	936	IV, random	SMD -0.28	-0.49 to -0.08	0.007	49%	⊕⊕⊕⊕ Low ^e	
Substance use excluded	6 ^{4,5,9,11,12,14}	799	IV, random	SMD -0.16	-0.30 to -0.02	0.02	0%	⊕⊕⊕⊕ Low ^e	
Substance use not excluded	3 ^{8,10,15}	137	IV, random	SMD -0.79	-1.14 to -0.44	< 0.0001	0%	⊕⊕⊕⊕ Very low ^{c,f}	0.07
Depression excluded	5 ^{4,5,9,12,14}	741	IV, random	SMD -0.16	-0.31 to -0.02	0.03	0%	⊕⊕⊕⊕ Low ^e	
Depression not excluded	4 ^{8,10,11,15}	195	IV, random	SMD -0.59	-1.01 to -0.16	0.007	53%	⊕⊕⊕⊕ Very low ^{c,f,g}	

*Negative MDs or SMDs and RRs >1 indicate beneficial effects of the experimental treatment

IV: Inversed variance; M-H: Mantel-Haenszel; N: Total number of participants; OCD: Obsessive-compulsive disorder; PTSD: Post-traumatic stress disorder; SMD: Standardised mean difference

[#]When including study data from the cross-over study of Ziegenhorn et al. (2009) in a pre-defined secondary sensitivity analyses (s. Cochrane Review¹⁶, the Chi² Test for subgroup differences reaches statistical significance (p=0.04). No substantial changes were observed if including Ziegenhorn et al. (2009) to any of the other relevant comparisons in terms of pooled effect estimates changing the direction of effect or crossing boundaries of confidence intervals.

a downgraded 2 levels due to trivial difference or no difference, with the possibility of difference in both directions

b downgraded 1 level due to a small sample size representing less than 70% but more than 50% of optimal information size (assumed as n≥400)

c downgraded 2 levels due to a small sample size less than 50% of optimal information size (assumed as n≥400)

d downgraded 3 levels due to a small sample size plus CI boundaries suggest very different inferences

e downgraded 2 levels due to the inclusion of studies with a high risk of bias (i.e., several domains with a high risk of bias or most domains with unclear risk of bias)

f downgraded 1 level due to the inclusion of studies with moderate risk of bias (i.e., high risk of bias for a maximum of one domain and/or unclear risk of bias for several domains)

g downgraded 1 level due to substantial heterogeneity (unexplained statistical heterogeneity not explained by clinical or methodological heterogeneity); χ^2 p values <.05 (if certainty had been downgraded for imprecision already due to small sample size, we downgraded only 1 level for inconsistency, even if considerable statistical heterogeneity was present)

h downgraded 1 level due to high risk of publication bias (cf. Fig. 4)

GRADE Working Group grades of evidence

High certainty: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate certainty: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low certainty: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low certainty: We are very uncertain about the estimate.

Supplement S7b: Effects of antidepressants on co-occurring depression, anxiety and dissociation.

Outcomes and subgroups	Trials	N	Statistical method	Effect size*	95% CI	p	I ²	Certainty of evidence (GRADE)	Chi ² Test for subgroup differences (p)
Depressive symptoms	5 ^{2,10,11,17,18}	187	IV, random	SMD -0.37	-0.82 to 0.08	0.11	52%	⊕⊕⊕⊕ Very low ^{a,b}	

Anxiety symptoms	4 ^{2,10,11,18}	164	IV, random	SMD -0.23	-0.58 to 0.12	0.20	17%	⊕⊕⊕⊕ Very low ^{a,b}
Dissociative symptoms	3 ^{10,11,18}	139	IV, random	SMD -0.22	-0.62 to 0.18	0.29	25%	⊕⊕⊕⊕ Very low ^{a,b}

*Negative MDs or SMDs and RRs >1 indicate beneficial effects by the experimental treatment; IV: Inversed variance; M-H: Mantel-Haenszel; N: Total number of participants SMD: Standardised mean difference

a downgraded 2 levels due to a small sample size of less than 50% of optimal information size (assumed as n≥400)

b downgraded 2 levels due to the inclusion of studies with a high risk of bias (i.e., several domains with a high risk of bias or most domains with unclear risk of bias)

GRADE Working Group grades of evidence

High certainty: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate certainty: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low certainty: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low certainty: We are very uncertain about the estimate.

Supplement S7c. Effects of anticonvulsants on co-occurring depression, anxiety and dissociation overall and in subgroups of exclusion criteria for included trials.

Outcomes and subgroups	Trials	N	Statistical method	Effect size*	95% CI	p	I ²	Certainty of evidence (GRADE)	Chi ² test for subgroup differences (p)
Depressive symptoms	6 ^{2,19-23}	344	IV, random	SMD -0.44	-0.80 to -0.08	0.02	46%	⊕⊕⊕⊕ Low ^b	0.42
Depression excluded	3 ¹⁹⁻²¹	65	IV, random	SMD -0.66	-1.19 to -0.12	0.02	0%	⊕⊕⊕⊕ Very low ^{a,b}	
Depression not excluded	3 ^{2,22,23}	279	IV, random	SMD -0.36	-0.86 to 0.15	0.16	66%	⊕⊕⊕⊕ Very low ^{a,b}	
Anxiety symptoms	3 ^{2,8,19}	104	IV, random	SMD -1.11	-1.60 to -0.62	<0.00001	24%	⊕⊕⊕⊕ Very low ^{a,b}	
Dissociative symptoms	3 ^{19,22,23}	270	IV, random	SMD -0.23	-0.66 to 0.20	0.30	51%	⊕⊕⊕⊕ Very low ^{a,b}	

*Negative MDs or SMDs and RRs >1 indicate beneficial effects of the experimental treatment

IV: Inversed variance; M-H: Mantel-Haenszel; N: Total number of participants SMD: Standardised mean difference

a downgraded 2 levels due to a small sample size of less than 50% of optimal information size (assumed as n≥400)

b downgraded 2 levels due to the inclusion of studies with a high risk of bias (i.e., several domains with a high risk of bias or most domains with unclear risk of bias)

GRADE Working Group grades of evidence

High certainty: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate certainty: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low certainty: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low certainty: We are very uncertain about the estimate.

References

- Black DW, Zanarini MC, Romine A, Shaw M, Allen J, Schulz SCC. Comparison of low and moderate dosages of extended-release quetiapine in borderline personality disorder: a randomized, double-blind, placebo-controlled trial. *Am J Psychiatry* 2014; **171**: 1174–82.
- Cowdry RW, Gardner DL. Pharmacotherapy of Borderline Personality Disorder: Alprazolam, Carbamazepine, Trifluoperazine, and Tranylcypromine. *Arch Gen Psychiatry* 1988; **45**: 111–9.
- Soler J, Pascual JC, Campins J, Barrachina J, Puigdemont D, Alvarez E, *et al.* Double-blind, placebo-controlled study of dialectical behavior therapy plus olanzapine for borderline personality disorder. *Am J Psychiatry* 2005; **162**: 1221–4.
- Schulz SC, Zanarini MC, Bateman A, Bohus M, Detke HC, Trzaskoma Q, *et al.* Olanzapine for the treatment of borderline personality disorder: Variable dose 12-week randomised double-blind placebo-controlled study. *Br J Psychiatry* 2008; **193**: 485–92.
- Goldberg CS, Schulz SC, Schulz MP, Resnick JR, Hamer MR, Friedel OR, *et al.* Borderline and schizotypal personality disorders treated with low-dose thiothixene vs placebo. *Arch Gen Psychiatry* 1986; **43**: 680–6.
- Grant JE, Valle S, Chesivoir E, Ehsan D, Chamberlain SR. A double-blind placebo-controlled study of brexpiprazole for the treatment of borderline personality disorder. *Br J Psychiatry* 2022; **220**: 58–63.
- Linehan MM, McDavid JP, Brown MZ, Sayrs JHR, Gallop RJ. Olanzapine plus dialectical behavior therapy for women with high irritability who Meet criteria for borderline personality disorder: A double-blind, placebo-controlled pilot study. *J Clin Psychiatry* 2008; **69**: 999–1005.

- 8 Nickel MK, Muehlbacher M, Nickel C, Kettler C, Gil FP, Bachler E, *et al.* Aripiprazole in the treatment of patients with borderline personality disorder: A double-blind, placebo-controlled study. *Am J Psychiatry* 2006; **163**: 833–8.
- 9 Pascual JC, Soler J, Puigdemont D, Pérez-Egea R, Tiana T, Alvarez E, *et al.* Ziprasidone in the treatment of borderline personality disorder: A double-blind, placebo-controlled, randomized study. *J Clin Psychiatry* 2008; **69**: 603–8.
- 10 Soloff PH, George A, Nathan RS, Schulz PM, Cornelius JR, Herring J, *et al.* Amitriptyline versus haloperidol in borderlines: Final outcomes and predictors of response. *J Clin Psychopharmacol* 1989; **9**: 238–46.
- 11 Soloff PH, Cornelius J, George A, Nathan S, Perel JM, Ulrich RF. Efficacy of Phenelzine and Haloperidol in Borderline Personality Disorder. *Arch Gen Psychiatry* 1993; **50**: 377–85.
- 12 Zanarini MC, Schulz SC, Detke HC, Tanaka Y, Zhao F, Lin D, *et al.* A dose comparison of olanzapine for the treatment of borderline personality disorder: A 12-week randomized double-blind placebo-controlled study. *Eur Psychiatry* 2007; **22**: S172–3.
- 13 Ziegenhorn AA, Roepke S, Schommer NC, Merkl A, Danker-Hopfe H, Perschel FH, *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; **29**: 170–3.
- 14 Bogenschutz MP, Nurnberg HG. Olanzapine versus placebo in the treatment of borderline personality disorder. *J Clin Psychiatry* 2004; **65**: 104–9.
- 15 Crawford MJ, Leeson VC, Evans R, Barrett B, McQuaid A, Cheshire J, *et al.* The clinical effectiveness and cost effectiveness of clozapine for inpatients with severe borderline personality disorder (CALMED study): a randomised placebo-controlled trial. *Ther Adv Psychopharmacol* 2022; **12**: 204512532210908.
- 16 Stoffers-Winterling JM, Storebø OJ, Pereira Ribeiro J, Kongerslev MT, Völm BA, Mattivi JT, *et al.* Pharmacological interventions for people with borderline personality disorder. *Cochrane Database Syst Rev* 2022; **2022**. doi:10.1002/14651858.CD012956.pub2.
- 17 Salzman C, Wolfson AN, Schatzberg A, Looper J, Henke R, Albanese M, *et al.* Effect of fluoxetine on anger in symptomatic volunteers with borderline personality disorder. *J Clin Psychopharmacol* 1995; **15**: 23–9.
- 18 Simpson EB, Yen S, Costello E, Rosen K, Begin A, Pistorello J, *et al.* Combined dialectical behavior therapy and fluoxetine in the treatment of borderline personality disorder. *J Clin Psychiatry* 2004; **65**: 379–85.
- 19 De la Fuente J, Lotstra F. A trial of carbamazepine in borderline personality disorder. *Eur Neuropsychopharmacol* 1994; **4**: 479–86.
- 20 Frankenburg FR, Zanarini MC. Divalproex sodium treatment of women with borderline personality disorder and bipolar II disorder: A double-blind placebo-controlled pilot study. *J Clin Psychiatry* 2002; **63**: 442–6.
- 21 Hollander E, Allen A, Lopez RP, Bienstock CA, Grossman R, Siever LJ, *et al.* A preliminary double-blind, placebo-controlled trial of divalproex sodium in borderline personality disorder. *J Clin Psychiatry* 2001; **62**: 199–203.
- 22 Crawford MJ, Sanatinia R, Barrett B, Cunningham G, Dale O, Ganguli P, *et al.* The clinical effectiveness and cost-effectiveness of lamotrigine in borderline personality disorder: a randomized placebo-controlled trial. *Am J Psychiatry* 2018; **175**: 756–64.
- 23 Loew TH, Nickel MK, Muehlbacher M, Kaplan P, Nickel C, Kettler C, *et al.* Topiramate treatment for women with borderline personality disorder: A double-blind, placebo-controlled study. *J Clin Psychopharmacol* 2006; **26**: 61–6.

Paper 12

CASE REPORT

WILEY

A clinical illustration of short-term mentalization-based therapy for borderline personality disorder

Sophie Juul^{1,2}  | Frederik W. Frandsen¹  | Sune Bo Hansen³  |
Per Sørensen¹ | Anthony Bateman⁴ | Sebastian Simonsen¹ 

¹Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark, Gentofte, Denmark

²Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital, Rigshospitalet, Capital Region of Denmark, Copenhagen, Denmark

³Psychiatric Research Unit, Region Zealand Psychiatry, Slagelse, Denmark

⁴Anna Freud Centre, Kantor Centre for Excellence, London, UK

Correspondence

Sebastian Simonsen, Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark, Gentofte 2820, Denmark.
Email: sebastian.simonsen@regionh.dk

Funding information

Mental Health Services Research Foundation, Capital Region; TrygFonden, Grant/Award Number: 123488

Abstract

Objective: Mentalization-based therapy (MBT) is an evidence-supported, long-term psychotherapy program developed to treat borderline personality disorder (BPD). A short-term, 20-week adaptation to the original MBT format including case formulation, psychoeducation, and group and individual therapy has recently been proposed. The current case material will illustrate how the recent adaptation to the mentalization-based practice can enhance personality functioning using a short-term format.

Methods: Case material is presented to demonstrate the clinical application of short-term MBT in the treatment of a young woman diagnosed with BPD who has a history of failed treatment attempts and who showed signs of affective dysregulation, unstable relationships, and intense abandonment anxiety.

Results: The case illustration shows how short-term MBT can facilitate improvement in personality functioning, specifically targeting situations in which the patient lost her temper and became overwhelmed by abandonment anxiety. By continuously employing therapeutic shifts toward greater autonomy and agency, and by maintaining a balanced empathetic therapeutic stance, the therapists were able to enhance the patients mentalizing and personality functioning.

Conclusions: Short-term MBT can be effectively implemented to enhance the mentalizing capacity and personality functioning in outpatients with BPD.

KEYWORDS

borderline personality disorder, group treatment, psychotherapy

1 | INTRODUCTION

Borderline personality disorder (BPD) is a serious and debilitating psychiatric disorder affecting around 2% of the population (Torgersen et al., 2001; Volkert et al., 2018). BPD is characterized by intense and unstable relationships, fear of abandonment, and identity diffusion and is associated with high rates of self-harming behavior and suicide-related mortality (Black et al., 2004). Today, psychotherapy, particularly long-term psychotherapy, is the most widely used treatment for BPD (Simonsen et al., 2019).

Despite long-term psychotherapy being the most widely used approach for BPD, the optimal treatment duration has been, and still is, continuously debated. Some argue that BPD patients only benefit from long-term psychotherapy because of the complexity and rigidity of their disorder (Leichsenring & Rabung, 2011), while others argue that short-term psychotherapy may be as sufficient as long-term treatment (McMain, 2015; Winston et al., 1994). Furthermore, some argue the need for a stepped-care approach, where patients are offered an initial psychotherapy program and are subsequently stepped up or down, depending on the outcome of the initial phase (Choi-Kain et al., 2016; Paris, 2013). However, to this date, the optimal psychotherapy duration for BPD remains unclear (Cristea et al., 2017; McMain, 2015). According to a recent Cochrane systematic review, long-term (i.e., lasting over 1 year) Dialectical Behavior Therapy (DBT) and Mentalization-Based Therapy (MBT) are considered the two treatment modalities with the best evidence base for patients with BPD (Storebø et al., 2020).

In this article, we will briefly outline the rationale for a short-term group-based MBT treatment for BPD. We will introduce basic clinical MBT concepts and illustrate these ideas as they pertain to the case of Sandra, who was a patient in the short-term MBT program focusing on key ingredients and change processes of the treatment.

1.1 | The theory of mentalizing in clinical practice

MBT is based on the theory of mentalizing, which refers to the process by which we make sense of each other and ourselves, implicitly and explicitly, in terms of subjective states and mental processes (Bateman & Fonagy, 2010). In more plain language, we mentalize when we reflect on what is going on inside our own mind (thoughts, feelings, needs, fantasies, etc.) and that of others. Given the generality of this definition, most psychiatric disorders will inevitably involve some difficulties with mentalizing. However, the key issue is whether mentalizing dysfunction is considered the core of the disorder and if a mentalizing focus is considered to provide an appropriate domain for therapeutic intervention. The theory of mentalizing proposes that anyone can lose the ability to reason about self and others, particularly when emotions run high. In BPD, these disorganized states of mind are more easily triggered by everyday stressors, such as a wrong look from someone or making a simple mistake, and are slower to resolve, especially when attachment processes are active (Bateman & Fonagy, 2010).

When mentalizing is lost, the mind relies on developmentally earlier ways of making sense of the world and these mental processing states are known as prementalizing modes. Prementalizing modes are characterized by (1) quick-fix or short-cut thinking or “outside-in” mind in which only fast and observable actions count, for example,

"If you don't do this, you don't love me" (called "teleological mode" in MBT), (2) "boom brain" or "inside out" experiences in which people lose the ability to tell the difference between internal and external reality, for example, "I feel stupid equals I am stupid" (this is called "psychic equivalence mode" in MBT), and (3) "bubble mode," which describes the feeling of being cut off from the outside world, in extreme cases causing dissociation or depersonalization, for example, "I don't feel real" (called "pretend mode" in MBT). In line with the theoretical assumption that patients with BPD are particularly susceptible to prementalizing modes when experiencing emotional distress, especially in the context of attachment relationships, the therapist's primary aim is to identify these mental shifts and to bring the patient back into a mental state in which the person can reason about and reflect on what is going on. When in a mentalizing mode, understandings of own and others' internal mental states are assumed to facilitate emotional and interpersonal functioning and social interaction (Bateman & Fonagy, 2016).

1.2 | Changing from long-term to short-term mentalization-based therapy

MBT currently has empirical support as an 18-month program for BPD (Storebø et al., 2020). The treatment comprises four essential components: (1) development of a case formulation at the beginning of treatment, which is continuously reassessed throughout; (2) psychoeducation; (3) weekly group therapy; and (4) weekly individual therapy (Bateman & Fonagy, 2016).

The MBT research program originally prescribed a treatment length of 18 months, but measurable change was found to occur after 6 months of treatment (Bateman & Fonagy, 2009). The long duration of the intervention combined with the high prevalence of BPD results in insufficient provision and limited access to evidence-based care. Therefore, an ongoing randomized clinical trial is currently underway, in which the beneficial and harmful effects of short-term compared with long-term MBT are being evaluated for outpatients with subthreshold or diagnosed BPD (Juul et al., 2019).

1.3 | The structure of short-term mentalization-based therapy

The short-term MBT program is very similar to the existing manualized long-term program but differs structurally in the following ways: (1) the length of treatment and the amount of individual psychotherapy, (2) the same therapists provide both group and individual sessions in the short-term program (conjoined psychotherapy), whereas group and individual therapy are provided by different therapists in the long-term program (combined psychotherapy), and (3) the short-term program is structured around closed groups, while the long-term program is structured around slow-open groups.

The short-term MBT program is a 20-week psychotherapy program consisting of five sessions of introductory MBT (MBT-I) followed by 15 sessions of group MBT (MBT-G) accompanied by conjoined individual sessions every second week and two psycho-educative meetings with patients and their relatives. Seven to nine patients and two therapists form each short-term MBT group. The groups are closed to enhance cohesion among the group participants.

Originally, MBT-I was a 12-session introductory psycho-educative program covering relevant topics, such as personality disorders (PDs), attachment, and mentalization (Karterud & Bateman, 2011). The original manual has been modified for our 5-week intervention. A copy of our modified manual is available upon request. After the completion of MBT-I, the same group of participants move on to MBT-G, which consists of 15 sessions of mentalization-based psychotherapy in groups, as manualized by Bateman and Fonagy (2016). In our short-term MBT program, group sessions are accompanied by individual psychotherapy every second week with one of the two group therapists. As part of the individual therapy, a case formulation is prepared and subsequently patients are encouraged to share and discuss the formulation with fellow patients in the group. The overall purpose of the

individual sessions is for the therapist and participant to develop a consensus about the participant's main difficulties and to establish psychotherapeutic focus points for the group therapy. Furthermore, patients are offered three individual follow-up sessions after the end of treatment.

1.4 | Termination in short-term MBT: A critical phase

In time-limited psychotherapy, and perhaps especially in psychotherapy of short duration, the termination phase of treatment inevitably takes on particular importance. Patients with BPD may challenge the termination phase in different ways, which in turn may impact the therapists' adherence to the manual. MBT is not immune to these challenges, but it may offer ways to understand and address them. We have previously proposed a mentalization-based approach to understand and intervene in relation to termination challenges (Juul et al., 2020). In short, we propose (1) that the therapist initiates a dialogue of the feelings associated with the loss of the therapeutic relationship; (2) that the therapist manages the treatment structure carefully, for example, by making the transition to the concluding phase more transparent; and (3) that the original case formulation in which the patient's main goals and mentalizing capacities are outlined at the beginning of treatment (Karterud & Kongerslev, 2019), is extended with a *termination formulation*. A termination formulation is a short document that the patient is asked to fill out at home and return to the therapist, in which the patient re-examines treatment outcomes, addresses how to amend mentalizing failures in the future, writes down future mentalizing goals, and attends to unresolved issues with either the therapists or other group members.

The case report illustrated in this article describes the course of treatment of a patient who received short-term MBT. To preserve the patient's anonymity, several details have been altered.

2 | CASE ILLUSTRATION

2.1 | Presenting problem and client description

Sandra is a 28-year-old Danish woman referred to the outpatient clinic for PDs at Stolpegaard Psychotherapy Centre under the Mental Health Services of the Capital Region of Denmark for specialist treatment of BPD.

Sandra was raised in a both emotionally and physically abusive environment. She lived with her two siblings and both parents until their divorce, which happened when Sandra was 7 years old. The divorce was extremely difficult and complex and left Sandra and her family without an appropriate closure of the difficult issues related to the divorce. The family lived in the suburbs of Copenhagen in what she described as an impoverished environment lacking in resources and stability. Sandra's mother was unemployed and lived on social welfare payments, while Sandra's father was employed as a construction worker. Sandra's father was an alcoholic, but she had positive memories of him from her early childhood. She remembered him as a caring father, and she mainly sought his attention and comfort. After the parents' divorce, her contact with the father diminished, and she lived with her mother and siblings. Sandra describes her mother as unable to provide any sense of security and comfort and with a tendency toward angry outbursts, which often escalated to her physically abusing Sandra in their recurring conflicts during her upbringing. From a mentalizing perspective, we suggest that Sandra's experiences with her parents are characterized by a pattern in which Sandra's emotions were rarely met by attachment figures who were able to contain and mirror her mental states (referred to as "incongruent and unmarked mirroring" in MBT). This may have had lasting effects on Sandra's ability to regulate her emotions and sustain lasting and mutually satisfying interpersonal relations (Fonagy & Luyten, 2009; Simonsen & Bo, 2021). This indicates that in treatment the clinician will need to pay particular attention to empathic validation of Sandra's narratives and experiences.

When Sandra was 10 years old, she was forcibly removed from the home by the social authorities due to significant neglect, violence, and abuse. The social authorities suggested that Sandra should avoid any contact with her family for an extended period, and Sandra was placed into foster care. While Sandra did not perceive her foster parents as abusive but found them supportive, her general experience of these years was overshadowed by the authorities' decision that she should avoid contact with her biological family. Sandra clearly remembers it as very overwhelming and fundamentally incomprehensible to her, at the age of 10 years, that all contact with her family was cut off so abruptly.

Today, Sandra has clear memories of the time when she was forcibly removed from her mother. Despite the emotional abuse she experienced as a small child, this episode stands out in her memory as the most traumatic event of her past. From a mentalization-based developmental perspective, we note that the effect of this separation, and Sandra's experience of it, is not only due to the objective circumstances of the separation itself. Sandra can be seen as particularly vulnerable to handle this experience due to the insecure attachment and impaired mentalizing abilities, which her earlier upbringing has left her with. Sandra now has some contact with both parents, but she still struggles in these relationships due to her father's continued drinking, her mother's sparse emotional availability, and her own mentalizing issues in close attachment relationships.

In her early 20s, Sandra attended individual psychotherapy with a practicing psychologist, to whom she was referred due to problems with her boyfriend and emotional dysregulation. She had ambivalent feelings about this previous treatment attempt. On the one hand, she felt she received help to achieve a better understanding of the psychological impact her early childhood experiences had on her ability to engage in close attachment relationships as an adult. On the other hand, she experienced the treatment as emotionally overwhelming, which left her reluctant to talk about emotionally difficult subjects. In her late 20s, Sandra often had aggressive outbursts at her boyfriend, which caused significant conflict and distress. Sandra had an early history of self-harming (cutting), but currently she is able to resist the impulse to self-harm. Instead, she occasionally consumes alcohol as a means of self-regulation. Based on her persistent interpersonal, emotional, and impulse problems, Sandra was diagnosed with BPD and was referred to two different outpatient treatments for PD under the public mental health service. However, on both occasions, she dropped out of treatment during the early phase due to difficulties with engaging in the group therapy.

In a third treatment attempt, Sandra was referred to a specialist MBT program at the outpatient clinic for PDs at Stolpegaard Psychotherapy Centre. In the initial interviews, Sandra presented herself as having gone through some positive developments before her referral to the clinic, which she found particularly evident in her ability to think less in black-and-white terms and having a better fundamental understanding of her PD diagnosis. However, she still struggled greatly with regulating her emotions, particularly in relation to her mother and to romantic partners. At the time when she was referred to the clinic, she was in a long-term relationship with a man, who in her perception had continuous doubts about their relationship. This often triggered strong emotions of anger, sadness, and intense abandonment anxiety. When overwhelmed with such emotions, Sandra had a history of feeling physically unable to leave her romantic partners, as the fear of ending up alone became too intense. When she was alone, suicidal thoughts often resurfaced, leading her to seek support in the psychiatric emergency ward.

Initially, Sandra presented as being somewhat quiet and shy, but as she increasingly opened up, a more direct and harsh part of her became evident. A tendency to switch from the more avoidant first impression to a more direct, confrontational relational style quickly made the staff realize that she was easily emotionally aroused. Despite Sandra's problems, she was able to function at a moderate level in everyday life. She attended higher education, combining social science and psychology studies, and she was able to maintain a small network of good friends. As part of the routine clinical assessment, Sandra filled out the Level of Personality Functioning Scale—Brief Form 2.0 (LPFS-BF 2.0) (Bach & Hutsebaut, 2018). The total score of the LPFS-BF 2.0 ranges from 0 to 36, with higher scores indicating higher personality pathology. Sandra's total score before treatment was 22 (self-functioning = 12; interpersonal functioning = 10) indicating a moderate level of personality dysfunction. Sandra

especially endorsed items centered on emotional dysregulation and fragile mentalizing, e.g., "My emotions change without me having a grip on them."

After completing the routine clinical assessment at the outpatient clinic, Sandra was referred to the short-term MBT program. As prescribed by the short-term MBT manual, she was assigned an individual therapist, with whom she attended two case formulation sessions before group therapy began. The mentalization-based case formulation for patients with BPD should focus on the primacy of separation distress and rage, the typical overinvolved attachment pattern, as well as the overall liability for mentalizing dysfunction and breakdown of the mentalizing capacity (Karterud & Kongerslev, 2019). The rationale and purpose of the MBT case formulation is to establish a therapeutic alliance and prevent the risk of early dropout by tailoring the mentalization-based therapy program (regardless of length or treatment context) to the individual patient's needs. The case formulation or part of it, especially pertaining to relational aspects, should always be presented by the patient in the group therapy as one way of creating a culture of shared attention to particular problems (called "we-ness" in MBT) (Bateman et al., 2021).

2.2 | Case formulation

The mentalization-based case formulation applied in Sandra's treatment was structured into the following sections: (1) background with a particular focus on attachment experiences; (2) presenting problems with an emphasis on mentalizing difficulties in current close relationships; (3) expected mentalizing difficulties in treatment; (4) focus points and treatment goals. As Sandra's background and presenting problems (Sections 1 and 2) are described earlier in this article, in the following we will focus on presenting Sections 3 and 4 of her case formulation.

2.2.1 | Expected mentalizing difficulties in treatment

Due to Sandra's earlier experience of being emotionally overwhelmed in therapy, she worried how sharing difficult emotions with others in group therapy would affect her. She feared a relapse of her symptoms as a result of talking about her experiences and listening to others in the group. The therapist probed this issue carefully to get a clearer picture of why Sandra experienced this fear. In this process, several important issues surfaced. In terms of a mentalizing focus, it became apparent that Sandra's ability to form accurate mental representations of other people's perceptions of her was often challenged in processes in which she got emotionally involved. In this state, she would oscillate between feeling confused, anxious, and angry and would ultimately conclude that she was somehow "damaged goods." Sandra and her therapist agreed to focus on this fear continuously and to formulate specific goals related to these difficulties. These goals entailed working on Sandra's basic ability to identify, endure, and regulate her emotions in the therapeutic process by approaching her here-and-now experience with curiosity in therapy. In relation to the group, Sandra would need to try to work out how her own emotions (often about herself) might have a negative impact on how she felt she was perceived by others. Sandra told the therapist how confusing her experiences with sharing emotions with other people often became. Sandra: "I am very sensitive, you know. When other people get upset, I go right along with them. I feel what they feel. But then at other times, if somebody is angry about something and telling me about it, I suddenly get paranoid and think the person is really angry with me. You know? It's really confusing."

She also expressed her fear that it might become difficult for her to develop trust and open up in the group, since maintaining a detached position would allow her to protect herself from being entangled in the others' emotions. In the individual sessions, Sandra shared these concerns with her therapist, and rather than seeing them as problems that would lead to her being detached in treatment, they became important goals for treatment in

themselves, as Sandra was helped to regard the group as an arena for training her mentalizing abilities in the here-and-now.

In the first case formulation sessions, Sandra was asked how she expected that the short-term treatment format would affect her, and whether she anticipated that the time format would cause any problems for her. Initially, she was thrilled that the treatment was not longer than 5 months, because that meant the other group members would probably not become too close to her, and it would be easier for her to separate their problems from her own. In other words, Sandra felt it would be easier for her to maintain a detached position in the group, because she knew that the duration would be brief. In the case formulation, the therapist noted this but also described that based on Sandra's relational track record, this detached position might at some point transform into a more overinvolved pattern. Sandra recognized this possibility, and they agreed that it might be something to look out for. After the first 3 months, Sandra and her individual therapist spent some time on reviewing her case formulation together. By this time, Sandra had become more engaged in the group and had begun to find confidence in the other group members. She had also begun to feel more secure in allowing herself to identify and express difficult emotions in the group, and the therapy felt as a "safe space" for her. This positive experience led her to reformulate her answer to the question of how she expected the short-term treatment format to affect her. She now felt an increased sense of anxiety when she thought about the upcoming termination phase and worried that ending this good process would cause her to deteriorate and re-experience painful emotions of separation anxiety, which typically caused her to self-harm.

2.2.2 | Focus points and treatment goals

Apart from the treatment goals related to Sandra's difficulties in the therapeutic process, as described above, Sandra and her therapist agreed on the following focus points and goals for her treatment:

1. Finding greater middle ground in group therapy, becoming neither detached and aloof nor overinvolved. Trying to find this balance was seen as important for avoiding early dropout.
2. Using the treatment context to train her ability to identify and regulate emotions, with a specific focus on
 - Developing a more robust sense of self while with others, especially being aware of hypervigilance about others' mental states about her.
 - Becoming more aware of her tendency to get lost in thoughts and associations by bringing attention to her bodily sensations.
 - Exploring what happens to her when she "freezes" and feels that it is impossible to react differently (especially when she feels abandoned by her romantic partner).
3. Being able to regard herself with greater compassion, as she was often overwhelmed by self-doubt and self-loathing.
4. Being better able to handle her current relationships constructively, especially with her mother and boyfriend. Sandra strongly stated that although she had a hard time being with her mother, maintaining the relationship was crucially important to her. In response, the therapist playfully noted that in stating this so strongly she almost seemed to be placing him in the role of the social authorities that had separated her from her mother as a child. Sandra laughed and remarked, "Don't even try to go there."

2.3 | Course of treatment

After the initial case formulation was completed, Sandra began in group therapy while still seeing her individual therapist every second week throughout, as prescribed by the short-term MBT manual. Consistent with her initial

expectation, she remained quiet and shy during the first couple of group sessions, and only spoke if she was directly spoken to. However, after a few weeks, her individual therapist (who was also one of the group therapists) noticed a change in Sandra's ability to open up and develop a social bond with the group and its individual members. Her therapist brought this up in the group. Sandra was also surprised by the change and agreed that she was now better able to talk in the group about her fear of trusting others and of getting overwhelmed in treatment. Sandra found comfort in the group's ability to understand her and help her further describe herself. This seemed to be happening mainly because of the MBT therapists' typically active and curious approach of focusing attention on particular aspects of Sandra's experiences in the group (Sharp et al., 2020).

Therapist: "Sandra, to me you seem more open and talkative than during the first group sessions. Am I getting that right about you?" Sandra: "Yes, I guess you are right about. I hadn't really thought about it." Therapist: "Tell us more about what that is like for you?" Sandra: "It feels a bit strange somehow. I am used to keeping my 'paranoia' about others to myself, but then when Melanie [another patient] talked about how she felt, it was like wow, that's how I feel. I think that makes me feel it's okay to share what is going on behind my façade."

Talking about her difficulties in the group enabled Sandra to open up more and to talk to the group about her difficulties outside treatment, focusing especially on issues in her romantic relationship. When asked further about what the cause of this positive development might be, Sandra replied that she had needed to observe the interactions in the group to feel safer. She mentioned a specific incident in the group, in which another group member had become emotionally overwhelmed and had to leave the group. One of the therapists then followed that group member, and after half an hour, they returned together. In keeping with the basic MBT model, the patient who returned to the group explained why she left and also how helpful it had been for her that one of the therapist had been present with her and had reassured her that she was not completely "nuts" (called "empathic validation" in MBT). Sandra explained how she could see herself in the other group member, but also that observing her process of calming down and expressing what was going on inside her was something new and worth aspiring to. This experience made Sandra feel safer, as she now allowed herself to experience her emotional states while being present in the group, without the fear of breaking down. Sandra was also able to connect this experience of being cared for by a therapist and being able to express emotions and move on to her own childhood experiences of being left too much to herself without a reassuring other person.

Sandra's work on her romantic relationships, in both group and individual therapy, involved a process of continuously opening up about both her previous relationship, which ended just before treatment began, and her current relationship. Sandra's therapist helped her to identify specific vulnerable incidents that triggered her abandonment anxiety and to undergo a reiterative process of micro-slicing these incidents, focusing on her affective experience in both the here and now and the then and there.

Therapist: "So, you and your boyfriend had another crisis during the weekend. Tell me all about it from your point of view, and we can see if we can make sense of it together." Sandra: "I was supposed to go to this party with some friends on Friday, but it got cancelled due to this bloody Coronavirus. I was pissed about it. My weekend plans gone down the drain, you know?" Therapist: "I think so, but tell me more. You were pissed?" Sandra: "And then I texted Tom [Sandra's boyfriend] about it, and when he did not answer right away I just started crying, and I felt so alone and stupid. And at that point, I was like, I don't even care if he answers me or not. So, when he finally did text me, I was like, whatever!"

By continuously reiterating the process of going through what had happened and asking about Sandra's feelings and how they seemed to shift, Sandra gained a more nuanced and mentalized understanding of her own reaction.

From the therapist perspective, this process was done with an attempt to balance having Sandra think about the situation then-and-there, while also moving her focus to her emotions in the here-and-now, all the while continuously tracking her arousal so she was not over-aroused. As treatment progressed, Sandra was surprised to find that her work in therapy gradually helped her become less reactive and better able to stop and reconsider her current reactions. This was particularly evident in her reactions toward her current boyfriend Tom, but Sandra reported an improved ability to maintain a more curious attitude toward herself and others across many situations—an ability she increasingly saw as connected to the therapeutic process of stopping, identifying her emotional reactions, and considering others' perspectives differently. Sandra's development in therapy became most evident toward the end of treatment when Tom decided to end their relationship. Sandra was able to talk about this in therapy, and while she found help in the group supporting and making sense of her reactions to the breakup, the most important perspective she gained from the group was how well she had actually handled herself. While Sandra obviously had become upset and sad, her abandonment anxiety was not as overwhelming as usual. Sandra described that she had experienced the same urge to stay with her boyfriend when he broke up with her, but she felt more in control of her emotional reaction and was eventually able to leave his apartment and seek comfort from a friend.

Toward the end of treatment, Sandra's presence was highly appreciated by the other group members. She always had relevant topics to bring up when the therapists started the session by doing a round in the group, and she connected with the other group members by validating their experiences and sensitively offering alternative perspectives on their situations. At some point, another group member raised the concern that the therapy was approaching an end at a time when she felt it was just getting started. The group's talk about the upcoming end of therapy occurred just a couple of weeks after Sandra's break-up, and in response to the other group member's concern about the approaching termination she became a little quiet. Sandra was only barely able to express that she felt overwhelmed by thinking of the therapy ending, and she did not speak much more for the rest of the group session.

This incident from the group became the focus of Sandra's next individual session. It became clearer to Sandra that the thought of ending therapy had become overwhelming because she had just become comfortable in the group and had experienced how important the therapeutic process had been for her ability to handle her break-up. The thought of having to deal with life on her own had sparked mixed feelings of abandonment anxiety and hopelessness. Sandra's therapist validated how quickly and independently she was able to make sense of her reaction in the group during the individual session. This move from dependency toward autonomy shifted Sandra's experience from a sense of hopelessness to a more realistic state of mind and increased her sense of agency. She was able to see her feelings about ending therapy as an example of her emotional reaction toward ending relationships more generally. In this way, Sandra was able to change her earlier perception of separation and abandonment as a bitter cocktail of excruciatingly painful self-states marked by shame and self-loathing into realistic sadness, both in relation to her ex-boyfriend as well as toward the termination of the group. This was also part of her written termination formulation, where she also highlighted the importance of making better use of her friends when she experienced separation and abandonment fears as well as the importance of expressing her emotions in the final group and individual sessions.

2.4 | Outcome and prognosis

Sandra ended treatment thinking that she was still vulnerable to reacting with feelings of abandonment anxiety in romantic relationships, but she felt more confident in the sense that experiences with better mentalizing of herself and reaching out to others had taught her that feelings of abandonment and shame were far from inevitable. It was a particularly important experience for her that she had been able to address her conflicted relationship with her boyfriend. They broke off their relationship during the final weeks of her treatment, but Sandra experienced that

she had been able to handle their break-up without the same degree of clinging, self-harm, and suicidal ideation that she had displayed in the past. Experiences with reaching out to the group and to a friend had increased her trust in others as well as her sense of agency: My feelings make sense, if I dare to share them with others. This experience of joint attention to mental states is often referred to as a “we-mode” in MBT and understood as a core ingredient in the process of shifting a fixed mindset (Bateman et al., 2021).

Sandra and her individual therapist developed a mutual understanding that while her earlier experiences with psychotherapy had been helpful for her in developing a better understanding of how her childhood experiences had left her vulnerable to experiences of abandonment, she had not yet gained a sufficient ability to regulate her emotional reactions by the time she was referred to the Stolpegaard MBT clinic. While she had previously gained a better rational understanding of her difficulties, she was still left highly aroused and without a sufficient capacity for emotional regulation. She was not able to properly represent and nuance her emotional experience, which felt too raw and too real (referred to as “psychic equivalence” in MBT). The focus on Sandra's experience in earlier treatment became important as a way to compare how she experienced the process in the current treatment, and it made it clear to Sandra how important it is for her to remain continuously aware of her current affective experience and arousal to mentalize herself properly. She also experienced that when she was able to do this, she tended to be much better at separating what belonged to herself and what belonged to other people. This had a major impact as it enabled her to see that her feeling that other people might reject or dislike her was often more about her own feelings than about how others actually felt. This insecurity was in part linked to childhood experiences of immense despair and guilt in connection with her parents' divorce and the subsequent forced separation. This enabled her to see clearly how certain interpersonal triggers from both past and present were linked to certain mental states that she had previously found incomprehensible and unavoidable. At the end of therapy, her total score on the Level of Personality Functioning Scale—Brief Form 2.0 was 14, which is a score in the low to moderate range of personality dysfunction. Several item responses had changed in desired directions (e.g., “I often feel very vulnerable when relations become more personal”), an increase in attachment security and an improved ability to sustain more stable mentalizing.

3 | CLINICAL PRACTICES AND SUMMARY

In this case report, we have briefly explained the rationale behind a short-term MBT program and provided a case example illustrating core components. Currently, there is no solid empirical basis for providing clear-cut indications and contraindications for different durations of treatment. Bach and Simonsen (2021) provide some tentative ideas about how to align ICD-11 PD severities to different MBT strategies. Although speculative at this time, we suspect that a main feature of patients who may do equally well in short- or long-term treatment would be those at low or moderate PD severity. In other words, patients who suffer from a higher level of severity may need longer time to develop the sense of “we-ness” that we consider essential for a good treatment process.

In the case of Sandra, three individual follow-up sessions were conducted before the short-term MBT program came to an end. In these sessions, both the case formulation and the termination formulation were revisited. In the last session, Sandra acknowledged that there was still work to be done, but she reported notable improvement in terms of the intensity of her emotional reactions, and she felt more confident about her ability to deal with emerging problems in the future.

This statement emphasized the clinical value of delivering MBT in a short-term format. Short-term MBT offers a critical set of specific methods to identify and work with core mentalizing issues, but given the short time frame, it requires structure and focus on the treatment goals in every single session. With no expectation of completely liberating patients from all their pathological personality problems, the treatment model serves as a valuable approach to identifying core mentalizing difficulties and to provide tools for preventing such difficulties from becoming overwhelming in the future. Just as the original long-term MBT format, the short-term format gives the

patient a learning experience of how the establishment of a therapeutic relationship that requires the generation and maintenance of trust can promote this development. However, while terminating this relationship sooner can understandably cause intense feelings of abandonment, the therapist's (and the patient's) delicate the management of such feelings and the shift toward more autonomy may in itself serve as an important therapeutic strategy for enhancing personality functioning and mentalizing over the long term.

ACKNOWLEDGMENTS

This study was funded by research grants from the TrygFonden foundation (grant no. 123488) and from the Mental Health Services Research Foundation, Capital Region of Denmark (grant no. N/A).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Sophie Juul  <http://orcid.org/0000-0002-6171-2904>

Frederik W. Frandsen  <http://orcid.org/0000-0001-6992-0794>

Sune Bo Hansen  <http://orcid.org/0000-0002-0498-2246>

Sebastian Simonsen  <http://orcid.org/0000-0002-5488-7562>

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/jcp.23378>

REFERENCES

- Bach, B., & Hutsebaut, J. (2018). Level of Personality Functioning Scale–Brief Form 2.0: Utility in capturing personality problems in psychiatric outpatients and incarcerated addicts. *Journal of Personality Assessment*, 100(6), 660–670.
- Bach, B., & Simonsen, S. (2021). How does level of personality functioning inform clinical management and treatment? Implications for ICD-11 classification of personality disorder severity. *Current Opinion in Psychiatry*, 34(1), 54–63.
- Bateman, A., Campbell, C., & Fonagy, P. (2021). Rupture and repair in mentalization-based group psychotherapy. *International Journal of Group Psychotherapy*, 71(2), 371–392.
- Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry*, 166(12), 1355–1364.
- Bateman, A., & Fonagy, P. (2010). Mentalization based treatment for borderline personality disorder. *World Psychiatry*, 9(1), 11–15.
- Bateman, A., & Fonagy, P. (2016). *Mentalization based treatment for personality disorders: A practical guide*. Oxford University Press
- Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: Prevalence, risk factors, prediction, and prevention. *Journal of Personality Disorders*, 18(3: Special issue), 226–239.
- Choi-Kain, L. W., Albert, E. B., & Gunderson, J. G. (2016). Evidence-based treatments for borderline personality disorder: Implementation, integration, and stepped care. *Journal of Personality Disorders*, 30(5), 342–356.
- Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder: A systematic review and meta-analysis. *Jama Psychiatry*, 74(4), 319–328.
- Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology*, 21(4), 1355–1381.
- Juul, S., Lunn, S., Poulsen, S., Sørensen, P., Salimi, M., Jakobsen, J. C., & Bateman, A. (2019). Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: A protocol for a randomized clinical trial. *Trials*, 20(1), 1–10.
- Juul, S., Simonsen, S., & Bateman, A. (2020). The capacity to end: Termination of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy*, 50(4), 331–338.
- Karterud, S., & Bateman, A. (2011). *Manual for mentalization based psycho educational group therapy (MBT-I)*. Gyldendal Akademisk.

- Karterud, S., & Kongerslev, M. T. (2019). Case formulations in mentalization-based treatment (MBT) for patients with borderline personality disorder. In U. Kramer (ed.), *Case formulation for personality disorders* (pp. 41–60). Academic Press
- Leichsenring, F., & Rabung, S. (2011). Long-term psychodynamic psychotherapy in complex mental disorders: Update of a meta-analysis. *The British Journal of Psychiatry*, 199(1), 15–22.
- McMain, S. F. (2015). Advances in the treatment of borderline personality disorder: An introduction to the special issue. *Journal of Clinical Psychology*, 71(8), 741–746.
- Paris, J. (2013). Stepped care: An alternative to routine extended treatment for patients with borderline personality disorder. *Psychiatric Services*, 64(10), 1035–1037.
- Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., & Fonagy, P. (2020). Learning to mentalize: A mediational approach for caregivers and therapists. *Clinical Psychology: Science and Practice*, 27(3), 50–105.
- Simonsen, S., Bateman, A., Bohus, M., Dalewijk, H. J., Doering, S., Kaera, A., Moran, P., Ribaudi, J. S., Taubner, S., Wilberg, T., & Mehlum, L. (2019). European guidelines for personality disorders: Past. *Borderline Personality Disorder and Emotion Dysregulation*, 6(1), 1–10.
- Simonsen, S., & Bo, S. (2021). Awakening to recognize the detrimental effects of social avoidance and the possibility of healing through reconnection. *Journal of Clinical Psychology*, 77(5), 1153–1161.
- Storebø, O. J., Stoffers-Winterling, J. M., Völlm, B. A., Kongerslev, M. T., Mattivi, J. T., Jørgensen, M. S., Simonsen, E., Faltinsen, E., Todorovac, A., Sales, C. P., Callesen, H. E., ... Lieb, K. (2020). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*, 5, CD012955.
- Torgersen, S., Kringlen, E., & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Archives of General Psychiatry*, 58(6), 590–596.
- Volkert, J., Gablonski, T. C., & Rabung, S. (2018). Prevalence of personality disorders in the general adult population in Western countries: Systematic review and meta-analysis. *The British Journal of Psychiatry*, 213(6), 709–715.
- Winston, A., Laikin, M., Pollack, J., Wallner Samstag, L., McCullough, L., & Muran, J. C. (1994). Short-term psychotherapy of personality disorders. *American Journal of Psychiatry*, 151(2), 190–194.

How to cite this article: Juul, S., Frandsen, F. W., Bo, S., Sørensen, P., Bateman, A., & Simonsen, S. (2022). A clinical illustration of short-term mentalization-based therapy for borderline personality disorder. *Journal of Clinical Psychology*, 78, 1567–1578. <https://doi.org/10.1002/jclp.23378>

Paper 13

Article

Adapting Short-Term Mentalization-Based Therapy to ICD-11 Personality Disorder in Adolescents

Sebastian Simonsen ^{1,*} , Emilie Hestbæk ¹  and Sophie Juul ^{1,2}

¹ Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, 2820 Gentofte, Denmark; emilie.hestbaek.jacobsen@regionh.dk (E.H.); sophie.juul@regionh.dk (S.J.)

² Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region of Denmark, Copenhagen University Hospital–Rigshospitalet, 2820 Copenhagen, Denmark

* Correspondence: sebastian.simonsen@regionh.dk

Abstract: Following the introduction of the 11th revision of the International Classification of Diseases (ICD-11), adolescents can now be diagnosed with a personality disorder based on severity ranging from mild to moderate to severe. This dimensional model has potential implications for treatment, as it allows clinicians and researchers to search for effective treatments targeting adolescents at different severity levels rather than offering all patients the same treatment. In this conceptual paper, we propose that the short-term mentalization-based therapy (MBT) program, originally developed to treat adults with borderline personality disorder (BPD), has potential clinical advantages for adolescents with ICD-11 personality disorder at the mild to moderate severity level. The short-term MBT program is a 5-month structured treatment approach including individual therapy, combined psychotherapy with the individual therapist also being one of the group therapists, and closed-group therapy to enhance cohesion and a feeling of security. The purpose of this paper is to make a case for the use of this format, as opposed to the traditional long-term MBT format, for adolescents with BPD. Future research should include large-scale randomized clinical trials powered to assess patient-important outcomes.

Keywords: personality pathology; ICD-11 personality disorder; short-term mentalization-based therapy; adolescents; short-term psychotherapy



Citation: Simonsen, S.; Hestbæk, E.; Juul, S. Adapting Short-Term Mentalization-Based Therapy to ICD-11 Personality Disorder in Adolescents. *Children* **2023**, *10*, 93. <https://doi.org/10.3390/children10010093>

Academic Editor: Matteo Alessio Chiappedi

Received: 21 November 2022

Revised: 23 December 2022

Accepted: 24 December 2022

Published: 2 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The 11th revision of the World Health Organization (WHO)’s International Classification of Diseases (ICD-11) represents a paradigm shift in the diagnosis of personality disorders from a categorical to a dimensional model [1]. When it comes to diagnosing children and adolescents with personality disorders, a developmentally informed position is maintained in ICD-11, as it is stated that “manifestations of personality disturbance tend to appear first in childhood, increase during adolescence [. . .]” [2] (p. 388). However, clinicians are simultaneously urged to exercise caution in assigning the personality disorder diagnosis to children and adolescents. The fact that clinicians can diagnose personality disorders in children and adolescents, but are advised not to, is not a major change from ICD-10 [3]. Instead, the major changes seem to lie both in the ICD-11’s capacity to differentiate different levels of personality disorder severity and in the provision of a new psychological language for describing individual differences in personality functioning applicable across a lifespan. Differentiating between mild, moderate, and severe personality disorder may cause clinicians to be less reluctant to diagnose personality disorder of, e.g., mild severity in adolescents, especially if they can refer these adolescents to relevant and effective treatments targeting aspects of self and interpersonal functioning. The change from symptom categories in ICD-10 to dimensions describing personality function (e.g., identity, affect regulation, perspective-taking, empathy, etc.) in ICD-11 is profound because it grounds personality disorders in a developmental and psychological language

that assumes continuity between normality and pathology rather than relying on arbitrary cut-offs. This change will hopefully reduce the negative effects of labeling and stigmatization and make the personality disorder diagnosis better understood and more acceptable to both the afflicted person and their surroundings, including clinicians, family, and friends. By providing a language grounded in normal development, the ICD-11 diagnostic system moves towards what might be labeled academic psychology, asking questions such as: How does identity develop? How do we learn to regulate emotions and understand one's own and other's minds? Instead of potentially distancing ourselves from the diagnosis by using somewhat alienating labels and descriptors, such as borderline, paranoid, or dependent, we are now forced to regard social context as part of both the cause and the cure in order to help the individual with a personality disorder. For an overview of examples in language for personality disorders between ICD-10 and ICD-11, see Table 1.

Table 1. A new psychological language for personality disorders: Crosswalk examples between ICD-10 diagnostic categorical criteria and potential aspects determining ICD-11 severity.

ICD-10	ICD-11
Tendency to bear grudges persistently, e.g., refusal to forgive insults and injuries or slights (Paranoid PD).	Interpersonal dysfunction (ability to manage conflict in relationships).
Apparent indifference to either praise or criticism (Schizoid PD).	Emotional manifestation (tendency to be emotionally underreactive).
Incapacity to experience guilt or to profit from experience, particularly punishment (Dissocial PD).	Emotional manifestation (range and appropriateness of emotional experience and expression).
Excessive efforts to avoid abandonment (BPD).	Interpersonal dysfunction (ability to develop and maintain close and mutually satisfying relationships).
Shallow and labile affectivity (Histrionic PD).	Emotional manifestations (range and appropriateness of emotional experience).
Perfectionism that interferes with task completion (Anankastic PD).	Cognitive manifestations (appropriate stability and flexibility of belief systems).
Excessive preoccupation with being criticized or rejected in social situations (Avoidant PD).	Interpersonal dysfunction (ability to understand and appreciate others' perspectives).
Limited capacity to make everyday decisions without an excessive amount of advice and reassurance from others (Dependent PD).	Self-dysfunction (capacity for self-direction).

Notes. Abbreviations: PD = personality disorder; BPD = borderline personality disorder.

A dimensional understanding grounded in developmental psychology also provides a new roadmap for intervention science aimed at alleviating suffering, because the model suggests that helping people with the disorder implies creating circumstances and processes around the person similar to those we know are important in normal development (salutogenesis) [4]. We believe that the concept of mentalization has unique relevance for understanding the personality functioning described in the ICD-11 and, furthermore, that mentalization-based therapy (MBT) can be adapted to fit the needs of patients across the entire dimension of personality functioning [5]. Mentalizing is a broad term overlapping with numerous other terms within the field of social cognition. At its core, the term concerns how we make sense of ourselves and others [6]. However, providing MBT to children and adolescents suffering from personality difficulties or personality disorders may require specific modifications of the original MBT models used in adult populations. The purpose of this conceptual paper is to develop the idea of why and how the short-term MBT program [7] can be adapted to fit the needs of adolescents battling personality disorder at a mild to moderate level of severity. In terms of the ICD-11 personality disorder classification, we will not cover all diagnostic requirements but focus on severity/functioning determinants.

2. Mentalization and Personality Functioning

The abandonment of the categorical approach based on symptom categories to a dimensional approach based on severity is a radical shift that carries both risks and opportunities. In the context of diagnosing and treating children and adolescents, one possibility is that the ICD-11 system will be able to capture young people at an earlier stage of disorder and that clinicians then can refer them to psychological treatment with a potential preventive impact. The ICD-11 personality functioning dimension is highly influenced by and congruent with the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) Alternative Model of Personality Disorders (AMPD) [8]. The AMPD defines personality functioning through two domains of self (identity and self-direction) and interpersonal capacities (empathy and intimacy). Bender and colleagues [9] based the content of the AMPD personality functioning dimension on a survey of clinician-rated measures of self and other pathology and acknowledged the concept of mentalization as central to the severity measure. We believe that the centrality of mentalization also holds true for the ICD-11 personality disorder guidelines. Understanding personality disorder severity through the lens of mentalization has important implications, as the manifestations of the disorder are developmental in origin, dynamic, and process-based rather than a reflection of static traits inherent in the individual. This helps position the problems of young people with problems of, for example, identity formation and self-worth, as contextual, and the process of enhancing mentalizing as something that is central to both the afflicted person and their surroundings [10]. Mentalization is a multidimensional concept that can be broken down into a set of polarities (implicit/explicit, self/other, affect/cognition, and internal/external), and determinants of personality disorder severity seem to fit nicely onto the mentalizing poles. As an example, consider the cognitive manifestation “accuracy of situational and interpersonal appraisals, especially under stress” [2] (p. 378). This characteristic may be viewed as a mentalizing problem of other, self, and cognitive polarity. The MBT approach would be to clearly outline situations and relationships in which this problem manifests and target increased mentalizing by interventions aimed at balancing the poles. In most cases when problems with the accuracy of situational and interpersonal appraisals are pervasive, the mentalizing therapist may also recognize a state of psychic equivalence in the person and will work hard to shift this state by focusing on process rather than content [11].

Another potentially attractive feature of the ICD-11 personality classification is the ability to differentiate the severity of disorder, which may help guide clinicians towards more tailored treatments. Thus, for example, in the case of mild personality disorder, under stress, the person would manifest with some distortions in appraisals, while distortions in appraisals would be marked in moderate personality disorder and extreme in severe cases. Although this probably resonates with most clinicians who recognize that there may be large variability in stress reactions, it is uncertain how well clinicians in clinical practice will be able to assess such differences among their patients. Even within research settings where reliable clinical measures have been developed and are used, the clinical utility and implications of differences in severity remain largely untested [12]. To the best of our knowledge, no randomized controlled trial has randomized patients with personality disorders of the same severity level into two distinctly different experimental and control groups, and based on the evidence of a strong common factor across bona fide treatments, it is not obvious that we should expect to find large effect differences across different levels of severity [13]. This knowledge gap in the implications of the ICD-11 for treatment may be even wider when it comes to children and adolescents. In this conceptual paper, we will, thus, primarily review the evidence base for MBT in children and adolescents with borderline personality disorder and focus on findings that may have implications for the treatment of ICD-11 personality disorder severity. We will then present the short-term MBT model, which we have developed for adults [14], and elaborate on how it can be useful and how it may be adapted to children and adolescents. Finally, we offer preliminary suggestions about how this format may or may not fit levels of severity in adolescents.

3. MBT for BPD in Adolescence

MBT has gained considerable popularity as a targeted treatment approach for parents, children, adolescents, and families across many levels of personality disorder severity, as evidenced by its wide use in both clinical and non-clinical settings [15,16]. Jørgensen and colleagues [17] recently systematically reviewed the beneficial and harmful effects of psychotherapies for BPD in adolescents. The systematic review identified two randomized clinical trials specifically assessing the effects of MBT for adolescents with self-harm or BPD. In a trial by Rossouw and Fonagy [18] there was evidence of a beneficial effect of MBT compared with treatment as usual in reducing self-harm and depression in adolescents who self-harmed (70% meeting criteria for BPD). The authors found that an increase in mentalizing, a reduction in attachment avoidance, and an improvement in BPD symptoms and traits accounted for the superior effect of MBT, although this evidence must be considered exploratory only. The MBT intervention lasted 12 months and consisted of weekly individual therapy and monthly family therapy. Treatment as usual was not manualized but based on guidance from the UK National Institute for Health. Importantly, the authors reported that there were no statistical differences between treatments in terms of either duration or the use of modalities. The mean duration of the received MBT intervention was approximately 20 h, and approximately one-third of the participants in the MBT intervention did not receive family therapy. Around 50% completed the full 12 months of treatment, which did not differ significantly from the treatment-as-usual group. Finally, in terms of BPD pathology, a large, significant difference emerged at 12 months with only 33% still meeting the criteria for BPD, while the same was true of 58% in the treatment-as-usual group.

Another trial assessing the effects of MBT for adolescents is the M-GAB trial by Beck and colleagues [19], in which 112 adolescents were randomized to 12 months of MBT in a group compared with treatment as usual. The MBT intervention consisted of three MBT introduction sessions, weekly MBT group sessions (slow-open), six MBT-Parent sessions, and five individual case formulation sessions. Treatment as usual consisted of at least 12 individual supportive sessions (not manualized but adapted to the needs of patients, including counseling, psychoeducation, and crisis management). Although positive changes in BPD symptomatology occurred, the results showed no statistical or clinically relevant differences between the MBT intervention and treatment as usual on primary or secondary outcome measures at either end of treatment or follow-up [20]. In terms of BPD psychopathology, only 29% dropped below the cut-off for BPD in both groups and drop-out rates were high: 25% in the treatment-as-usual group and 45% in the MBT group. The authors speculated that group-based MBT may not be suitable for adolescents with high levels of psychopathology and low social functioning and that, following the staging model suggested by Chanen and colleagues [21], group-based intervention may be more appropriate as an early first-stage intervention for adolescents with low levels of severity, possibly corresponding to ICD-11 PD, mild to moderate severity. Subsequent explorative analyses of drop-out and outcomes from this trial found that lower reflective functioning, i.e., mentalizing, predicted drop-out in the MBT intervention [22] and that a more internalizing profile, as opposed to an externalizing one, predicted better outcomes at a two-year follow-up [23]. Overall, these results have led experts in MBT to suggest that more individual and family therapy should be offered to patients with higher personality disorder severity [19], should be of higher intensity and duration [5], and should include a greater focus on and involvement of the social environment (social-ecological approach) [24–26].

The effects of MBT have also been assessed in observational studies. In a pilot study by Laurensen and colleagues [27], 15 inpatients were assessed before and after MBT for adolescents with BPD. The average treatment length was 11 months. The MBT program consisted of four weekly group therapy sessions, one weekly individual session, art therapy, writing therapy, mentalizing cognitive therapy, and a family therapy session every third week. At 12 months' follow-up, there were significant decreases in symptoms and improved personality functioning. Another small, non-controlled study with 34 patients was performed by Bo and colleagues [28], who reported promising results on several

measures of 12 months of MBT group therapy for adolescents with BPD. In addition to weekly group therapy sessions (34 sessions), the adolescents were offered two individual case-formulation sessions and six MBT psychoeducational group sessions, while parents were offered seven sessions of MBT-Parenting, introducing mentalization, personality disorders, attachment, and self-regulation. The authors reported that for 52% of the adolescents, borderline symptoms had dropped below the clinical cut-off at follow-up. The drop-out rate was 26%. However, the results of observational studies should be interpreted with great caution.

4. Short-Term MBT for BPD in Adults

MBT currently has empirical support as an 18-month program for adults with BPD [29]. The treatment comprises four essential components: (1) the development of a case formulation at the beginning of treatment, which is continuously reassessed throughout treatment, (2) psychoeducation, (3) weekly group therapy, and (4) weekly individual therapy [11]. While the original 18-month MBT program is the approach prescribed by the MBT manual, this duration is rarely available, and the long and costly treatment combined with a highly prevalent disorder results in insufficient access to evidence-based care.

Hence, a randomized clinical trial is currently underway in which the beneficial and harmful effects of short-term (20-week) compared with long-term (14-month) MBT are being evaluated for adult outpatients with subthreshold or diagnosed BPD [14].

The short-term MBT program is a 20-week psychotherapy course consisting of five sessions of introductory MBT (MBT-I) followed by 15 sessions of group MBT (MBT-G) accompanied by conjoint individual sessions every second week and two psychoeducative meetings with patients and relatives. The groups are closed to enhance cohesion and security among group participants, meaning that all patients start and finish their 20-week program together. Originally, MBT-I was a 12-session introductory psychoeducative program covering relevant topics such as personality disorders, attachment, and mentalization. The original manual has been modified for the five-week intervention [30]. After the completion of MBT-I, the same group of participants will move on to MBT-G, consisting of 15 sessions of MBT in groups, as manualized by Bateman and Fonagy [11]. A copy of our modified manual is available upon request. In our short-term MBT program, group sessions will be accompanied by individual psychotherapy every second week with one of the two group therapists. As part of individual therapy, a case formulation will be prepared and subsequently shared by the participants in the group. The overall purpose of the individual sessions is for the therapist and participant to develop a consensus on the participant's main difficulties and to establish psychotherapeutic focus points for the group therapy. Patients are furthermore offered three individual follow-up sessions after the end of treatment [7].

5. Adapting Short-Term MBT to Adolescents with BPD

The short-term MBT format could potentially have four major advantages for the treatment of adolescents with BPD. First, the use of individual therapy throughout the course of treatment could be an important component for supporting mentalizing, regardless of baseline reflective functioning level. In patients with lower levels of reflective functioning, it may help to prevent drop-out, and in patients with higher levels of reflective functioning, it may help patients to feel continuously supported to share sensitive topics with their individual therapist before sharing with the group. Second, having one of the group therapists double as the individual therapist might enhance the feeling of security, as the adolescent patient will only need to form a therapeutic relationship with two primary therapists throughout the course of treatment, and the individual therapist can help facilitate a safe environment for the patient in the group. Although, we also acknowledge that this format may also be considered a potential disadvantage, as the adolescent may fear that information divulged in individual therapy may somehow be used in group therapy. Third, in contrast to previous MBT formats, the short-term format consists of closed groups, which

may facilitate group cohesion and a sense of togetherness and belonging; this latter point may be of particular importance for young patients [24]. Fourth, the short format has the potential to benefit more patients, both from an administrative point of view (more patients will have access to treatment, thus minimizing waitlists) and from a clinical point of view, given that the short timeframe may seem less overwhelming from a youth's perspective. Five months may seem less of an intrusion into the person's everyday life compared with the traditional long-term program.

However, considering the immense importance of social environment and family for adolescent development and personality function [31], we believe that further focus on and inclusion of the most important people in the young person's life will be of the utmost importance. This may simply include offering family therapy, as reported in the Rossouw and Fonagy [18] trial, or it may include one or several network approaches, as described by Bo and colleagues [26] and in the Adolescent Mentalization-Based Integrative Therapy (AMBIT) intervention for the hardest-to-reach adolescents [32]. However, it should also be noted that in the Rossouw and Fonagy trial [18], which arguably has provided the most favorable results, one-third of the patients in the MBT group did not receive the family or network interventions. This could indicate that the importance of providing an individual space for scaffolding and developing mentalizing should not be underrated and that we should not only look at the social system around the adolescents but also more carefully consider the specific psychotherapy processes that are experienced as helpful and foster a more elaborate theory of mind. We suspect that disorganized attachment may have an important role to play in how adolescents can make use of psychotherapy. Recently, Talia and colleagues [33] presented data suggesting that disorganized adult patients display patterns of speech that may hinder the development of epistemic trust in their therapists, which may be an important obstacle to using the therapy process as a vehicle for enhancing a sense of agency and identity. If these results are replicated in adolescents, studying how therapists can intervene would be an important next step in studying the most severe personality-disordered patients described in the ICD-11.

6. Translating MBT to the ICD-11 Personality Disorder Model

Hutsebaut and colleagues [5] have presented a very persuasive staging model describing how BPD should not be viewed as a single categorical disorder but instead as a dimensional marker of personality dysfunction disposing the person to a long list of problems, psychopathologies, and psychosocial impairment. Following Hutsebaut and colleagues, we suggest that the short-term MBT program could be used at what the authors label as the "MBT-early" stage of BPD treatment. In categorical terms, this stage is suggested for adolescents aged 13–15 with three to five positive diagnostic BPD criteria. In ICD-11 dimensional terms, this would generally fit the descriptions of mild to moderate personality disorder. Hutsebaut and colleagues suggest that the treatment should focus on individual therapy but also include family sessions, case management, and medication review. In terms of doses, MBT-early consists of an active treatment phase with 16 weekly individual sessions and 3–4 family sessions followed by a booster phase with four booster sessions over 6 months. There are many similarities between this model and the short-term MBT program for adults. However, they differ in two ways: First, the exclusion of MBT group therapy as a modality, which is only suggested at the most severe level in the model proposed by Hutsebaut and colleagues. Secondly, they differ in the lack of structured psychoeducation, which is suggested mainly as a preventive, subclinical intervention in Hutsebaut and colleagues' MBT-early program. Decisions about the inclusion or exclusion of modalities such as group therapy and psychoeducation at different levels of ICD-11 severity is ultimately an empirical question. However, from a clinical perspective, including both group therapy and psychoeducation seems very justified, perhaps especially at the mild–moderate severity level. First, group therapy may provide exactly the kind of training arena for developing the mentalizing skills that are needed for disturbances at this level of ICD-11 PD. For example, the ICD-11 states that "relationships may be characterized

by dependence and avoidance of conflict by giving in to others, even at some cost to themselves" [2] (p. 379). In the MBT group, this problem would in all probability come to the surface and could be mentalized both in the group and individual therapy. In the short-term MBT program [7], this work is highly facilitated by the fact the individual therapist is also one of the group therapists. Thus, instead of relying solely on the individual young patient's account of interpersonal events, the therapist may support mentalizing by probing different situations in which he or she felt that something important was going on with the person. In many ways, this seems to mimic what an active, engaged, and mentalizing parent does in a family [34]. Second, although the evidence for psychoeducation as an intervention modality is not strong for either adults or adolescents [35], psychoeducation does play a very important role in the general MBT model. It is often described as an initial component in developing epistemic trust, as it involves the therapist providing information or a model of the mind that matches something in the patient's experience and often furthers their understanding of themselves [36]. This may happen through something as simple as providing labels for experiences—for example, that a particular experience is called a "developmental trauma", or explaining that problems with understanding other people may be partly attributed to not always being well-understood by those around you. Furthermore, psychoeducation may be viewed as an example of not just how the mind works, but, from a more general mentalizing perspective, of the development of epistemic trust—how personally and culturally useful information is exchanged from mind to mind [6]. We would argue that psychoeducational information is not just useful in itself—as a modality of therapy, it also, and more importantly, provides a key context for identifying and eventually alleviating problems with epistemic trust.

A major future challenge in matching ICD-11 personality disorder severity with optimal treatments along the full spectrum of disorder is finding valid and reliable ways to demarcate the different levels. This may be even more challenging in younger people, simply due to their shorter life experience. Consider some of the specific examples of severe personality disorder in ICD-11: "The individual is largely unable to set and pursue realistic goals" or "The individual is unwilling or unable to sustain regular work due to lack of interest or effort, poor performance (e.g., failure to complete assignments or perform expected roles, unreliability), interpersonal difficulties, or inappropriate behaviour (e.g., fits of temper, insubordination)" [2] (p. 381). After how many years is it reasonable to assess that such problems are clearly present? At what point should such problems primarily be reduced to a personality problem, as opposed to a contextual or even a societal problem? Clinicians quite rightly hesitate to assign a personality disorder diagnosis to adolescents, but this reluctance may also prevent adolescents from accessing potentially effective treatments at the time when the treatment is most effective [37]. The ICD-11 personality disorder severity descriptions have come a long way in describing personality pathology in a more psychological and non-stigmatizing language, but, as shown by the examples above, certain examples may still be difficult for clinicians to assess, especially in child and adolescent populations. In these cases, and especially in assigning an ICD-11 personality disorder diagnosis of severe degree, the general note used throughout the PD diagnostic guidelines specifying that "the list of examples is not exhaustive and not intended to suggest that all items will be present in any single individual" [2] (p. 381) seems extremely useful. In severe cases, clinicians may focus more on psychological and distress descriptors, which may ultimately help to channel those in need of evidence-supported treatments such as MBT.

7. Conclusions and Future Directions

The short-term MBT program, originally developed to treat adult BPD, can be a potentially effective psychological intervention when adapted to adolescents with ICD-11 personality disorder. Short-term MBT includes individual therapy, uses a combined format in which the group therapist is also the individual therapist, is run in closed groups where patients start and finish the group together, and is shorter than existing psychotherapy

options for adolescents with personality disorders. Together, we speculate that these factors could benefit an adolescent population, as they enhance a feeling of security and cohesion, support mentalizing in a safe environment, and are less overwhelming from a youth's perspective. We speculate that this adaptation could potentially prevent drop-out and enhance outcomes for adolescents with ICD-11 personality disorder. Developing efficient short-term treatment for a diagnosis that traditionally has been suggested to require treatment of a longer duration could potentially have a large impact on public health policies and guidelines.

A recent systematic review assessing the effects of psychotherapies for adolescents with BPD [17] concluded that it is currently difficult to conclude whether psychotherapy, in general, is effective for adolescents with BPD, due to the high risk of bias, high attrition rates, and small trials underpowered to confirm or reject realistic intervention effects on patient-important outcomes. Therefore, there is a need for more high-quality trials with larger samples to identify potentially effective psychological interventions for adolescents with BPD. As we have outlined in this paper, the short-term MBT program can be a promising intervention for adolescents with ICD-11 personality disorder. To test this assumption, we suggest that the next step be a large, randomized clinical trial randomizing adolescents into an adapted short-term MBT program, as outlined in this paper, or a control intervention using a psychological sham intervention. Outcomes would be selected based on what is most important to patients (e.g., symptoms, quality of life, level of functioning) but would also include a comprehensive assessment of harmful effects, including objective outcomes such as suicide, suicide attempts, self-harm, and serious adverse events, and the trial should be adequately powered to assess these outcomes. That way, we could move closer to potentially establishing short-term MBT as an evidence-based intervention for adolescents with personality disorders, which would significantly benefit the care of individual patients in the future. We acknowledge that a major limitation of the ideas presented in this conceptual paper is that the empirical literature was not systematically reviewed. The reader should take this into account when assessing the potential benefits and harms of the ideas presented.

Author Contributions: Conceptualization, S.S. and S.J.; methodology, S.J.; writing—original draft preparation, S.S.; writing—review and editing, S.S.; visualization, S.S., S.J. and E.H.; project administration, E.H.; funding acquisition, S.J. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by a research grant from Trygffonden (grant no. 123488).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bach, B.; Kramer, U.; Doering, S.; di Giacomo, E.; Hutsebaut, J.; Kaera, A.; De Panfilis, C.; Schmahl, C.; Swales, M.; Taubner, S.; et al. The ICD-11 classification of personality disorders: A European perspective on challenges and opportunities. *Bord. Personal. Disord. Emot. Dysregulation* **2022**, *9*, 12. [CrossRef] [PubMed]
2. World Health Organization. ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. 2022. Available online: <https://gcp.network/en/private/icd-11-guidelines> (accessed on 16 September 2022).
3. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*; World Health Organization: Geneva, Switzerland, 1992.
4. Lindström, B.; Eriksson, M. Contextualizing salutogenesis and Antonovsky in public health development. *Health Promot. Int.* **2006**, *21*, 238–244. [CrossRef] [PubMed]
5. Hutsebaut, J.; Debbané, M.; Sharp, C. Designing a range of mentalizing interventions for young people using a clinical staging approach to borderline pathology. *Bord. Personal. Disord. Emot. Dysregulation* **2020**, *7*, 6. [CrossRef] [PubMed]

6. Luyten, P.; Campbell, C.; Allison, E.; Fonagy, P. The mentalizing approach to psychopathology: State of the art and future directions. *Annu. Rev. Clin. Psychol.* **2020**, *16*, 297–325. [[CrossRef](#)] [[PubMed](#)]
7. Juul, S.; Frandsen, F.W.; Hansen, S.B.; Sørensen, P.; Bateman, A.; Simonsen, S. A clinical illustration of short-term mentalization-based therapy for borderline personality disorder. *J. Clin. Psychol.* **2022**, *78*, 1567–1578. [[CrossRef](#)] [[PubMed](#)]
8. Bach, B.; Mulder, R. Empirical foundation of the ICD-11 classification of personality disorders. In *Personality Disorders and Pathology: Integrating Clinical Assessment and Practice in the DSM-5 and ICD-11 Era*; Huprich, S.K., Ed.; American Psychological Association: Washington, DC, USA, 2022; pp. 27–52. [[CrossRef](#)]
9. Bender, D.S.; Morey, L.C.; Skodol, A.E. Toward a Model for Assessing Level of Personality Functioning in DSM–5, Part I: A Review of Theory and Methods. *J. Personal. Assess.* **2011**, *93*, 332–346. [[CrossRef](#)]
10. Fonagy, P.; Campbell, C.; Constantinou, M.; Higgitt, A.; Allison, E.; Luyten, P. Culture and psychopathology: An attempt at reconsidering the role of social learning. *Dev. Psychopathol.* **2021**, *34*, 1205–1220. [[CrossRef](#)]
11. Bateman, A.; Fonagy, P. *Mentalization-Based Treatment for Personality Disorders: A Practical Guide*; Oxford University Press: Oxford, UK, 2016.
12. Bach, B.; Tracy, M. Clinical utility of the alternative model of personality disorders: A 10th year anniversary review. *Personal. Disord. Theory Res. Treat.* **2022**, *13*, 369. [[CrossRef](#)]
13. Zimmerman, M. Should the demonstration of improved patient outcome be necessary to overhaul diagnostic approaches? *Comment Bach Tracy* **2022**, *13*, 387–391. [[CrossRef](#)]
14. Juul, S.; Lunn, S.; Poulsen, S.; Sørensen, P.; Salimi, M.; Jakobsen, J.C.; Bateman, A.; Simonsen, S. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: A protocol for a randomized clinical trial. *Trials* **2019**, *20*, 196. [[CrossRef](#)]
15. Volkert, J.; Taubner, S.; Byrne, G.; Rossouw, T.; Midgley, N. Introduction to mentalization-based approaches for parents, children, youths, and families. *Am. J. Psychother.* **2022**, *75*, 4–11. [[CrossRef](#)] [[PubMed](#)]
16. Midgley, N.; Mortimer, R.; Cirasola, A.; Batra, P.; Kennedy, E. The evidence-base for psychodynamic psychotherapy with children and adolescents: A narrative synthesis. *Front. Psychol.* **2021**, *12*, 662671. [[CrossRef](#)] [[PubMed](#)]
17. Jørgensen, M.S.; Storebø, O.J.; Stoffers-Winterling, J.M.; Faltinsen, E.; Todorovac, A.; Simonsen, E. Psychological therapies for adolescents with borderline personality disorder (BPD) or BPD features—A systematic review of randomized clinical trials with meta-analysis and Trial Sequential Analysis. *PLoS ONE* **2021**, *16*, 0245331. [[CrossRef](#)] [[PubMed](#)]
18. Rossouw, T.I.; Fonagy, P. Mentalization-based treatment for self-harm in adolescents: A randomized controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* **2012**, *51*, 1304–1313. [[CrossRef](#)]
19. Beck, E.; Bo, S.; Jørgensen, M.S.; Gondan, M.; Poulsen, S.; Storebø, O.J.; Fjellerad Andersen, C.; Folmo, E.; Sharp, C.; Pedersen, J.; et al. Mentalization-based treatment in groups for adolescents with borderline personality disorder: A randomized controlled trial. *J. Child Psychol. Psychiatry* **2020**, *61*, 594–604. [[CrossRef](#)] [[PubMed](#)]
20. Jørgensen, M.S.; Storebø, O.J.; Bo, S.; Poulsen, S.; Gondan, M.; Beck, E.; Chanen, A.M.; Bateman, A.; Pedersen, J.; Simonsen, E. Mentalization-based treatment in groups for adolescents with Borderline Personality Disorder: 3-and 12-month follow-up of a randomized controlled trial. *Eur. Child Adolesc. Psychiatry* **2021**, *30*, 699–710. [[CrossRef](#)]
21. Chanen, A.M.; Berk, M.; Thompson, K. Integrating early intervention for borderline personality disorder and mood disorders. *Harv. Rev. Psychiatry* **2016**, *24*, 330–341. [[CrossRef](#)]
22. Jørgensen, M.S.; Bo, S.; Vestergaard, M.; Storebø, O.J.; Sharp, C.; Simonsen, E. Predictors of dropout among adolescents with borderline personality disorder attending mentalization-based group treatment. *Psychother. Res.* **2021**, *31*, 950–961. [[CrossRef](#)]
23. Simonsen, E.; Vestergaard, M.; Storebø, O.J.; Bo, S.; Jørgensen, M.S. Prediction of Treatment Outcome of Adolescents with Borderline Personality Disorder: A 2-Year Follow-Up Study. *J. Personal. Disord.* **2021**, *35*, 111–130. [[CrossRef](#)]
24. Bo, S.; Vilmar, J.W.; Jensen, S.L.; Jørgensen, M.S.; Kongerslev, M.; Lind, M.; Fonagy, P. What works for adolescents with borderline personality disorder: Towards a developmentally informed understanding and structured treatment model. *Curr. Opin. Psychol.* **2021**, *37*, 7–12. [[CrossRef](#)]
25. Jensen, S.L.; Bo, S.; Vilmar, J.W. What is behind the closed door? A case illustration of working with social isolation in adolescents using adaptive mentalization-based integrative treatment (AMBIT). *J. Clin. Psychol.* **2021**, *77*, 1189–1204. [[CrossRef](#)] [[PubMed](#)]
26. Bo, S.; Sharp, C.; Kongerslev, M.T.; Luyten, P.; Fonagy, P. Improving treatment outcomes for adolescents with borderline personality disorder through a socioecological approach. *Bord. Personal. Disord. Emot. Dysregulation* **2022**, preprint. [[CrossRef](#)] [[PubMed](#)]
27. Laurensen, E.M.; Hutsebaut, J.; Feenstra, D.J.; Bales, D.L.; Noom, M.J.; Busschbach, J.J.; Verheul, R.; Luyten, P. Feasibility of mentalization-based treatment for adolescents with borderline symptoms: A pilot study. *Psychotherapy* **2014**, *51*, 159–166. [[CrossRef](#)] [[PubMed](#)]
28. Bo, S.; Sharp, C.; Beck, E.; Pedersen, J.; Gondan, M.; Simonsen, E. First empirical evaluation of outcomes for mentalization-based group therapy for adolescents with BPD. *Personal. Disord. Theory Res. Treat.* **2017**, *8*, 396–401. [[CrossRef](#)]
29. Storebø, O.J.; Stoffers-Winterling, J.M.; Völlm, B.A.; Kongerslev, M.T.; Mattivi, J.T.; Jørgensen, M.S.; Faltinsen, E.; Todorovac, A.; Sales, C.P.; Callesen, H.E.; et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst. Rev.* **2020**, *5*, 1–512. [[CrossRef](#)]
30. Karterud, S.; Bateman, A. *Manual for Mentalization Based Psychoeducational Group Therapy (MBT-I)*; Gyldendal Akademisk: Oslo, Norway, 2011.

31. Skabeikyte-Norkiene, G.; Sharp, C.; Kulesz, P.A.; Barkauskiene, R. Personality pathology in adolescence: Relationship quality with parents and peers as predictors of the level of personality functioning. *Bord. Personal. Disord. Emot. Dysregulation* **2022**, *9*, 1–10. [[CrossRef](#)]
32. Bevington, D.; Fuggle, P.; Fonagy, P.; Target, M.; Asen, E. Innovations in practice: Adolescent mentalization-based integrative therapy (AMBIT)—A new integrated approach to working with the most hard to reach adolescents with severe complex mental health needs. *Child Adolesc. Ment. Health* **2013**, *18*, 46–51. [[CrossRef](#)]
33. Talia, A.; Taubner, S.; Miller-Bottome, M.; Muurholm, S.D.; Winther, A.; Frandsen, F.W.; Harpøth, T.; Onofri, A.; Kongerslev, M.T.; Simonsen, S.; et al. The in-session discourse of unresolved/disorganized psychotherapy patients: An exploratory study of an attachment classification. *Front. Psychol.* **2022**, *13*, 985685. [[CrossRef](#)]
34. Sharp, C.; Shohet, C.; Givon, D.; Penner, F.; Marais, L.; Fonagy, P. Learning to mentalize: A mediational approach for caregivers and therapists. *Clin. Psychol. Sci. Pract.* **2020**, *27*, 50. [[CrossRef](#)]
35. Ridolfi, M.E.; Rossi, R.; Occhialini, G.; Gunderson, J.G. A clinical trial of a psychoeducation group intervention for patients with borderline personality disorder. *J. Clin. Psychiatry* **2019**, *81*, 16069. [[CrossRef](#)]
36. Fonagy, P.; Luyten, P.; Allison, E. Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *J. Pers. Disord.* **2015**, *29*, 575–609. [[CrossRef](#)] [[PubMed](#)]
37. Rameckers, S.A.; Verhoef, R.E.; Grasman, R.P.; Cox, W.R.; van Emmerik, A.A.; Engelmoer, I.M.; Arntz, A. Effectiveness of psychological treatments for borderline personality disorder and predictors of treatment outcomes: A multivariate multilevel Meta-analysis of data from all design types. *J. Clin. Med.* **2021**, *10*, 5622. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Paper 14

External validity of randomised clinical trials including outpatients with borderline personality disorder

Caroline Barkholt Kamp, Psychologist, PhD^{1,2}, Sebastian Simonsen, Psychologist, PhD, Associate Professor^{2,3}, Sophie Juul, Psychologist, PhD, Assistant Professor^{1,2}

Affiliations:

¹ Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark

² Stolpegaard Psychotherapy Centre, Mental Health Services, The Capital Region, Gentofte, Denmark

³ Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Corresponding author:

Caroline Barkholt Kamp: caroline.kamp@ctu.dk

Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital – Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen Ø, Denmark

Email addresses:

Caroline Barkholt Kamp: caroline.kamp@ctu.dk

Sebastian Simonsen: sebastian.simonsen@regionh.dk

Sophie Juul: sophie.juul@ctu.dk

Word count: 925

Keywords: external validity, randomised clinical trials, borderline personality disorder

When treating patients with psychiatric disorders, clinicians and administrators should turn to randomised clinical trials to look for evidence-based interventions to guide clinical decision-making.¹ Randomised clinical trials are generally thought to be the gold standard when evaluating the effects of any healthcare intervention.² High-quality trials are conducted with methodological rigour to limit the risk of systematic and random errors and thereby increase the internal validity. However, the external validity, i.e. the generalisability of a trial, is equally important. The assessment of external validity of trials may be based on several factors, for example how the screening procedure was implemented, the number and type of eligibility criteria, the patient characteristics, how the intervention was implemented, or the number and characteristics of the participants who discontinued the intervention. While these factors provide some insight into the external validity of a trial, a critical aspect is often overlooked: which patients are offered participation in randomised clinical trials? The selection process can introduce substantial problems with external validity, especially if certain patient groups are systematically less likely to be included in a trial, even though they frequently seek treatment in a real-world clinical setting. Hence, an arbitrary participant selection process in clinical trials can cause research participants to differ markedly from clinical populations, which negatively impacts the generalisability of the trial results.

Borderline personality disorder is an example of a prevalent and burdensome psychiatric disorder often treated in outpatient psychotherapy clinics,³ for which the external validity of the evidence-based interventions may be questioned. A previous systematic review found that the total proportion of eligible patients with borderline personality disorder who declined to enter a trial varied between (4.4-49.4%).⁴ Hence, already at the screening stage, there is a risk of selection bias as the patients who consent to participate in trials may potentially differ from patients, who decline to participate. However, the proportion of patients who do not reach the trial screening stage is unclear.

Based on a screening of all the included trials from the most recent Cochrane review of psychological interventions for borderline personality disorder,⁵ we found that none of the included outpatient trials provided information about the time gap from referral to trial inclusion. However, eligible patients may potentially discontinue before they reach the screening and inclusion stage of a trial, particularly if the waiting period for the initial consultation is significant.

We assessed the proportion of outpatients with borderline personality disorder who were offered an initial consultation in the Mental Health Services of the Capital Region of Denmark from 2017 to 2023; the proportion of outpatients who attended the initial consultation; and the proportion of outpatients who subsequently initiated treatment (**Figure 1**). Of the 3,751 patients offered an initial consultation for treatment of borderline personality disorder at an outpatient clinic during this period, 11% did not attend the initial consultation. Only 65% of the patients offered initial consultation were also offered treatment, while 56% initiated treatment. Consequently, we can assume that between 11% and 44% of potentially eligible patients were not informed about the opportunity to participate in a randomised clinical trial, if such trial was implemented at the clinic. Some of these patients may have been rightfully excluded because of failure to confirm the diagnosis. However, as they were offered an initial consultation, we must assume that a large proportion of these patients were potentially eligible.

When assessing the external validity of a trial, it may be relevant to consider the screening and enrolment process, but also to assess the proportion of eligible patients who do not reach the screening stage. Although there is often limited information about the pre-screening phase of trials, such information may offer important insights about the generalisability of trials. Figure 1 illustrates

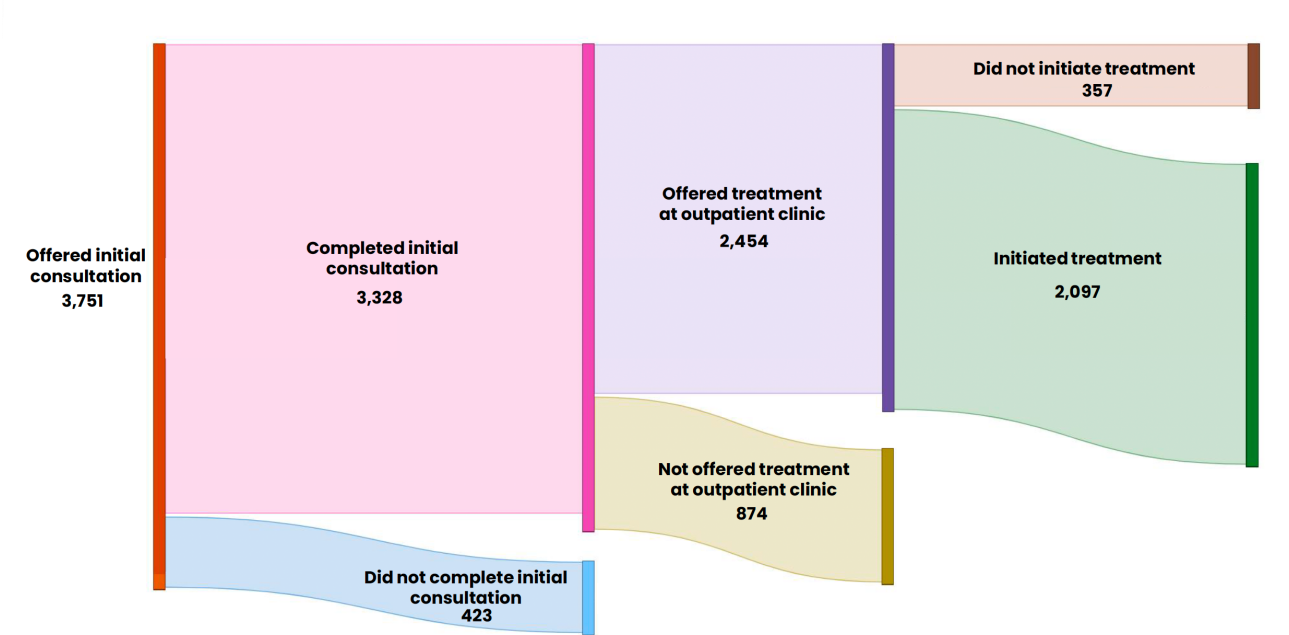
challenges with generalisability that may be invisible in most trial publications of borderline personality disorder: many patients are unintentionally not offered participation in trials, as they discontinue their clinical course before being screened for eligibility in outpatient psychological trials. As a result, we have limited knowledge about whether the evidence-based interventions for borderline personality disorder are effective for this unrepresented patient group. In addition, borderline personality disorder trials often face high dropout rates during the study period,⁶ but this estimate may be even higher if all eligible patients referred to the clinics were included. Ultimately, this raises the question of whether borderline personality disorder trials are generally conducted on a narrow subset of the broader population, resulting in what the GRADE framework would refer to as indirectness of the trial results.⁷

While we still lack information about the comparability of the patients who discontinue their clinical course prior to being screened versus patients enrolled in the trials, we hypothesise that such patients potentially show different clinical characteristics, for example higher levels of impulsivity or clinical or functional instability. When trial participants do not accurately reflect the diversity of the real-world population, the assessed interventions may only be evidence-based for a subgroup of patients. This ultimately limits clinicians' ability to make informed decisions and reduces patients' chances of receiving effective care.

A potential solution could be to implement screening, informed consent collection, baseline assessments and randomisation procedures at the referral stage or right after the diagnosis has been confirmed, instead of waiting for the treatment to commence. Although this approach may present challenges, it could significantly enhance the understanding of external validity of borderline

personality disorder trials. As a minimum, trialists should implement detailed screening logs at the referral stage to properly account for the external validity of the included trial population.

Figure 1: Sankey diagram of referred outpatients with borderline personality disorder in the Mental Health Services of the Capital Region of Denmark from 2017 to 2023



Conflict of Interest Disclosures

The authors report no conflicts of interest.

Funding/Support

No funding was received for this work.

Author Contributions

CBK and SS had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: CBK, SS, SJ. Drafting of the manuscript: CBK, SJ. Critical revision of the manuscript for important intellectual content: CBK, SS, SJ.

Additional Contributions

We thank Delal Yücel for her contributions and assistance with data extraction.

References:

1. Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. *N Eng J Med*. 2020;382(7):674-678.
2. Higgins J, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019). Cochrane, 2019. 2019.
3. Conklin CZ, Westen D. Borderline personality disorder in clinical practice. *Am J Psychiatry*. 2005;162:867-875.
4. Lana F, Fernández-San Martín MI. To what extent are specific psychotherapies for borderline personality disorders efficacious? A systematic review of published randomised controlled trials. *Actas Esp Psiquiatr*. 2013;41(4):242-252.
5. Storebø OJ, Stoffers-Winterling JM, Völlm BA, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2020;(5)
6. Iliakis EA, Ilagan GS, Choi-Kain LW. Dropout rates from psychotherapy trials for borderline personality disorder: A meta-analysis. *Personality Disorders: Theory, Research, and Treatment*. 2021;12(3):193.
7. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol*. Dec 2011;64(12):1303-10. doi:10.1016/j.jclinepi.2011.04.014

Paper 15

Keywords Missing outcome data, Randomised clinical trials, Psychological interventions

Introduction

Although the randomised clinical trial is the gold standard in clinical intervention research, the validity of the trial results may be compromised by missing outcome data. The higher the proportions of missing outcome data in a trial, the more problematic it is to handle statistically, and the more biased the conclusions potentially become [1]. The extent of this potential bias due to missing data depends on the mechanisms causing the data to be missing and the statistical methods applied [2, 3]. Three mechanisms are causing missing data: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) [2, 4, 5]. Data are MCAR when the mechanism causing the missing data depends neither on the observed data nor the missing data [2, 5], for example if paper questionnaires are lost in the mail or blood samples are accidentally damaged in the lab. MCAR causes inflated standard errors due to the reduced sample size but does not cause bias ('systematic error'), as the incomplete dataset is representative of the entire dataset [2]. If the missingness mechanism only depends on the observed data, then the missing data are MAR [2]. MAR allows the prediction of the missing values based on the participants with complete data [2]. If the missingness depends on the missing data, and this dependency remains even given the observed data, then data are classified as MNAR [2, 5]. The MAR and MNAR conditions cannot be distinguished based on the observed data because, by definition, the missing data are unknown and it can, therefore, not be confirmed or rejected if the observed data can predict the unknown data [2, 5].

Trial data should be analysed including data from all randomised participants, i.e., the intention-to-treat principle. Omitting participants who do not have complete data from the analysis is known as complete case (or available case) analysis [1]. When only a few observations are missing, little harm will be done. Still, when many are missing it will lead to a loss of statistical power and bias if the characteristics of the participants with missing data are different from the characteristics of the participants with available data. One alternative approach to complete case analysis is to use single- or multiple imputations, whereby missing values are replaced by some plausible estimated values predicted from the available data [1]. The validity of single imputation depends on specific assumptions that the missing values, for example, are identical to the last observed value [5]. These assumptions are often unrealistic and single imputation is, therefore, a potentially biased method and should be used with caution [5–7]. Using multiple imputation instead,

multiple plausible imputed data sets are created, and results obtained from each are appropriately combined [2, 3]. Multiple imputation is a valid general method for handling missing data in randomised clinical trials, and this method is available for most types of data [2, 6–8]. However, multiple imputation only leads to valid results if data are *not* MNAR, and as mentioned, it can never be determined if data are MAR or MNAR. Other valid statistical methods exist for handling missing data in trials [3], but no completely satisfactory solution for handling data MNAR is available, and missing outcome data are always a threat to the validity of trial results.

The extent of missing data could vary based on the nature of outcomes; for instance, conducting interviews or completing questionnaires is time-consuming and resource-intensive, potentially leading to higher rates of uncollected data. Outcomes that are simpler to evaluate, for example all-cause mortality, which can be retrieved through medical records, may lead to lower levels of missing data [9]. Additionally, specific participant demographics, for example impulsive patients or patients with low levels of functioning, may exhibit higher rates of non-attendance at follow-up assessment [10]. The length of trial interventions could also contribute to increased missing data at follow-up – it might be easier for participants to show up for outcome assessment if the trial period is short. The type of psychological intervention may also contribute to the risk of missing data, for example if participants are dissatisfied with the treatment and prematurely drops out of the whole trial [11].

The present study sought to determine the proportion of missing outcome data in randomised clinical trials of psychological interventions published in six high impact factor psychiatry journals; if missing data patterns differed per types of outcomes, participants, intervention lengths; and psychological intervention types; how missing outcome data were handled in the statistical analyses; and if trialists discussed missing outcome data in the discussion section of the manuscript.

Methods

Search strategy and selection criteria

Two investigators (SJ, PF) independently searched for randomised clinical trials of psychological interventions (for example psychodynamic therapy, cognitive behavioural therapy, or mentalisation-based therapy) for any type of mental health disorder published from 2017 to 2022 (all years inclusive) on the websites of six high-journal impact factor psychiatric journals:

- World Psychiatry.

- Lancet Psychiatry.
- JAMA Psychiatry.
- Psychotherapy and Psychosomatics.
- American Journal of Psychiatry.
- British Journal of Psychiatry.

We chose these peer-reviewed journals due to their high impact factor (ranging from 10.5 to 73.3 in Web of Science for the year 2022). The period was chosen to provide an overview of the current research practice. Two investigators (SJ, PF) independently screened titles and abstracts. Any discrepancies were resolved through discussion or, if required, through discussion with a third investigator (JCJ). Full-texts were retrieved for all trial reports.

Eligibility criteria

We included any randomised clinical trial (as defined by trialists) assessing the effects of a psychological intervention for any mental health disorder, i.e. we included (1) trials comparing psychological interventions with non-psychological interventions (for example drugs, no intervention, or wait-list) and (2) trials comparing two or more psychological interventions with each other. We excluded prevention trials assessing healthy participants.

Data extraction

Six independent investigators (SJ, PF, RKA, CBK, JJP, FS) extracted data and performed risk of bias assessments in pairs. The final data extractions and risk of bias assessments were reached through consensus. Any disagreements following the independent data extractions or risk of bias assessments were resolved through discussion or, if required, through discussion with a third investigator (JCJ). Both the published trial reports and supplementary materials were used for data extraction.

We extracted data from the assessment timepoint defined as primary by the trialists. If no primary timepoint was defined by the trialists, we used the timepoint closest to the end of treatment in both groups. We included feasibility studies but only used the clinical outcome(s).

Assessment of risk of bias due to missing outcome data

For each trial, we assessed domain 3 of the Cochrane Risk of Bias tool – version 2 (RoB2) [12], which is “*Bias due to missing outcome data*”. Low risk of bias was assessed, if outcome data was available for all, or nearly all, participants (usually 95% complete data). High risk of bias was assessed, if missing data exceeded 5%, there was no evidence that the result was not biased, and the missingness could, and was likely to, depend on the true value [12].

Classification of the reporting of the missing data

We classified how the trial reported the extent of the missing data into three categories:

1. ‘Fully reported’: when the trial reported in detail the extent of missing data for the primary outcome.
2. ‘Partially reported’: when the overall proportion of missing data was reported in a flowchart, but the extent of missing data for the primary outcome was unclear.
3. ‘Not reported’: when the extent of missing data was not reported.

Classification of outcomes

We only extracted data for each trial’s primary outcome(s). We classified the different types of primary outcomes as:

1. ‘Hard binary outcomes’: defined as patient-important binary outcomes that are conclusive regarding the disease progression, demonstrating a patient’s feelings, functionality, or survival [13]. Examples of hard binary outcomes are all-cause mortality, suicides, suicide attempts, psychiatric hospitalizations, and self-harm.
2. ‘Symptom severity scales’: defined as any scale (for example interviews or questionnaires) measuring psychiatric symptoms. If a trial assessed a dichotomised version of a continuous outcome (for example response or remission), we also classified this as a symptom severity scale. Examples of symptom severity scales are Hamilton Rating Scale for Depression and Clinical Global Impressions Scale.
3. ‘Count data’: defined as any outcome using countable quantities. Examples of count data are number of binge eating episodes and number of days without drinking alcohol.
4. ‘Other types of outcomes’: any type of outcome not included in the above-mentioned classifications. Examples of other types of outcomes are urine toxicology and other lab results.

Classification of participants

We classified if the trial participants belonged to the following mental health disorder categories: Addiction disorders and comorbidities (for example, substance use disorder and post-traumatic stress disorder; or alcohol use disorder and depression); affective disorders and anxiety disorders; eating disorders and comorbidities; neurodevelopmental disorders; personality disorders; post-traumatic stress disorder; psychosis and

schizophrenia spectrum disorders; sleeping disorders; transdiagnostic; or other disorders.

Classification of intervention length

We classified if the intervention lengths were under 1 month; 1–3 months; 3–6 months; 6–12 months; over 12 months; or if there was no information regarding intervention length.

Classification of psychological intervention types

We classified if the psychological trial interventions (both the experimental and control groups) belonged to the following categories: cognitive and behavioural therapies; humanistic therapy; psychodynamic therapy; supportive psychotherapy; systemic/family therapy.

Classification of statistical analyses

To determine the group of participants included in the analyses, we categorised the included population as:

- 1) 'Intention-to-treat population': if all randomised participants with available data were included in the primary analysis;
- 2) 'Modified intention-to-treat population': if participants who did not initiate the interventions were excluded from the analysis;
- 3) 'Per protocol population': If only participants who completed the intervention and follow-up assessments were analysed;
- 4) 'Wrong classification of the population': if the outcome was reported as being analysed according to the intention-to-treat principle, but the trialists erroneously excluded certain participants from the follow-up assessments and analyses (for example due to protocol violations, adverse effects, or no effects of the intervention). This classification was based on consensus ratings by two review authors (SJ, JJP), as the trialists' reporting of their analyses was often unclear or ambiguous;
- 5) 'Unclear population': if it was unclear or not reported which population was included in the primary analysis.

We assessed if the potential impact of missing data was assessed in sensitivity analyses to evaluate the robustness of the primary analyses in trials, where missing data exceeded 5%. Additionally, we grouped the primary statistical analyses used to handle missing data in the following categories: (1) complete case analysis, (2) multiple imputation or similar methods, (3) other types of imputation (for example last observation carried forward, single imputation, worst-case imputation), (4) full information maximum likelihood, (5) regression analyses or similar methods (e.g. mixed-effects modelling, multilevel linear

regression, linear mixed-effects models), or (6) unclear/not reported.

Additionally, we assessed whether the trialists discussed the potential strengths or limitations of missing outcome data in the discussion section of the published trial report.

Data analysis

We calculated the mean missing data along with 95% confidence intervals (CI). Confidence intervals of proportions were calculated using one sample proportions test with continuity correction and presented as percentages and 95% CI. We assessed the difference between missing data proportions with Fisher's exact test. We used the unadjusted threshold (0.05) as the threshold for statistical significance. The analyses were carried out using R version 4.2.1 (R Core Team, Vienna, Austria).

Results

Characteristics of the included trials

A total of 182 randomised clinical trials of psychological interventions were identified, assessing 233 primary outcomes. Table 1 provides an overview of the included trials. The included trials may be found in Supplementary material 1.

Risk of missing outcome data bias

Of the 233 outcomes, 206 outcomes (88.4%) were assessed at high risk of bias due to missing data. 27 outcomes (11.6%) were assessed as at low risk of bias (Table 2).

Proportion of missing outcome data in randomised clinical trials of psychological interventions

The overall mean percentage of missing data was 18.3% (95% confidence interval: 16.7–20%) for all primary outcomes (Fig. 1). The percentage of missing data was fully reported for 93 outcomes (39.9%), partially reported for 124 outcomes (53.2%), and not reported for 16 outcomes (6.9%) (Table 2). Figure 1 includes data for the fully reported and partially reported outcomes.

Missing outcome data stratified by type of outcome

Fisher's exact test showed that the type of outcome was associated with the proportion of missing outcome data ($p < 0.001$). The percentage of missing data was 18.9% (95% CI: 17.1–20.6%; 180 outcomes) for symptom severity scales; 1.8% (95% CI: 2.3–3.3%; 6 outcomes) for 'hard' binary outcomes (all-cause mortality, self-harm, psychiatric hospitalizations, and out-of-home placement) [14–19]; and 17.8% (95% CI 12.7–22.9%; 27 outcomes) for count data outcomes (Fig. 2a).

Table 1 Trial characteristics

	Overall <i>n</i> = 182
Journal	
American Journal of Psychiatry	18 (9.9%)
British Journal of Psychiatry	27 (14.8%)
JAMA Psychiatry	51 (28.0%)
Lancet Psychiatry	38 (20.9%)
Psychotherapy and Psychosomatics	47 (25.8%)
World Psychiatry	1 (0.5%)
Design	
Cluster RCT	3 (1.6%)
RCT	179 (98.4%)
Number randomized	
median [IQR]	197.50 [120.00;297.75]
mean (95%CI)	288.81 (230.94;346.68)
Type of participants	
Addiction disorders and comorbidities	14 (7.7%)
Affective disorders and anxiety disorders	77 (42.3%)
Eating disorders and comorbidities	9 (4.9%)
Neurodevelopmental disorders	6 (3.3%)
Personality disorders	5 (2.7%)
Post traumatic stress disorder	16 (8.8%)
Psychosis and schizophrenia spectrum disorders	15 (8.2%)
Sleeping disorders	14 (7.7%)
Transdiagnostic	20 (11.0%)
Other	6 (3.3%)
Trial intervention length	
<1 month	9 (4.9%)
>1–3 month	74 (40.7%)
>3–6 month	50 (27.5%)
>6–12 month	29 (15.9%)
>12 month	10 (5.5%)
No information	10 (5.5%)
Method of collecting primary outcomes	
Interview	63 (34.6%)
Questionnaire	104 (57.1%)
Registry	11 (6.0%)
Other	4 (2.2%)
Number of timepoints	
2	37 (20.3%)
3	74 (40.7%)
4	35 (19.2%)
5	19 (10.4%)
6	7 (3.8%)
8	3 (1.6%)
9	2 (1.1%)
10	2 (1.1%)
11	1 (0.5%)
16	1 (0.5%)
17	1 (0.5%)
Number of primary outcomes	
1	145 (79.7%)
2	25 (13.7%)
3	10 (5.5%)

Table 1 (continued)

	Overall <i>n</i> = 182
4	2 (1.1%)
Discussion about missing outcome data in trial report (Yes/No)	
No	92 (50.5%)
Yes	90 (49.5%)

Table 2 Outcome characteristics

	Overall <i>n</i> = 233
Type of primary outcome	
Count data	29 (12.4%)
Hard binary	6 (2.6%)
Scale	196 (84.1%)
Other	2 (0.9%)
Table 2. Statistical analyses in the included trials	
Intention-to-treat	58 (24.9%)
Modified intention-to-treat	12 (5.2%)
Per protocol	16 (6.9%)
Wrong classification	129 (55.4%)
Unclear/Not reported	18 (7.7%)
RoB2 missing data item	
High	206 (88.4%)
Low	27 (11.6%)
Missingness	
Fully reported	93 (39.9%)
Not reported	16 (6.9%)
Partially reported	124 (53.2%)
Proportion of missing data	
0–1%	8 (3.4%)
1–5%	21 (9.0%)
5–10%	28 (12.0%)
10–20%	73 (31.3%)
20–30%	46 (19.7%)
>30%	38 (16.3%)
No information	19 (8.2%)
Reasons for missing data provided?	
No	139 (59.7%)
Yes	94 (40.3%)

Missing outcome data stratified by type of participants

Fisher's exact test showed that the type of diagnosis was associated with the proportion of missing outcome data ($p < 0.001$). The percentage of missing outcome data was 33.1% (95% CI: 22.3–43.9%; 3 outcomes) for participants with personality disorders (all included borderline personality disorder); 19.6% (95% CI 13.2–26.0%; 18 outcomes) for addiction disorders and comorbidities; 18.8% (95% CI 16.4–21.1%; 89 outcomes) for affective disorders and anxiety disorders; 18.7% (95% CI 11.8–25.6%; 8 outcomes) for eating disorders and comorbidities; 13.0% (95% CI -4.3–30.3%; 6 outcomes) for neurodevelopmental disorders; 20.0% (95% CI 15.1–24.9%; 18 outcomes) for post-traumatic stress disorders; 14.8% (95%

CI 9.9–20.7%; 15 outcomes) for psychosis and schizophrenia spectrum disorders; 25.3% (95% CI 18.5–32.2%; 24 outcomes) for sleeping disorders; and 12.0% (95% CI 7.7–16.4%; 25 outcomes) for the transdiagnostic category (Fig. 2c).

Missing outcome data stratified by intervention length

Fisher's exact test showed that intervention lengths was associated with the proportion of missing outcome data ($p < 0.001$). The percentage of missing data was as follows: 14.5% (with a 95% confidence interval [CI] from 3.9 to 25.2%; across 11 outcomes) for interventions shorter than one month; 17.7% (95% CI 15.3–20.2%; covering 98 outcomes) for those lasting one to three months; 20.9% (95% CI 17.4–24.5%; for 50 outcomes) for durations of four to six months; 18.6% (95% CI 14.9–22.4%; across 35 outcomes) for six to twelve months; and 19.5% (95% CI 10.3–28.6%; involving 10 outcomes) for interventions extending beyond twelve months (Fig. 2b).

Missing outcome data stratified by psychological intervention type

Fisher's exact test showed that the type of psychological intervention was associated with the proportion of missing outcome data ($p < 0.001$). The percentage of missing data was 35% (95% CI 34–36%; 176 outcomes) for cognitive and behavioral therapies; 26% (95% CI 21–33%; 2 outcomes) for humanistic therapy; 24% (95% CI 21–28%; 14 outcomes) for psychodynamic therapy; 13% (95% CI 11–15%; 11 outcomes) for supportive psychotherapy; 5% (95% CI 4–7%; 7 outcomes) for systemic/family therapy (primarily due to two trials using 'hard' binary outcomes with little to no missing data) (Fig. 2d).

Missing outcome data stratified by other interventions (e.g. different types of control interventions) are presented in the Supplementary material 2.

Statistical analyses of the outcome data and statistical procedures used to handle missing outcome data

In 58 outcomes (24.9%) an intention-to-treat population was analysed; in 12 outcomes (5.2%) it was a modified intention-to-treat population; in 16 outcomes (6.9%) it was a per protocol population; in 129 outcomes (55.4%) a wrong classification of the population was used; and in 18 outcomes (7.7%) the analysed population was unclear (Table 2).

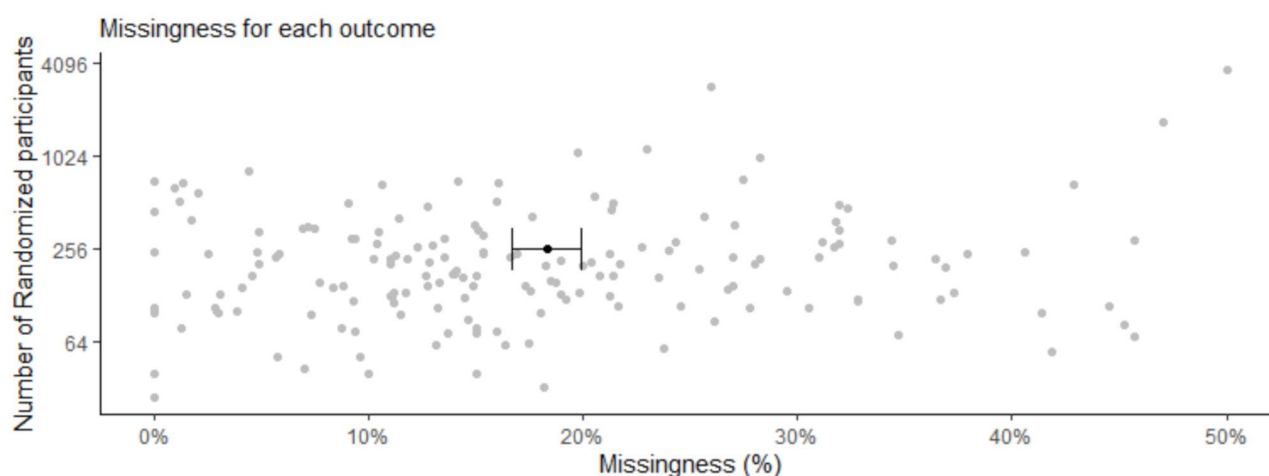


Fig. 1 Total proportion of missing outcome data in randomised clinical trials of psychological interventions

Most trials used either regression analyses (or similar methods) (67 outcomes, 28.8%) or multiple imputation (or similar methods) (69 outcomes, 29.6%) to statistically handle missing outcome data in their primary analysis.

The robustness of the findings was explored with a sensitivity analysis for 48 outcomes (20.6%). The most common sensitivity analysis was multiple imputation (29 outcomes, 12.4%) (Table 3). Statistical analysis methods stratified by sensitivity analysis can be found in Supplementary material 3.

Discussion of missing outcome data in the published trial report

90 trials (49.5%) discussed the strengths or limitations of missing outcome data in their published trial report. The remaining trials did not discuss missing outcome data in the trial report.

Discussion

Missing outcome data is a considerable problem in randomised clinical trials of psychological interventions published in major psychiatry journals. We identified a total of 182 randomised clinical trials assessing 233 primary outcomes and found a mean proportion of missing outcome data of 18.3% across all outcomes. 206 outcome results (88.4%) were assessed at high risk of bias due to missing data.

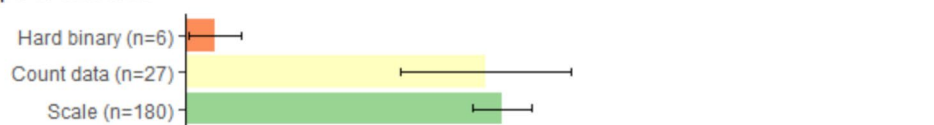
Of the 233 primary outcome, 180 outcomes (77%) were symptom severity scales with a mean percentage of missing data of 18.9%. Six outcomes (2.6%) were hard binary outcomes with a mean percentage of missing data of 1.8%. Trials including participants with borderline personality disorder had the highest proportion of missing data (33.1%) compared with other mental health disorders. The intervention lengths seemed to affect the percentage of missing outcome data, but no clear

missing data pattern could be identified. The percentage of missing data was highest for cognitive and behavioural therapies (35%) compared with other psychological interventions, but this result must be considered exploratory as the number of outcomes assessing different psychological interventions varied a lot (e.g. cognitive and behavioural therapies were assessed in 176 outcomes and humanistic therapy was assessed in two outcomes).

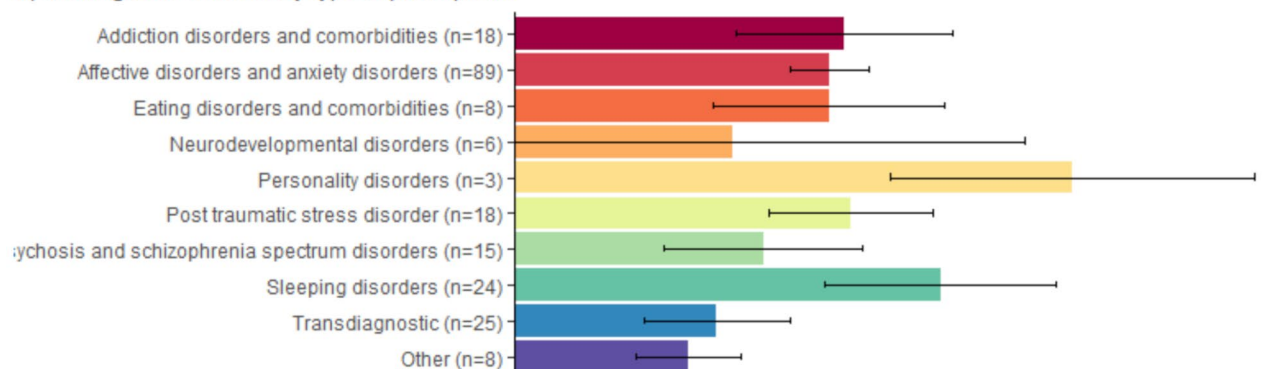
Regression analyses or similar methods (67 outcomes, 28.8%) and multiple imputation or similar methods (69 outcomes, 29.6%) were frequently used to statistically handle missing data in the primary analysis. The robustness of trial results was only explored with a sensitivity analysis for 48 outcomes (20.6%). 90 trials (49.5%) discussed the potential bias risk from missing data in the published trial report.

The results from the present study may inform future outcome selection and measurement in psychological trials to avoid bias arising from missing outcome data. Considering the course rule of thumb that missing data exceeding 5% is considered problematic [3, 12], the total proportion of missing data of 18.3% found in this study must warrant careful concern. Based on our results, it is advisable to incorporate 'hard' binary outcomes into the outcome hierarchy of a trial to minimize missing outcome data. Examples of 'hard' binary outcomes relevant for mental health populations could be all-cause mortality, serious adverse events, suicides, suicide attempts, serious self-harm incidents, employment status, on social welfare or not, or out-of-home placement (for children). These outcomes can be retrieved through medical records or public registries by blinded research personnel, thus minimizing the risk of missing information. Using 'hard' binary outcomes as primary outcomes would often increase the sample size. Still, sample sizes based on hard outcomes in psychological trials are achievable,

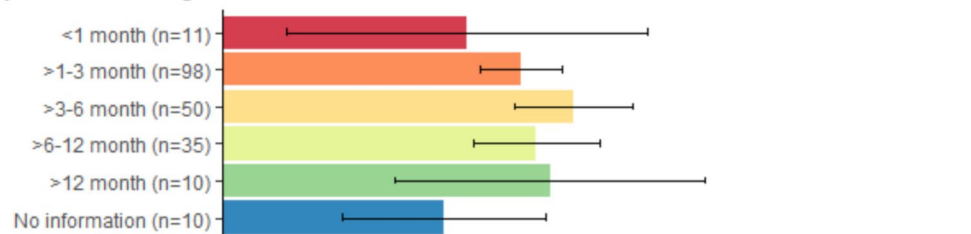
A) Missingness stratified by type of outcome



B) Missingness stratified by type of participants



C) Missingness stratified by intervention length



D) Missingness stratified by type of psychological intervention

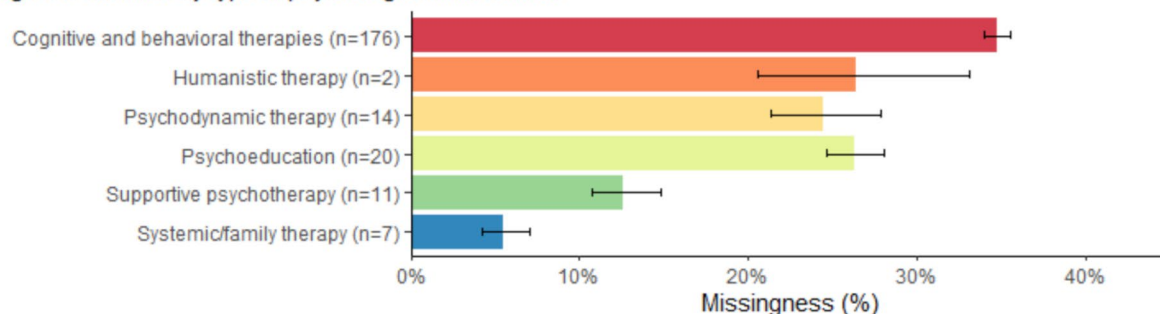


Fig. 2 Missingness stratified by (a) type of outcome; (b) type of participants (c) intervention length, (d) type of psychological intervention

as demonstrated in other areas of medicine [20]. Rather than conducting numerous small trials with flawed methodology, it would be more sensible to focus on fewer trials with larger sample sizes and sound methodology [20].

Our results indicate that patients with personality disorders (borderline personality disorder in particular) had the highest missing data proportion compared with other mental health disorders. This result is similar to a previous systematic review assessing dropout rates from psychotherapy trials for borderline personality disorder, where up to 28.2% dropout rates were found for outpatient settings [21]. We speculate that many patients who drop out of the intervention also do not attend follow-up assessments. The typical reasons for dropout, including

dissatisfaction with treatment, expulsion from treatment, and lack of motivation, may also explain the high missing data proportions found in this study. Recent recommendations for randomised clinical trials in personality disorders recognized missing outcome data as a serious limitation in the field [22, 23].

The present study has limitations. First, if a trial assessed outcomes at various timepoints, we evaluated the proportion of missing data from the trialists designated primary timepoint. If the trialists did not specify a primary timepoint, we evaluated data from the timepoint closest to 'end of treatment'. This methodology was chosen based on the assumption that the trialists' primary focus would align with this timepoint. However, it can be

Table 3 Statistical analysis methods

	Overall n = 233
Primary analysis - method	
Complete case analysis	16 (6.9%)
Full information maximum likelihood	4 (1.7%)
Multiple imputation or similar methods	69 (29.6%)
Other types of imputation (Last observation carried forward)	11 (4.7%)
Other types of imputation (Single imputation)	6 (2.6%)
Other types of imputation (Worst-case imputation)	4 (1.7%)
Regression analyses or similar methods	67 (28.8%)
Unclear/Not reported	56 (24.0%)
Sensitivity analysis	
No	185 (79.4%)
Yes	48 (20.6%)
	Overall n = 48
Type of sensitivity analysis	
Complete case	2 (4.2%)
Last observation carried forward	5 (10.4%)
Multiple imputation	29 (60.4%)
Single imputation	1 (2.1%)
Unclear	11 (22.9%)

argued whether ‘end of treatment’ represents a clinically meaningful timepoint for patients, and that the assessment of beneficial and harmful intervention effects possibly are more patient important at longer-term follow-up. Second, the missingness in trials reporting more than one primary outcome may be correlated, which we did not consider in our analyses. Third, we did not publish a protocol and a statistical analysis plan prior to this retrospective study of published trial reports, as we judged that publishing a protocol would not alter the final results of this review. Fourth, we may have missed important trials with different missing data patterns due to our selection of journals. Journals with high journal impact factors are known to have lower risks of bias compared to journals with low journal impact factor [24, 25]. Hence, there is a possibility that the trials we included may have underestimated the ‘true’ bias arising from missing outcome data, and that the proportions of missing data are in fact even higher. Fifth, we only focused on psychological interventions although trials assessing other interventions (for example pharmacological interventions) for mental health patients may have comparable problems with missing outcome data [26, 27]. Sixth, we explored missing data patterns in types of outcomes, participants, intervention lengths, and psychological intervention types resulting in multiple comparisons. Still, we did not adjust the thresholds for significance according to multiple testing. Hence, the results of our stratified analyses must be considered hypothesis-generating only.

In sum, some recommendations for optimizing data acquisition in psychological trials can be made based on

the results of this study. Trialists should consider incorporating hard binary outcomes into their outcome hierarchy, particularly in trials assessing personality disorders at high risk of missing outcome data. Furthermore, we recommend trialists to optimise their use and reporting of sensitivity analyses, for example by exploring the impact of missing outcome data on the results by testing the missing data mechanisms, i.e. if data seem to be missing under MCAR, MAR, or MNAR. In the present review, the robustness of the findings was only reported as being explored with a sensitivity analysis for 48 outcomes (20.6%) – a result, which is similar to somatic trials [28]. As mentioned, it is impossible to differentiate the MAR and MNAR assumptions. However, a simple way to differentiate the MCAR from MAR and MNAR assumptions could be to create two additional baseline-tables: a baseline table that only includes participants with available outcome data, and another that only includes participants with missing outcome data. By visually inspecting these two baseline tables, it may be possible to judge whether the two tables seem similar, i.e., if data are MCAR or MAR/MNAR.

Conclusion

Missing outcome data is a considerable problem in randomised clinical trials of psychological interventions. Missing outcome data was lowest for hard binary outcomes and highest for symptom severity scales. Participants with borderline personality disorder had the highest proportion of missing data. These results should be considered when designing trials of psychological

interventions. Trialists should improve the use and reporting of sensitivity analyses to explore the robustness of the primary analyses for outcomes at high risk of missing outcome data bias.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12888-024-06263-4>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

Not applicable.

Author contributions

SJ and JCJ initiated the study and wrote the first manuscript draft. SJ and PF screened the journal websites for eligible trials. SJ, PF, RKA, JJP, CBK and FS independently extracted data and performed risk of bias assessments. MHO performed statistical analyses. All authors read, commented on, and approved the final manuscript.

Funding

This study received no specific funding. Some authors receive salaries from the Copenhagen Trial Unit, Centre for Clinical Intervention Research which is funded by the Danish state.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

The analyses in this study were performed on publicly available, published trial data that have already been approved by the respective ethical committees that originally assessed the included trials. Hence, it was not applicable to obtain ethical approval nor participation consent, e.g. in accordance with the Declaration of Helsinki, for this study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

²Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Gentofte, Denmark

³Department of Psychology, University of Copenhagen, Copenhagen, Denmark

⁴Department of Neuroanaesthesiology, The Neuroscience Centre, The Capital Region, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark

⁵Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark

⁶St Joseph's Healthcare Hamilton, Hamilton, ON, Canada

⁷Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada

References

1. Altman DG, Bland JM. Missing data. *BMJ*. 2007;334(7590):424.
2. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338.
3. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.
4. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. 2012;367(14):1355–60.
5. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J Biology Med*. 2013;86(3):343.
6. Zhang Y, Alyass A, Vanniyasingam T, Sadeghirad B, Flórez ID, Pichika SC, et al. A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants with missing outcome data for continuous outcomes in randomized controlled trials. *J Clin Epidemiol*. 2017;88:67–80.
7. Jørgensen AW, Lundstrøm LH, Wetterslev J, Astrup A, Gøtzsche PC. Comparison of results from different imputation techniques for missing data from an anti-obesity drug trial. *PLoS ONE*. 2014;9(11):e111964.
8. Horton NJ, Lipsitz SR. Multiple imputation in practice: comparison of software packages for regression models with missing variables. *Am Stat*. 2001;55(3):244–54.
9. Mendelsohn AB, Dreyer NA, Mattox PW, Su Z, Swenson A, Li R, et al. Characterization of missing data in clinical registry studies. *Ther Innov Reg Sci*. 2014;49(1):146–54.
10. Fenger M, Mortensen EL, Poulsen S, Lau M. No-shows, drop-outs and completers in psychotherapeutic treatment: demographic and clinical predictors in a large sample of non-psychotic patients. *Nord J Psychiatry*. 2011;65(3):183–91.
11. Fernandez E, Salem D, Swift JK, Ramtahal N. Meta-analysis of dropout from cognitive behavioral therapy: magnitude, timing, and moderators. *J Consult Clin Psychol*. 2015;83(6):1108–22.
12. Sterne J, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:14898.
13. Grey P, Grey A, Bolland MJ. Outcomes, interventions and funding in randomised research published in high-impact journals. *Trials*. 2018;19(1):592.
14. Cheryl A, King P, corresponding author1 Alejandra Arango, MS,1 Anne Kramer, LMSW,1 Danielle Busby, PhD,1 Ewa Czyz, PhD,1 Cynthia Ewell Foster, PhD,1 and, Gillespie BW. PhD2, for the YST Study Team. Association of the Youth-Nominated Support Team Intervention for Suicidal Adolescents With 11- to 14-Year Mortality Outcomes. *JAMA Psychiatry*. 2019;76(5):492–8.
15. Tinland A, Loubiere S, Mougeot F, Jouet E, Pontier M, Baumstarck K, et al. Effect of Psychiatric Advance directives facilitated by peer workers on compulsory admission among people with Mental illness: a Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(8):752–9.
16. O'Connor RC, Ferguson E, Scott F, Smyth R, McDaid D, Park AL, et al. A brief psychological intervention to reduce repetition of self-harm in patients admitted to hospital following a suicide attempt: a randomised controlled trial. *Lancet Psychiatry*. 2017;4(6):451–60.
17. Fonagy P, Butler S, Cottrell D, Scott S, Pilling S, Eisler I, et al. Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. *Lancet Psychiatry*. 2018;5(2):119–33.
18. Cottrell DJ, Wright-Hughes A, Collinson M, Boston P, Eisler I, Fortune S, et al. Effectiveness of systemic family therapy versus treatment as usual for young people after self-harm: a pragmatic, phase 3, multicentre, randomised controlled trial. *Lancet Psychiatry*. 2018;5(3):203–16.
19. Gillard S, Bremner S, Patel A, Goldsmith L, Marks J, Foster R, et al. Peer support for discharge from inpatient mental health care versus care as usual in England (ENRICH): a parallel, two-group, individually randomised controlled trial. *Lancet Psychiatry*. 2022;9(2):125–36.
20. Juul S, Siddiqui F, Faltermeier P, Petersen JJ, Kamp CB, Jakobsen RH et al. Less focus on symptom scales in psychiatric trials: it is time to ensure research equality between psychiatry and other medical specialities. *Lancet Reg Health – Eur*. 2024;43.
21. Iliakis EA, Ilagan GS, Choi-Kain LW. Dropout rates from psychotherapy trials for borderline personality disorder: a meta-analysis. *Personal Disord: Theory Res Treat*. 2021;12(3):193.

Received: 30 August 2024 / Accepted: 6 November 2024

Published online: 14 November 2024

22. Stoffers-Winterling J, Kongerslev M, Simonsen E, Lieb K, Storebø OJ. Clinical trials in personality disorders: recommendations for study design and conduct. *Personal Dis: Theor Res Treat*. 2023;14(1):127.
23. Juul S. Improving the methodological quality of randomized clinical trials assessing psychotherapy for borderline personality disorder: recommendations for the future. *Front Psychiatry*. 2022;13:1053844.
24. Bala MM, Akl EA, Sun X, Bassler D, Mertz D, Mejza F, et al. Randomized trials published in higher vs. lower impact journals differ in design, conduct, and analysis. *J Clin Epidemiol*. 2013;66(3):286–95.
25. Gluud LL, Sørensen TI, Gøtzsche PC, Gluud C. The journal impact factor as a predictor of trial quality and outcomes: cohort study of hepatobiliary randomized clinical trials. *Am J Gastroenterol*. 2005;100(11):2431–5.
26. Kamp CB, Petersen JJ, Faltermeier P, Juul S, Siddiqui F, Barbateskovic M et al. Beneficial and harmful effects of tricyclic antidepressants for adults with major depressive disorder: a systematic review with meta-analysis and trial sequential analysis. *BMJ Mental Health*. 2024;27(1).
27. Jakobsen JC, Katakam KK, Schou A, Hellmuth SG, Stallknecht SE, Leth-Møller K, et al. Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis. *BMC Psychiatry*. 2017;17(1):58.
28. Wood A, White I, Thompson S. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials*. 2004;1(4):368–76.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary material 1 – references to the included trials [1-183]

1. Shawyer F, Farhall J, Thomas N, Hayes SC, Gallop R, Copolov D, et al. Acceptance and commitment therapy for psychosis: randomised controlled trial. *Br J Psychiatry*. 2017;210(2):140-8.
2. Ducasse D, Jaussent I, Arpon-Brand V, Vienot M, Laglaoui C, Beziat S, et al. Acceptance and Commitment Therapy for the Management of Suicidal Patients: A Randomized Controlled Trial. *Psychother Psychosom*. 2018;87(4):211-22.
3. Sirey JA, Banerjee S, Marino P, Bruce ML, Halkett A, Turnwald M, et al. Adherence to Depression Treatment in Primary Care: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(11):1129-35.
4. Grabski M, McAndrew A, Lawn W, Marsh B, Raymen L, Stevens T, et al. Adjunctive Ketamine With Relapse Prevention–Based Psychological Therapy in the Treatment of Alcohol Use Disorder. *American Journal of Psychiatry*. 2022;179(2):152-62.
5. Morrison AP, Law H, Carter L, Sellers R, Emsley R, Pyle M, et al. Antipsychotic drugs versus cognitive behavioural therapy versus a combination of both in people with psychosis: a randomised controlled pilot and feasibility study. *Lancet Psychiatry*. 2018;5(5):411-23.
6. Morrison AP, Pyle M, Maughan D, Johns L, Freeman D, Broome MR, et al. Antipsychotic medication versus psychological intervention versus a combination of both in adolescents with first-episode psychosis (MAPS): a multicentre, three-arm, randomised controlled pilot and feasibility study. *Lancet Psychiatry*. 2020;7(9):788-800.
7. Cheryl A. King P, corresponding author¹ Alejandra Arango, MS,¹ Anne Kramer, LMSW,¹ Danielle Busby, PhD,¹ Ewa Czyz, PhD,¹ Cynthia Ewell Foster, PhD,¹ and Brenda W. Gillespie, PhD², for the YST Study Team. Association of the Youth-Nominated Support Team Intervention for Suicidal Adolescents With 11- to 14-Year Mortality Outcomes. *JAMA Psychiatry*. 2019;76(5):492–8.
8. Hyett MP, Bank SR, Lipp OV, Erceg-Hurn DM, Alvares GA, Maclaine E, et al. Attenuated Psychophysiological Reactivity following Single-Session Group Imagery Rescripting versus Verbal Restructuring in Social Anxiety Disorder: Results from a Randomized Controlled Trial. *Psychother Psychosom*. 2018;87(6):340-9.
9. Freeman D, Haselton P, Freeman J, Spanlang B, Kishore S, Albery E, et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. *Lancet Psychiatry*. 2018;5(8):625-32.
10. Freeman D, Lambe S, Kabir T, Petit A, Rosebrock L, Yu LM, et al. Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): a multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. *Lancet Psychiatry*. 2022;9(5):375-88.
11. Craig TK, Rus-Calafell M, Ward T, Leff JP, Huckvale M, Howarth E, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. *Lancet Psychiatry*. 2018;5(1):31-40.
12. Zagorscak P, Heinrich M, Sommer D, Wagner B, Knaevelsrud C. Benefits of Individualized Feedback in Internet-Based Interventions for Depression: A Randomized Controlled Trial. *Psychother Psychosom*. 2018;87(1):32-45.
13. Weersing VR, Brent DA, Rozenman MS, Gonzalez A, Jeffreys M, Dickerson JF, et al. Brief Behavioral Therapy for Pediatric Anxiety and Depression in Primary Care: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(6):571-8.

14. Sloan DM, Marx BP, Lee DJ, Resick PA. A Brief Exposure-Based Treatment vs Cognitive Processing Therapy for Posttraumatic Stress Disorder: A Randomized Noninferiority Clinical Trial. *JAMA Psychiatry*. 2018;75(3):233-9.
15. Wahlund T, Mataix-Cols D, Olofsdotter Lauri K, de Schipper E, Ljotsson B, Aspvall K, et al. Brief Online Cognitive Behavioural Intervention for Dysfunctional Worry Related to the COVID-19 Pandemic: A Randomised Controlled Trial. *Psychother Psychosom*. 2021;90(3):191-9.
16. O'Connor RC, Ferguson E, Scott F, Smyth R, McDaid D, Park AL, et al. A brief psychological intervention to reduce repetition of self-harm in patients admitted to hospital following a suicide attempt: a randomised controlled trial. *Lancet Psychiatry*. 2017;4(6):451-60.
17. Zerwas SC, Watson HJ, Hofmeier SM, Levine MD, Hamer RM, Crosby RD, et al. CBT4BN: A Randomized Controlled Trial of Online Chat and Face-to-Face Group Therapy for Bulimia Nervosa. *Psychother Psychosom*. 2017;86(1):47-53.
18. Boelen PA, Lenferink LIM, Spuij M. CBT for Prolonged Grief in Children and Adolescents: A Randomized Clinical Trial. *Am J Psychiatry*. 2021;178(4):294-304.
19. Berry C, Hodgekins J, French P, Clarke T, Shepstone L, Barton G, et al. Clinical and cost-effectiveness of social recovery therapy for the prevention and treatment of long-term social disability among young people with emerging severe mental illness (PRODIGY): randomised controlled trial. *Br J Psychiatry*. 2022;220(3):154-62.
20. Creswell C, Violato M, Fairbanks H, White E, Parkinson M, Abitabile G, et al. Clinical outcomes and cost-effectiveness of brief guided parent-delivered cognitive behavioural therapy and solution-focused brief therapy for treatment of childhood anxiety disorders: a randomised controlled trial. *Lancet Psychiatry*. 2017;4(7):529-39.
21. Graham AK, Greene CJ, Kwasny MJ, Kaiser SM, Lieponis P, Powell T, et al. Coached Mobile App Platform for the Treatment of Depression and Anxiety Among Primary Care Patients: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(9):906-14.
22. Wilkinson ST, Rhee TG, Joormann J, Webler R, Ortiz Lopez M, Kitay B, et al. Cognitive Behavioral Therapy to Sustain the Antidepressant Effects of Ketamine in Treatment-Resistant Depression: A Randomized Clinical Trial. *Psychother Psychosom*. 2021;90(5):318-27.
23. Wood JJ, Kendall PC, Wood KS, Kerns CM, Seltzer M, Small BJ, et al. Cognitive Behavioral Treatments for Anxiety in Children With Autism Spectrum Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(5):474-83.
24. Kleinstaub M, Allwang C, Bailer J, Berking M, Brunahl C, Erkcic M, et al. Cognitive Behaviour Therapy Complemented with Emotion Regulation Training for Patients with Persistent Physical Symptoms: A Randomised Clinical Trial. *Psychother Psychosom*. 2019;88(5):287-99.
25. Goodyer IM, Reynolds S, Barrett B, Byford S, Dubicka B, Hill J, et al. Cognitive behavioural therapy and short-term psychoanalytical psychotherapy versus a brief psychosocial intervention in adolescents with unipolar major depressive disorder (IMPACT): a multicentre, pragmatic, observer-blind, randomised controlled superiority trial. *The Lancet Psychiatry*. 2017;4(2):109-19.
26. Goldstein LH, Robinson EJ, Mellers JDC, Stone J, Carson A, Reuber M, et al. Cognitive behavioural therapy for adults with dissociative seizures (CODES): a pragmatic, multicentre, randomised controlled trial. *Lancet Psychiatry*. 2020;7(6):491-505.
27. van Passel B, Danner UN, Dingemans AE, Aarts E, Sternheim LC, Becker ES, et al. Cognitive Remediation Therapy Does Not Enhance Treatment Effect in Obsessive-Compulsive Disorder and Anorexia Nervosa: A Randomized Controlled Trial. *Psychother Psychosom*. 2020;89(4):228-41.

28. Salzer S, Stefini A, Kronmuller KT, Leibing E, Leichsenring F, Henningsen P, et al. Cognitive-Behavioral and Psychodynamic Therapy in Adolescents with Social Anxiety Disorder: A Multicenter Randomized Controlled Trial. *Psychother Psychosom.* 2018;87(4):223-33.
29. Hilbert A, Petroff D, Neuhaus P, Schmidt R. Cognitive-Behavioral Therapy for Adolescents with an Age-Adapted Diagnosis of Binge-Eating Disorder: A Randomized Clinical Trial. *Psychother Psychosom.* 2020;89(1):51-3.
30. Dekker K, Benjamins JS, Maksimovic T, Filardi M, Hofman WF, van Straten A, et al. Combined Internet-Based Cognitive-Behavioral and Chronobiological Intervention for Insomnia: A Randomized Controlled Trial. *Psychother Psychosom.* 2020;89(2):117-8.
31. Green J, Leadbitter K, Ellis C, Taylor L, Moore HL, Carruthers S, et al. Combined social communication therapy at home and in education for young autistic children in England (PACT-G): a parallel, single-blind, randomised controlled trial. *Lancet Psychiatry.* 2022;9(4):307-20.
32. Collins SE, Duncan MH, Saxon AJ, Taylor EM, Mayberry N, Merrill JO, et al. Combining behavioral harm-reduction treatment and extended-release naltrexone for people experiencing homelessness and alcohol use disorder in the USA: a randomised clinical trial. *Lancet Psychiatry.* 2021;8(4):287-300.
33. Williams C, McClay CA, Matthews L, McConnachie A, Haig C, Walker A, et al. Community-based group guided self-help intervention for low mood and stress: randomised controlled trial. *Br J Psychiatry.* 2018;212(2):88-95.
34. Wong SY, Zhang DX, Li CC, Yip BH, Chan DC, Ling YM, et al. Comparing the Effects of Mindfulness-Based Cognitive Therapy and Sleep Psycho-Education with Exercise on Chronic Insomnia: A Randomised Controlled Trial. *Psychother Psychosom.* 2017;86(4):241-53.
35. Rief W, Bleichhardt G, Dannehl K, Euteneuer F, Wambach K. Comparing the Efficacy of CBASP with Two Versions of CBT for Depression in a Routine Care Center: A Randomized Clinical Trial. *Psychother Psychosom.* 2018;87(3):164-78.
36. Malla A, Joober R, Iyer S, Norman R, Schmitz N, Brown T, et al. Comparing three-year extension of early intervention service to regular care following two years of early intervention service in first-episode psychosis: a randomized single blind clinical trial. *World Psychiatry.* 2017;16(3):278-86.
37. Freeman D, Emsley R, Diamond R, Collett N, Bold E, Chadwick E, et al. Comparison of a theoretically driven cognitive therapy (the Feeling Safe Programme) with befriending for the treatment of persistent persecutory delusions: a parallel, single-blind, randomised controlled trial. *Lancet Psychiatry.* 2021;8(8):696-707.
38. Jahoda A, Hastings R, Hatton C, Cooper S-A, Dagnan D, Zhang R, et al. Comparison of behavioural activation with guided self-help for treatment of depression in adults with intellectual disabilities: a randomised controlled trial. *The Lancet Psychiatry.* 2017;4(12):909-19.
39. Ougrin D, Corrigan R, Poole J, Zundel T, Sarhane M, Slater V, et al. Comparison of effectiveness and cost-effectiveness of an intensive community supported discharge service versus treatment as usual for adolescents with psychiatric emergencies: a randomised controlled trial. *Lancet Psychiatry.* 2018;5(6):477-85.
40. Fortney JC, Bauer AM, Cerimele JM, Pyne JM, Pfeiffer P, Heagerty PJ, et al. Comparison of Teleintegrated Care and Telereferral Care for Treating Complex Psychiatric Disorders in Primary Care: A Pragmatic Randomized Comparative Effectiveness Trial. *JAMA Psychiatry.* 2021;78(11):1189-99.
41. Cooney P, Jackman C, Coyle D, O'Reilly G. Computerised cognitive-behavioural therapy for adults with intellectual disability: randomised controlled trial. *Br J Psychiatry.* 2017;211(2):95-102.

42. Ruglass LM, Lopez-Castro T, Papini S, Killeen T, Back SE, Hien DA. Concurrent Treatment with Prolonged Exposure for Co-Occurring Full or Subthreshold Posttraumatic Stress Disorder and Substance Use Disorders: A Randomized Clinical Trial. *Psychother Psychosom*. 2017;86(3):150-61.
43. Hamdani SU, Huma ZE, Rahman A, Wang D, Chen T, van Ommeren M, et al. Cost-effectiveness of WHO Problem Management Plus for adults with mood and anxiety disorders in a post-conflict area of Pakistan: randomised controlled trial. *Br J Psychiatry*. 2020;217(5):623-9.
44. Smits ML, Feenstra DJ, Eeren HV, Bales DL, Laurensen EMP, Blankers M, et al. Day hospital versus intensive out-patient mentalisation-based treatment for borderline personality disorder: multicentre randomised clinical trial. *Br J Psychiatry*. 2020;216(2):79-84.
45. Fuhr DC, Weobong B, Lazarus A, Vanobberghen F, Weiss HA, Singla DR, et al. Delivering the Thinking Healthy Programme for perinatal depression through peers: an individually randomised controlled trial in India. *Lancet Psychiatry*. 2019;6(2):115-27.
46. Sikander S, Ahmad I, Atif N, Zaidi A, Vanobberghen F, Weiss HA, et al. Delivering the Thinking Healthy Programme for perinatal depression through volunteer peers: a cluster randomised controlled trial in Pakistan. *Lancet Psychiatry*. 2019;6(2):128-39.
47. Pearce S, Scott L, Attwood G, Saunders K, Dean M, De Ridder R, et al. Democratic therapeutic community treatment for personality disorder: randomised controlled trial. *Br J Psychiatry*. 2017;210(2):149-56.
48. Bohus M, Kleindienst N, Hahn C, Muller-Engelmann M, Ludascher P, Steil R, et al. Dialectical Behavior Therapy for Posttraumatic Stress Disorder (DBT-PTSD) Compared With Cognitive Processing Therapy (CPT) in Complex Presentations of PTSD in Women Survivors of Childhood Abuse: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(12):1235-45.
49. Keefe RSE, Canadas E, Farlow D, Etkin A. Digital Intervention for Cognitive Deficits in Major Depression: A Randomized Controlled Trial to Assess Efficacy and Safety in Adults. *Am J Psychiatry*. 2022;179(7):482-9.
50. Zoellner LA, Roy-Byrne PP, Mavissakalian M, Feeny NC. Doubly Randomized Preference Trial of Prolonged Exposure Versus Sertraline for Treatment of PTSD. *Am J Psychiatry*. 2019;176(4):287-96.
51. von Kanel R, Barth J, Princip M, Meister-Langraf RE, Schmid JP, Znoj H, et al. Early Psychological Counseling for the Prevention of Posttraumatic Stress Induced by Acute Coronary Syndrome: The MI-SPRINT Randomized Controlled Trial. *Psychother Psychosom*. 2018;87(2):75-84.
52. Chanen AM, Betts JK, Jackson H, Cotton SM, Gleeson J, Davey CG, et al. Effect of 3 Forms of Early Intervention for Young People With Borderline Personality Disorder: The MOBY Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(2):109-19.
53. Ritterband LM, Thorndike FP, Ingersoll KS, Lord HR, Gonder-Frederick L, Frederick C, et al. Effect of a Web-Based Cognitive Behavior Therapy for Insomnia Intervention With 1-Year Follow-up: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(1):68-75.
54. Manning V, Garfield JBB, Staiger PK, Lubman DI, Lum JAG, Reynolds J, et al. Effect of Cognitive Bias Modification on Early Relapse Among Adults Undergoing Inpatient Alcohol Withdrawal Treatment: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(2):133-40.
55. Zhou ES, Ritterband LM, Bethea TN, Robles YP, Heeren TC, Rosenberg L. Effect of Culturally Tailored, Internet-Delivered Cognitive Behavioral Therapy for Insomnia in Black Women: A Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(6):538-49.
56. Rosner R, Rimane E, Frick U, Gutermann J, Hagl M, Renneberg B, et al. Effect of Developmentally Adapted Cognitive Processing Therapy for Youth With Symptoms of

Posttraumatic Stress Disorder After Childhood Sexual and Physical Abuse: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(5):484-91.

57. Espie CA, Emsley R, Kyle SD, Gordon C, Drake CL, Siriwardena AN, et al. Effect of Digital Cognitive Behavioral Therapy for Insomnia on Health, Psychological Well-being, and Sleep-Related Quality of Life: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(1):21-30.

58. Schramm E, Kriston L, Zobel I, Bailer J, Wambach K, Backenstrass M, et al. Effect of Disorder-Specific vs Nonspecific Psychotherapy for Chronic Depression: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(3):233-42.

59. Resick PA, Wachen JS, Dondanville KA, Pruiksma KE, Yarvis JS, Peterson AL, et al. Effect of Group vs Individual Cognitive Processing Therapy in Active-Duty Military Seeking Treatment for Posttraumatic Stress Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(1):28-36.

60. McDonell MG, Hirschak KA, Herron J, Lyons AJ, Alcover KC, Shaw J, et al. Effect of Incentives for Alcohol Abstinence in Partnership With 3 American Indian and Alaska Native Communities: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(6):599-606.

61. Axelsson E, Andersson E, Ljotsson B, Bjorkander D, Hedman-Lagerlof M, Hedman-Lagerlof E. Effect of Internet vs Face-to-Face Cognitive Behavior Therapy for Health Anxiety: A Randomized Noninferiority Clinical Trial. *JAMA Psychiatry*. 2020;77(9):915-24.

62. de Zwaan M, Herpertz S, Zipfel S, Svaldi J, Friederich HC, Schmidt F, et al. Effect of Internet-Based Guided Self-help vs Individual Face-to-Face Treatment on Full or Subsyndromal Binge Eating Disorder in Overweight or Obese Patients: The INTERBED Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(10):987-95.

63. Kahlon MK, Aksan N, Aubrey R, Clark N, Cowley-Morillo M, Jacobs EA, et al. Effect of Layperson-Delivered, Empathy-Focused Program of Telephone Calls on Loneliness, Depression, and Anxiety Among Adults During the COVID-19 Pandemic: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(6):616-22.

64. Constantino MJ, Boswell JF, Coyne AE, Swales TP, Kraus DR. Effect of Matching Therapists to Patients vs Assignment as Usual on Adult Psychotherapy Outcomes: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(9):960-9.

65. Van Lieshout RJ, Layton H, Savoy CD, Brown JSL, Ferro MA, Streiner DL, et al. Effect of Online 1-Day Cognitive Behavioral Therapy-Based Workshops Plus Usual Care vs Usual Care Alone for Postpartum Depression: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(11):1200-7.

66. Svensson M, Nilsson T, Perrin S, Johansson H, Viborg G, Falkenstrom F, et al. The Effect of Patient's Choice of Cognitive Behavioural or Psychodynamic Therapy on Outcomes for Panic Disorder: A Doubly Randomised Controlled Preference Trial. *Psychother Psychosom*. 2021;90(2):107-18.

67. Tinland A, Loubiere S, Mougeot F, Jouet E, Pontier M, Baumstarck K, et al. Effect of Psychiatric Advance Directives Facilitated by Peer Workers on Compulsory Admission Among People With Mental Illness: A Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(8):752-9.

68. Osborn TL, Ventura-Conerly KE, Arango GS, Roe E, Rodriguez M, Alemu RG, et al. Effect of Shamiri Layperson-Provided Intervention vs Study Skills Control Intervention for Depression and Anxiety Symptoms in Adolescents in Kenya: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(8):829-37.

69. McMain SF, Chapman AL, Kuo JR, Dixon-Gordon KL, Guimond TH, Labrish C, et al. The Effectiveness of 6 versus 12 Months of Dialectical Behavior Therapy for Borderline Personality Disorder: A Noninferiority Randomized Clinical Trial. *Psychother Psychosom*. 2022;91(6):382-97.

70. Baumeister H, Paganini S, Sander LB, Lin J, Schlicker S, Terhorst Y, et al. Effectiveness of a Guided Internet- and Mobile-Based Intervention for Patients with Chronic Back Pain and Depression (WARD-BP): A Multicenter, Pragmatic Randomized Controlled Trial. *Psychother Psychosom.* 2021;90(4):255-68.
71. Kramer LV, Grunzig SD, Baumeister H, Ebert DD, Bengel J. Effectiveness of a Guided Web-Based Intervention to Reduce Depressive Symptoms before Outpatient Psychotherapy: A Pragmatic Randomized Controlled Trial. *Psychother Psychosom.* 2021;90(4):233-42.
72. Maselko J, Sikander S, Turner EL, Bates LM, Ahmad I, Atif N, et al. Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years postnatal: a cluster randomised trial in Pakistan. *Lancet Psychiatry.* 2020;7(9):775-87.
73. Lubman DI, Grigg J, Reynolds J, Hall K, Baker AL, Staiger PK, et al. Effectiveness of a Stand-alone Telephone-Delivered Intervention for Reducing Problem Alcohol Use: A Randomized Clinical Trial. *JAMA Psychiatry.* 2022;79(11):1055-64.
74. Bryant RA, Dawson KS, Keyan D, Azevedo S, Yadav S, Tran J, et al. Effectiveness of a Videoconferencing-Delivered Psychological Intervention for Mental Health Problems during COVID-19: A Proof-of-Concept Randomized Clinical Trial. *Psychother Psychosom.* 2022;91(1):63-72.
75. Claus BB, Scherbaum N, Bonnet U. Effectiveness of an Adjunctive Psychotherapeutic Intervention Developed for Enhancing the Placebo Effect of Antidepressants Used within an Inpatient-Treatment Program of Major Depression: A Pragmatic Parallel-Group, Randomized Controlled Trial. *Psychother Psychosom.* 2020;89(4):258-60.
76. Serfaty M, King M, Nazareth I, Moorey S, Aspden T, Mannix K, et al. Effectiveness of cognitive-behavioural therapy for depression in advanced cancer: CanTalk randomised controlled trial. *Br J Psychiatry.* 2020;216(4):213-21.
77. Cornelis J, Barakat A, Blankers M, Peen J, Lommerse N, Eikelenboom M, et al. The effectiveness of intensive home treatment as a substitute for hospital admission in acute psychiatric crisis resolution in the Netherlands: a two-centre Zelen double-consent randomised controlled trial. *Lancet Psychiatry.* 2022;9(8):625-35.
78. Rollman BL, Herbeck Belnap B, Abebe KZ, Spring MB, Rotondi AJ, Rothenberger SD, et al. Effectiveness of Online Collaborative Care for Treating Mood and Anxiety Disorders in Primary Care: A Randomized Clinical Trial. *JAMA Psychiatry.* 2018;75(1):56-64.
79. Arntz A, Jacob GA, Lee CW, Brand-de Wilde OM, Fassbinder E, Harper RP, et al. Effectiveness of Predominantly Group Schema Therapy and Combined Individual and Group Schema Therapy for Borderline Personality Disorder: A Randomized Clinical Trial. *JAMA Psychiatry.* 2022;79(4):287-99.
80. Bockting CLH, Klein NS, Elgersma HJ, van Rijsbergen GD, Slofstra C, Ormel J, et al. Effectiveness of preventive cognitive therapy while tapering antidepressants versus maintenance antidepressant treatment versus their combination in prevention of depressive relapse or recurrence (DRD study): a three-group, multicentre, randomised controlled trial. *Lancet Psychiatry.* 2018;5(5):401-10.
81. Knapstad M, Lervik LV, Saether SMM, Aaro LE, Smith ORF. Effectiveness of Prompt Mental Health Care, the Norwegian Version of Improving Access to Psychological Therapies: A Randomized Controlled Trial. *Psychother Psychosom.* 2020;89(2):90-105.
82. Jordans MJD, Luitel NP, Garman E, Kohrt BA, Rathod SD, Shrestha P, et al. Effectiveness of psychological treatments for depression and alcohol use disorder delivered by community-based counsellors: two pragmatic randomised controlled trials within primary healthcare in Nepal. *Br J Psychiatry.* 2019;215(2):485-93.

83. Tjaden C, Mulder CL, den Hollander W, Castelein S, Delespaul P, Keet R, et al. Effectiveness of Resource Groups for Improving Empowerment, Quality of Life, and Functioning of People With Severe Mental Illness: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(12):1309-18.
84. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of Self-guided App-Based Virtual Reality Cognitive Behavior Therapy for Acrophobia: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(7):682-90.
85. Morin CM, Edinger JD, Beaulieu-Bonneau S, Ivers H, Krystal AD, Guay B, et al. Effectiveness of Sequential Psychological and Medication Therapies for Insomnia Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(11):1107-15.
86. Biesheuvel-Leliefeld KEM, Dijkstra-Kersten SMA, van Schaik DJF, van Marwijk HWJ, Smit F, van der Horst HE, et al. Effectiveness of Supported Self-Help in Recurrent Depression: A Randomized Controlled Trial in Primary Care. *Psychother Psychosom*. 2017;86(4):220-30.
87. Cottrell DJ, Wright-Hughes A, Collinson M, Boston P, Eisler I, Fortune S, et al. Effectiveness of systemic family therapy versus treatment as usual for young people after self-harm: a pragmatic, phase 3, multicentre, randomised controlled trial. *Lancet Psychiatry*. 2018;5(3):203-16.
88. Dorsey S, Lucid L, Martin P, King KM, O'Donnell K, Murray LK, et al. Effectiveness of Task-Shifted Trauma-Focused Cognitive Behavioral Therapy for Children Who Experienced Parental Death and Posttraumatic Stress in Kenya and Tanzania: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(5):464-73.
89. Marchant NL, Barnhofer T, Coueron R, Wirth M, Lutz A, Arenaza-Urquijo EM, et al. Effects of a Mindfulness-Based Intervention versus Health Self-Management on Subclinical Anxiety in Older Adults with Subjective Cognitive Decline: The SCD-Well Randomized Superiority Trial. *Psychother Psychosom*. 2021;90(5):341-50.
90. Gieselmann A, Bockermann M, Sorbi M, Pietrowsky R. The Effects of an Internet-Based Imagery Rehearsal Intervention: A Randomized Controlled Trial. *Psychother Psychosom*. 2017;86(4):231-40.
91. Miklowitz DJ, Schneek CD, Walshaw PD, Singh MK, Sullivan AE, Suddath RL, et al. Effects of Family-Focused Therapy vs Enhanced Usual Care for Symptomatic Youths at High Risk for Bipolar Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(5):455-63.
92. Freeman D, Sheaves B, Goodwin GM, Yu LM, Nickless A, Harrison PJ, et al. The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. *Lancet Psychiatry*. 2017;4(10):749-58.
93. Christensen TN, Wallstrom IG, Stenager E, Bojesen AB, Gluud C, Nordentoft M, et al. Effects of Individual Placement and Support Supplemented With Cognitive Remediation and Work-Focused Social Skills Training for People With Severe Mental Illness: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(12):1232-40.
94. Woud ML, Blackwell SE, Shkreli L, Wurtz F, Cwik JC, Margraf J, et al. The Effects of Modifying Dysfunctional Appraisals in Posttraumatic Stress Disorder Using a Form of Cognitive Bias Modification: Results of a Randomized Controlled Trial in an Inpatient Setting. *Psychother Psychosom*. 2021;90(6):386-402.
95. Bruijniks SJE, Lemmens L, Hollon SD, Peeters F, Cuijpers P, Arntz A, et al. The effects of once- versus twice-weekly sessions on psychotherapy outcomes in depressed patients. *Br J Psychiatry*. 2020;216(4):222-30.
96. Dunlop BW, Kelley ME, Aponte-Rivera V, Mletzko-Crowe T, Kinkead B, Ritchie JC, et al. Effects of Patient Preferences on Outcomes in the Predictors of Remission in Depression to

Individual and Combined Treatments (PReDICT) Study. *American Journal of Psychiatry*. 2017;174(6):546-56.

97. Davis AK, Barrett FS, May DG, Cosimano MP, Sepeda ND, Johnson MW, et al. Effects of Psilocybin-Assisted Therapy on Major Depressive Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(5):481-9.

98. Burger H, Verbeek T, Aris-Meijer JL, Beijers C, Mol BW, Hollon SD, et al. Effects of psychological treatment of mental health problems in pregnant women to protect their offspring: randomised controlled trial. *Br J Psychiatry*. 2020;216(4):182-8.

99. Garety P, Ward T, Emsley R, Greenwood K, Freeman D, Fowler D, et al. Effects of SlowMo, a Blended Digital Therapy Targeting Reasoning, on Paranoia Among People With Psychosis: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(7):714-25.

100. Hanlon C, Medhin G, Dewey ME, Prince M, Assefa E, Shibre T, et al. Efficacy and cost-effectiveness of task-shared care for people with severe mental disorders in Ethiopia (TaSCS): a single-blind, randomised, controlled, phase 3 non-inferiority trial. *Lancet Psychiatry*. 2022;9(1):59-71.

101. Wilhelm S, Phillips KA, Greenberg JL, O'Keefe SM, Hoepfner SS, Keshaviah A, et al. Efficacy and Posttreatment Effects of Therapist-Delivered Cognitive Behavioral Therapy vs Supportive Psychotherapy for Adults With Body Dysmorphic Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(4):363-73.

102. Silverstein M, Diaz-Linhart Y, Cabral H, Beardslee W, Hegel M, Haile W, et al. Efficacy of a Maternal Depression Prevention Strategy in Head Start: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(8):781-9.

103. Ilgen MA, Coughlin LN, Bohnert ASB, Chermack S, Price A, Kim HM, et al. Efficacy of a Psychosocial Pain Management Intervention for Men and Women With Substance Use Disorders and Chronic Pain: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(12):1225-34.

104. Dittmer N, Voderholzer U, Monch C, Cuntz U, Jacobi C, Schlegl S. Efficacy of a Specialized Group Intervention for Compulsive Exercise in Inpatients with Anorexia Nervosa: A Randomized Controlled Trial. *Psychother Psychosom*. 2020;89(3):161-73.

105. Myin-Germeys I, van Aubel E, Vaessen T, Steinhart H, Klippel A, Lafit G, et al. Efficacy of Acceptance and Commitment Therapy in Daily Life in Early Psychosis: Results from the Multi-Center INTERACT Randomized Controlled Trial. *Psychother Psychosom*. 2022;91(6):411-23.

106. Wilhelm S, Weingarden H, Greenberg JL, Hoepfner SS, Snorrason I, Bernstein EE, et al. Efficacy of App-Based Cognitive Behavioral Therapy for Body Dysmorphic Disorder with Coach Support: Initial Randomized Controlled Clinical Trial. *Psychother Psychosom*. 2022;91(4):277-85.

107. Aust S, Brakemeier EL, Spies J, Herrera-Melendez AL, Kaiser T, Fallgatter A, et al. Efficacy of Augmentation of Cognitive Behavioral Therapy With Transcranial Direct Current Stimulation for Depression: A Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(6):528-37.

108. McCauley E, Berk MS, Asarnow JR, Adrian M, Cohen J, Korslund K, et al. Efficacy of Dialectical Behavior Therapy for Adolescents at High Risk for Suicide: A Randomized Clinical Trial. *JAMA Psychiatry*. 2018;75(8):777-85.

109. Felder JN, Epel ES, Neuhaus J, Krystal AD, Prather AA. Efficacy of Digital Cognitive Behavioral Therapy for the Treatment of Insomnia Symptoms Among Pregnant Women: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(5):484-92.

110. Norman SB, Trim R, Haller M, Davis BC, Myers US, Colvonen PJ, et al. Efficacy of Integrated Exposure Therapy vs Integrated Coping Skills Therapy for Comorbid Posttraumatic

Stress Disorder and Alcohol Use Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(8):791-9.

111. Stigsdotter UK, Corazon SS, Sidenius U, Nyed PK, Larsen HB, Fjorback LO. Efficacy of nature-based therapy for individuals with stress-related illnesses: randomised controlled trial. *Br J Psychiatry*. 2018;213(1):404-11.

112. Rauch SAM, Kim HM, Powell C, Tuerk PW, Simon NM, Acierno R, et al. Efficacy of Prolonged Exposure Therapy, Sertraline Hydrochloride, and Their Combination Among Combat Veterans With Posttraumatic Stress Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(2):117-26.

113. Wolfling K, Muller KW, Dreier M, Ruckes C, Deuster O, Batra A, et al. Efficacy of Short-term Treatment of Internet and Computer Game Addiction: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(10):1018-25.

114. Janse A, Worm-Smeitink M, Bleijenberg G, Donders R, Knoop H. Efficacy of web-based cognitive-behavioural therapy for chronic fatigue syndrome: randomised controlled trial. *Br J Psychiatry*. 2018;212(2):112-8.

115. Simon NM, Hofmann SG, Rosenfield D, Hoepfner SS, Hoge EA, Bui E, et al. Efficacy of Yoga vs Cognitive Behavioral Therapy vs Stress Education for the Treatment of Generalized Anxiety Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(1):13-20.

116. Gumley AI, Bradstreet S, Ainsworth J, Allan S, Alvarez-Jimenez M, Aucott L, et al. The EMPOWER blended digital intervention for relapse prevention in schizophrenia: a feasibility cluster randomised controlled trial in Scotland and Australia. *Lancet Psychiatry*. 2022;9(6):477-86.

117. Aardema F, Bouchard S, Koszycki D, Lavoie ME, Audet JS, O'Connor K. Evaluation of Inference-Based Cognitive-Behavioral Therapy for Obsessive-Compulsive Disorder: A Multicenter Randomized Controlled Trial with Three Treatment Modalities. *Psychother Psychosom*. 2022;91(5):348-59.

118. Delgadillo J, de Jong K, Lucock M, Lutz W, Rubel J, Gilbody S, et al. Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. *Lancet Psychiatry*. 2018;5(7):564-72.

119. Herzog W, Wild B, Giel KE, Junne F, Friederich H-C, Resmark G, et al. Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in female outpatients with anorexia nervosa (ANTOP study): 5-year follow-up of a randomised controlled trial in Germany. *The Lancet Psychiatry*. 2022;9(4):280-90.

120. Lazarov A, Pine DS, Bar-Haim Y. Gaze-Contingent Music Reward Therapy for Social Anxiety Disorder: A Randomized Controlled Trial. *Am J Psychiatry*. 2017;174(7):649-56.

121. Cruwys T, Haslam C, Rathbone JA, Williams E, Haslam SA, Walter ZC. Groups 4 Health versus cognitive-behavioural therapy for depression and loneliness in young people: randomised phase 3 non-inferiority trial with 12-month follow-up. *Br J Psychiatry*. 2022;220(3):140-7.

122. Gureje O, Oladeji BD, Montgomery AA, Araya R, Bello T, Chisholm D, et al. High-versus low-intensity interventions for perinatal depression delivered by non-specialist primary maternal care providers in Nigeria: cluster randomised controlled trial (the EXPONATE trial). *Br J Psychiatry*. 2019;215(3):528-35.

123. Stulz N, Wyder L, Maeck L, Hilpert M, Lerzer H, Zander E, et al. Home treatment for acute mental healthcare: randomised controlled trial. *Br J Psychiatry*. 2020;216(6):323-30.

124. Boterhoven de Haan KL, Lee CW, Fassbinder E, van Es SM, Menninga S, Meewisse ML, et al. Imagery rescripting and eye movement desensitisation and reprocessing as treatment for adults with post-traumatic stress disorder from childhood trauma: randomised clinical trial. *Br J Psychiatry*. 2020;217(5):609-15.

125. Schmid C, Hansen K, Kroner-Borowik T, Steil R. Imagery Rescripting and Imaginal Exposure in Nightmare Disorder Compared to Positive Imagery: A Randomized Controlled Trial. *Psychother Psychosom.* 2021;90(5):328-40.
126. Cristea IA, Ioannidis JPA. Improving Disclosure of Financial Conflicts of Interest for Research on Psychosocial Interventions. *JAMA Psychiatry.* 2018;75(6):541-2.
127. Thase ME, Wright JH, Eells TD, Barrett MS, Wisniewski SR, Balasubramani GK, et al. Improving the Efficiency of Psychotherapy for Depression: Computer-Assisted Versus Standard CBT. *Am J Psychiatry.* 2018;175(3):242-50.
128. Schag K, Rennhak SK, Leehr EJ, Skoda EM, Becker S, Bethge W, et al. IMPULS: Impulsivity-Focused Group Intervention to Reduce Binge Eating Episodes in Patients with Binge Eating Disorder - A Randomised Controlled Trial. *Psychother Psychosom.* 2019;88(3):141-53.
129. Bormann JE, Thorp SR, Smith E, Glickman M, Beck D, Plumb D, et al. Individual Treatment of Posttraumatic Stress Disorder Using Mantram Repetition: A Randomized Clinical Trial. *Am J Psychiatry.* 2018;175(10):979-88.
130. van Gelderen MJ, Nijdam MJ, Haagen JFG, Vermetten E. Interactive Motion-Assisted Exposure Therapy for Veterans with Treatment-Resistant Posttraumatic Stress Disorder: A Randomized Controlled Trial. *Psychother Psychosom.* 2020;89(4):215-27.
131. Arjadi R, Nauta MH, Scholte WF, Hollon SD, Chowdhary N, Suryani AO, et al. Internet-based behavioural activation with lay counsellor support versus online minimal psychoeducation without support for treatment of depression: a randomised controlled trial in Indonesia. *Lancet Psychiatry.* 2018;5(9):707-16.
132. Lindsater E, Axelsson E, Salomonsson S, Santoft F, Ejeby K, Ljotsson B, et al. Internet-Based Cognitive Behavioral Therapy for Chronic Stress: A Randomized Controlled Trial. *Psychother Psychosom.* 2018;87(5):296-305.
133. Weise C, Kaiser G, Janda C, Kues JN, Andersson G, Strahler J, et al. Internet-Based Cognitive-Behavioural Intervention for Women with Premenstrual Dysphoric Disorder: A Randomized Controlled Trial. *Psychother Psychosom.* 2019;88(1):16-29.
134. Camacho EM, Davies LM, Hann M, Small N, Bower P, Chew-Graham C, et al. Long-term clinical and cost-effectiveness of collaborative care (versus usual care) for people with mental-physical multimorbidity: cluster-randomised trial. *Br J Psychiatry.* 2018;213(2):456-63.
135. McKenna L, Marks EM, Hallsworth CA, Schaette R. Mindfulness-Based Cognitive Therapy as a Treatment for Chronic Tinnitus: A Randomized Controlled Trial. *Psychother Psychosom.* 2017;86(6):351-61.
136. Torres-Platas SG, Escobar S, Belliveau C, Wu J, Sasi N, Fotso J, et al. Mindfulness-Based Cognitive Therapy Intervention for the Treatment of Late-Life Depression and Anxiety Symptoms in Primary Care: A Randomized Controlled Trial. *Psychother Psychosom.* 2019;88(4):254-6.
137. Stein A, Netsi E, Lawrence PJ, Granger C, Kempton C, Craske MG, et al. Mitigating the effect of persistent postnatal depression on child outcomes through an intervention to treat depression and improve parenting: a randomised controlled trial. *Lancet Psychiatry.* 2018;5(2):134-44.
138. Singla DR, MacKinnon DP, Fuhr DC, Sikander S, Rahman A, Patel V. Multiple mediation analysis of the peer-delivered Thinking Healthy Programme for perinatal depression: findings from two parallel, randomised controlled trials. *Br J Psychiatry.* 2021;218(3):143-50.
139. Fonagy P, Butler S, Cottrell D, Scott S, Pilling S, Eisler I, et al. Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): 5-year follow-up of a pragmatic, randomised, controlled, superiority trial. *Lancet Psychiatry.* 2020;7(5):420-30.

140. Fonagy P, Butler S, Cottrell D, Scott S, Pilling S, Eisler I, et al. Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. *Lancet Psychiatry*. 2018;5(2):119-33.
141. Grilo CM, Lydecker JA, Fineberg SK, Moreno JO, Ivezaj V, Gueorguieva R. Naltrexone-Bupropion and Behavior Therapy, Alone and Combined, for Binge-Eating Disorder: Randomized Double-Blind Placebo-Controlled Trial. *American Journal of Psychiatry*. 2022;179(12):927-37.
142. Schonenberg M, Wiedemann E, Schneidt A, Scheeff J, Logemann A, Keune PM, et al. Neurofeedback, sham neurofeedback, and cognitive-behavioural group therapy in adults with attention-deficit hyperactivity disorder: a triple-blind, randomised, controlled trial. *Lancet Psychiatry*. 2017;4(9):673-84.
143. Nidich S, Mills PJ, Rainforth M, Heppner P, Schneider RH, Rosenthal NE, et al. Non-trauma-focused meditation versus exposure therapy in veterans with post-traumatic stress disorder: a randomised controlled trial. *Lancet Psychiatry*. 2018;5(12):975-86.
144. Van der Zweerde T, Lancee J, Slottje P, Bosmans JE, Van Someren EJW, van Straten A. Nurse-Guided Internet-Delivered Cognitive Behavioral Therapy for Insomnia in General Practice: Results from a Pragmatic Randomized Clinical Trial. *Psychother Psychosom*. 2020;89(3):174-84.
145. Zwerenz R, Becker J, Knickenberg RJ, Siepmann M, Hagen K, Beutel ME. Online Self-Help as an Add-On to Inpatient Psychotherapy: Efficacy of a New Blended Treatment Approach. *Psychother Psychosom*. 2017;86(6):341-50.
146. Segal ZV, Dimidjian S, Beck A, Boggs JM, Vanderkruik R, Metcalf CA, et al. Outcomes of Online Mindfulness-Based Cognitive Therapy for Patients With Residual Depressive Symptoms: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(6):563-73.
147. Gillard S, Bremner S, Patel A, Goldsmith L, Marks J, Foster R, et al. Peer support for discharge from inpatient mental health care versus care as usual in England (ENRICH): a parallel, two-group, individually randomised controlled trial. *Lancet Psychiatry*. 2022;9(2):125-36.
148. Barkham M, Saxon D, Hardy GE, Bradburn M, Galloway D, Wickramasekera N, et al. Person-centred experiential therapy versus cognitive behavioural therapy delivered in the English Improving Access to Psychological Therapies service for the treatment of moderate or severe depression (PRaCTICED): a pragmatic, randomised, non-inferiority trial. *Lancet Psychiatry*. 2021;8(6):487-99.
149. Irwin MR, Carrillo C, Sadeghi N, Bjurstrom MF, Breen EC, Olmstead R. Prevention of Incident and Recurrent Major Depression in Older Adults With Insomnia: A Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(1):33-41.
150. DeRubeis RJ, Zajecka J, Shelton RC, Amsterdam JD, Fawcett J, Xu C, et al. Prevention of Recurrence After Recovery From a Major Depressive Episode With Antidepressant Medication Alone or in Combination With Cognitive Behavioral Therapy: Phase 2 of a 2-Phase Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(3):237-45.
151. Rossouw J, Yadin E, Alexander D, Seedat S. Prolonged exposure therapy and supportive counselling for post-traumatic stress disorder in adolescents: task-shifting randomised controlled trial. *Br J Psychiatry*. 2018;213(4):587-94.
152. Forsell E, Jernelov S, Blom K, Kraepelien M, Svanborg C, Andersson G, et al. Proof of Concept for an Adaptive Treatment Strategy to Prevent Failures in Internet-Delivered CBT: A Single-Blind Randomized Clinical Trial With Insomnia Patients. *Am J Psychiatry*. 2019;176(4):315-23.

153. Lely JCG, Knipscheer JW, Moerbeek M, Ter Heide FJJ, van den Bout J, Kleber RJ. Randomised controlled trial comparing narrative exposure therapy with present-centred therapy for older patients with post-traumatic stress disorder. *Br J Psychiatry*. 2019;214(6):369-77.
154. Pigeon WR, Crean HF, Cerulli C, Gallegos AM, Bishop TM, Heffner KL. A Randomized Clinical Trial of Cognitive-Behavioral Therapy for Insomnia to Augment Posttraumatic Stress Disorder Treatment in Survivors of Interpersonal Violence. *Psychother Psychosom*. 2022;91(1):50-62.
155. Kiluk BD, Nich C, Buck MB, Devore KA, Frankforter TL, LaPaglia DM, et al. Randomized Clinical Trial of Computerized and Clinician-Delivered CBT in Comparison With Standard Outpatient Treatment for Substance Use Disorders: Primary Within-Treatment and Follow-Up Outcomes. *Am J Psychiatry*. 2018;175(9):853-63.
156. Hildebrandt T, Michaelides A, Mayhew M, Greif R, Sysko R, Toro-Ramos T, et al. Randomized Controlled Trial Comparing Health Coach-Delivered Smartphone-Guided Self-Help With Standard Care for Adults With Binge Eating. *Am J Psychiatry*. 2020;177(2):134-42.
157. JGL AT, Morina N, Topper M, Emmelkamp PMG. A Randomized Controlled Trial in Routine Clinical Practice Comparing Acceptance and Commitment Therapy with Cognitive Behavioral Therapy for the Treatment of Major Depressive Disorder. *Psychother Psychosom*. 2018;87(3):154-63.
158. Best MW, Milanovic M, Iftene F, Bowie CR. A Randomized Controlled Trial of Executive Functioning Training Compared With Perceptual Training for Schizophrenia Spectrum Disorders: Effects on Neurophysiology, Neurocognition, and Functioning. *Am J Psychiatry*. 2019;176(4):297-306.
159. Fallon BA, Ahern DK, Pavlicova M, Slavov I, Skritskya N, Barsky AJ. A Randomized Controlled Trial of Medication and Cognitive-Behavioral Therapy for Hypochondriasis. *Am J Psychiatry*. 2017;174(8):756-64.
160. Luby JL, Barch DM, Whalen D, Tillman R, Freedland KE. A Randomized Controlled Trial of Parent-Child Psychotherapy Targeting Emotion Development for Early Childhood Depression. *Am J Psychiatry*. 2018;175(11):1102-10.
161. Sullivan MA, Bisaga A, Pavlicova M, Carpenter KM, Choi CJ, Mishlen K, et al. A Randomized Trial Comparing Extended-Release Injectable Suspension and Oral Naltrexone, Both Combined With Behavioral Therapy, for the Treatment of Opioid Use Disorder. *Am J Psychiatry*. 2019;176(2):129-37.
162. Lynch TR, Hempel RJ, Whalley B, Byford S, Chamba R, Clarke P, et al. Refractory depression - mechanisms and efficacy of radically open dialectical behaviour therapy (Reframed): findings of a randomised trial on benefits and harms. *Br J Psychiatry*. 2020;216(4):204-12.
163. Danielson CK, Adams Z, McCart MR, Chapman JE, Sheidow AJ, Walker J, et al. Safety and Efficacy of Exposure-Based Risk Reduction Through Family Therapy for Co-occurring Substance Use Problems and Posttraumatic Stress Disorder Symptoms Among Adolescents: A Randomized Clinical Trial. *JAMA Psychiatry*. 2020;77(6):574-86.
164. Rafanelli C, Gostoli S, Buzzichelli S, Guidi J, Sirri L, Gallo P, et al. Sequential Combination of Cognitive-Behavioral Treatment and Well-Being Therapy in Depressed Patients with Acute Coronary Syndromes: A Randomized Controlled Trial (TREATED-ACS Study). *Psychother Psychosom*. 2020;89(6):345-56.
165. Dakwar E, Nunes EV, Hart CL, Foltin RW, Mathew SJ, Carpenter KM, et al. A Single Ketamine Infusion Combined With Mindfulness-Based Behavioral Modification to Treat Cocaine Dependence: A Randomized Clinical Trial. *Am J Psychiatry*. 2019;176(11):923-30.

166. Dakwar E, Levin F, Hart CL, Basaraba C, Choi J, Pavlicova M, et al. A Single Ketamine Infusion Combined With Motivational Enhancement Therapy for Alcohol Use Disorder: A Randomized Midazolam-Controlled Pilot Trial. *Am J Psychiatry*. 2020;177(2):125-33.
167. Gilbody S, Peckham E, Bailey D, Arundel C, Heron P, Crosland S, et al. Smoking cessation for people with severe mental illness (SCIMITAR+): a pragmatic randomised controlled trial. *Lancet Psychiatry*. 2019;6(5):379-90.
168. Fowler D, Hodgekins J, French P, Marshall M, Freemantle N, McCrone P, et al. Social recovery therapy in combination with early intervention services for enhancement of social recovery in patients with first-episode psychosis (SUPEREDEN3): a single-blind, randomised controlled trial. *Lancet Psychiatry*. 2018;5(1):41-50.
169. Delgadillo J, Ali S, Fleck K, Agnew C, Southgate A, Parkhouse L, et al. Stratified Care vs Stepped Care for Depression: A Cluster Randomized Clinical Trial. *JAMA Psychiatry*. 2022;79(2):101-8.
170. Holt RIG, Gossage-Worrall R, Hind D, Bradburn MJ, McCrone P, Morris T, et al. Structured lifestyle education for people with schizophrenia, schizoaffective disorder and first-episode psychosis (STEPWISE): randomised controlled trial. *Br J Psychiatry*. 2019;214(2):63-73.
171. Dennis CL, Grigoriadis S, Zupancic J, Kiss A, Ravitz P. Telephone-based nurse-delivered interpersonal psychotherapy for postpartum depression: nationwide randomised controlled trial. *Br J Psychiatry*. 2020;216(4):189-96.
172. Gilbody S, Brabyn S, Lovell K, Kessler D, Devlin T, Smith L, et al. Telephone-supported computerised cognitive-behavioural therapy: REEACT-2 large-scale pragmatic randomised controlled trial. *Br J Psychiatry*. 2017;210(5):362-7.
173. Kall A, Back M, Welin C, Aman H, Bjerkander R, Wanman M, et al. Therapist-Guided Internet-Based Treatments for Loneliness: A Randomized Controlled Three-Arm Trial Comparing Cognitive Behavioral Therapy and Interpersonal Psychotherapy. *Psychother Psychosom*. 2021;90(5):351-8.
174. Nordh M, Wahlund T, Jolstedt M, Sahlin H, Bjureberg J, Ahlen J, et al. Therapist-Guided Internet-Delivered Cognitive Behavioral Therapy vs Internet-Delivered Supportive Therapy for Children and Adolescents With Social Anxiety Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78(7):705-13.
175. Hollis C, Hall CL, Jones R, Marston L, Novere ML, Hunter R, et al. Therapist-supported online remote behavioural intervention for tics in children and adolescents in England (ORBIT): a multicentre, parallel group, single-blind, randomised controlled trial. *The Lancet Psychiatry*. 2021;8(10):871-82.
176. Costa MA, Goncalves FG, Tatton-Ramos T, Fonseca N, Schwinn JK, Alves SG, et al. A Three-Arm Randomized Clinical Trial Comparing the Efficacy of a Mindfulness-Based Intervention with an Active Comparison Group and Fluoxetine Treatment for Adults with Generalized Anxiety Disorder. *Psychother Psychosom*. 2021;90(4):269-79.
177. Reinholt N, Hvenegaard M, Christensen AB, Eskildsen A, Hjorthoj C, Poulsen S, et al. Transdiagnostic versus Diagnosis-Specific Group Cognitive Behavioral Therapy for Anxiety Disorders and Depression: A Randomized Controlled Trial. *Psychother Psychosom*. 2022;91(1):36-49.
178. Leerssen J, Lakbila-Kamal O, Dekkers LMS, Ikelaar SLC, Albers ACW, Blanken TF, et al. Treating Insomnia with High Risk of Depression Using Therapist-Guided Digital Cognitive, Behavioral, and Circadian Rhythm Support Interventions to Prevent Worsening of Depressive Symptoms: A Randomized Controlled Trial. *Psychother Psychosom*. 2022;91(3):168-79.

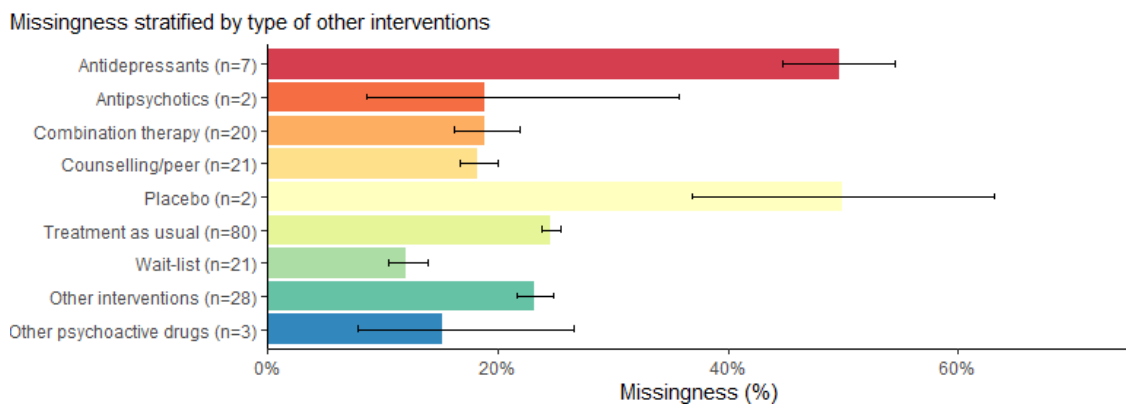
179. Norman LJ, Mannella KA, Yang H, Angstadt M, Abelson JL, Himle JA, et al. Treatment-Specific Associations Between Brain Activation and Symptom Reduction in OCD Following CBT: A Randomized fMRI Trial. *Am J Psychiatry*. 2021;178(1):39-47.
180. Schramm E, Kriston L, Elsaesser M, Fangmeier T, Meister R, Bausch P, et al. Two-Year Follow-Up after Treatment with the Cognitive Behavioral Analysis System of Psychotherapy versus Supportive Psychotherapy for Early-Onset Chronic Depression. *Psychother Psychosom*. 2019;88(3):154-64.
181. Barlow DH, Farchione TJ, Bullis JR, Gallagher MW, Murray-Latin H, Sauer-Zavala S, et al. The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders Compared With Diagnosis-Specific Protocols for Anxiety Disorders: A Randomized Clinical Trial. *JAMA Psychiatry*. 2017;74(9):875-84.
182. Bouchard S, Dumoulin S, Robillard G, Guitard T, Klinger E, Forget H, et al. Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. *Br J Psychiatry*. 2017;210(4):276-83.
183. Pot-Kolder R, Geraets CNW, Veling W, van Beilen M, Staring ABP, Gijsman HJ, et al. Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: a single-blind randomised controlled trial. *Lancet Psychiatry*. 2018;5(3):217-26.

Supplementary material 2

Missing outcome data stratified by other interventions

Fisher's exact test showed that different types of other interventions affected the proportion of missing outcome data ($p = 0.00049$). The proportion of missing data was 50% (95% CI 45% to 55%; 7 outcomes) for outcomes assessing antidepressants; 19% (95% CI 9% to 36%; 2 outcomes) for antipsychotics; 24% (95% CI 20% to 29%; 20 outcomes) for combination therapy; 18% (95% CI 17% to 20%; 21 outcomes) for counselling/peer-delivered interventions; 50% (95% CI 37% to 63%; 2 outcomes) for placebo; 26% (95% CI 25% to 28%; 20 outcomes) for psychoeducation; 25% (95% CI 24% to 26%; 80 outcomes) for treatment as usual; 12% (95% CI 10% to 14%; 21 outcomes) for wait-list; 23% (95% CI 22% to 25%; 28 outcomes) for other interventions; 15% (95% CI 8% to 27%; 3 outcomes) for other psychoactive drugs.

Supplementary figure 1



Supplementary material 3: Statistical analysis methods stratified by sensitivity analysis

	No <i>n</i> = 185	Yes <i>n</i> = 48
Primary analysis - method		
Complete case analysis	11 (5.9%)	5 (10.4%)
Full information maximum likelihood	4 (2.2%)	0 (0.0%)
Multiple imputation or similar methods	58 (31.4%)	11 (22.9%)
Other types of imputation (Last observation carried forward)	11 (5.9%)	0 (0.0%)
Other types of imputation (Single imputation)	6 (3.2%)	0 (0.0%)
Other types of imputation (Worst-case imputation)	3 (1.6%)	1 (2.1%)
Regression analyses or similar methods	50 (27.0%)	17 (35.4%)
Unclear/Not reported	42 (22.7%)	14 (29.2%)

Paper 16



OPEN ACCESS

EDITED BY
Majse Lind,
University of Florida, United States

REVIEWED BY
Sune Bo,
Psychiatric Research Unit, Psychiatry
Region Zealand, Denmark
Benjamin Hummelen,
Oslo University Hospital, Norway

*CORRESPONDENCE
Sophie Juul
✉ sophie.juul@ctu.dk

SPECIALTY SECTION
This article was submitted to
Personality Disorders,
a section of the journal
Frontiers in Psychiatry

RECEIVED 26 September 2022
ACCEPTED 30 November 2022
PUBLISHED 14 December 2022

CITATION
Juul S (2022) Improving
the methodological quality
of randomized clinical trials assessing
psychotherapy for borderline
personality disorder:
Recommendations for the future.
Front. Psychiatry 13:1053844.
doi: 10.3389/fpsy.2022.1053844

COPYRIGHT
© 2022 Juul. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Improving the methodological quality of randomized clinical trials assessing psychotherapy for borderline personality disorder: Recommendations for the future

Sophie Juul^{1,2*}

¹Copenhagen Trial Unit – Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark, ²Stolpegaard Psychotherapy Centre, Mental Health Services in the Capital Region of Denmark, Copenhagen, Denmark

KEYWORDS

personality pathology, borderline personality disorder, psychotherapy, clinical research methodology, evidence-based medicine, evidence-based health care

Introduction

Borderline personality disorder (BPD) is a debilitating psychiatric disorder affecting 1–6% of the population (1, 2). BPD symptoms include identity diffusion, affective dysregulation, and often high rates of self-harming behavior and suicide-related mortality (3, 4). Today, psychotherapy is the most widely used intervention for treatment of BPD, and the efficacy of different specialized psychotherapies for BPD has already been assessed in many randomized clinical trials (5).

Evidence-based medicine is based on the fundamental principle that high-quality research should guide practice and decision-making in the care of individual patients (6). The systematic review of randomized clinical trials is considered the gold standard when estimating intervention effects (7). However, even at the top of the evidence hierarchy, methodological challenges may still occur in the design, conduct, analysis, interpretation, and publication of trial results. These challenges from individual trials may ultimately skew the results and conclusions drawn from systematic reviews of these trials.

A Cochrane review by Storebø et al. assessing the beneficial and harmful effects of all psychological therapies for BPD was published in 2020 (5). The authors

concluded that dialectical behavioral therapy (DBT) and mentalization-based therapy (MBT) were more effective than treatment as usual on a number of patient-important outcomes (5). However, all results were based on low-quality evidence (5). The authors used the original Cochrane risk of bias tool to assess bias (8), and the Grades of Recommendations, Assessment, Development, and Evaluation (GRADE) guideline to assess the certainty of evidence (9).

Based on these methodological assessments (5) combined with my personal experiences with conducting a randomized clinical trial assessing psychotherapy for BPD (10, 11), I will in the following present some key methodological limitations along with recommendations to improve clinical research, particularly individual, parallel-group, randomized clinical trial methodology, within the field.

Balancing beneficial and harmful effects

“First do no harm” is an important injunction in all medical interventions (12). It implies that both beneficial *and* harmful effects of any intervention should always be assessed. While beneficial effects (e.g., symptom reduction and quality of life) have been thoroughly assessed in psychotherapy trials for BPD, the harmful effects of psychotherapy for BPD are currently unclear due to lack of data (5).

A universal definition of a harmful effect of psychotherapy currently does not exist, and therefore, the appropriate way of assessing harmful effects of psychotherapy can be discussed (13, 14). Relevant harmful effects for BPD patients could be self-harm, suicidality, and, more broadly, serious adverse events (15). There are many other types of harmful effects that have been reported, for example clinical deterioration or treatment non-response (14). The problem with outcomes such as response, non-response, or deterioration is that they are often based on a dichotomization of a continuous scale. For example, trialists often dichotomize the Hamilton Depression Rating Scale (HDRS) by transforming the overall HDRS score between 0 and 52 into a dichotomous score comparing responders ($\geq 50\%$ improvement) to non-responders ($\leq 50\%$ improvement) (16). However, several studies have shown that dichotomization of continuous data can bias results (17, 18). For example, a participant who improves by $\geq 50\%$ is defined as a responder, whereas a participant who improves by 49% is defined as a non-responder, and thus the apparent difference becomes inflated (16).

A more objective way of assessing harmful effects is by applying the proportion of participants with one or more serious adverse events as defined by the International Conference on Harmonization of technical requirements—Good Clinical Practice (ICH-GCP) guidelines (19) as an

outcome. According to ICH-GCP, a serious adverse event is defined as any adverse event that results in death, is life-threatening, requires hospitalization or prolongation of existing hospitalization, or results in persistent or significant disability or incapacity (19). When assessing serious adverse events with the ICH-GCP definition, two blinded investigators should independently go through patient medical records and record all events according to these criteria. An advantage of this method is that all events should be classified regardless of their etiology, and the investigators thus avoid having to differentiate a disease-specific event (e.g., a suicide or suicide-attempts) from a non-disease-related event (e.g., an accident) (20). When employing the ICH-GCP definition, a potential suicide or suicide-attempt will be included as an event, and can be categorized in a serious adverse events table. If the number of randomized participants is high enough, the non-disease-related serious adverse events will be equal in both groups, and the “true” difference between the assessed interventions on the disease-related events becomes isolated. The inclusion of serious adverse events and other methods to assess harmful effects should be improved in future trials.

Minimizing systematic errors (bias)

Causal inferences from randomized clinical trials can be undermined by errors in the design, conduct, analyses, and reporting leading to skewed estimates of the true intervention effects (bias) (21). Meta-epidemiological studies have shown that biased trial results typically overestimate beneficial effects and underestimate harmful effects of the experimental intervention (21, 22). Therefore, trials with high risk of bias may ultimately mislead clinical decisions.

In the Cochrane review of psychotherapy for BPD, all trials were assessed as at unclear or high risk of bias in a least one domain (5). The most biased domains were blinding of outcome assessors, incomplete outcome data, selective outcome reporting, and other bias (covering what the authors refer to as: attention, affiliation, or adherence bias, all of which can be understood as a bias arising from the trialists allegiance to one of the assessed interventions) (5). The authors did not assess bias associated with blinding of participants or clinicians. However, blinding of all key persons involved in a trial should be implemented whenever possible (23). Some key persons are more easily blinded than others, e.g., outcome assessors, data managers, statisticians, the data safety and monitoring committee, and the decision makers (23), and the blinding status of these persons could easily be implemented in future trials. Whether participants and clinicians could (and should) be blinded can be discussed. There are obviously practical challenges associated with delivering a

treatment without being aware of its theoretical foundation. Furthermore, there are clinical challenges as self-confirming response expectancies associated with lack of blinding can be considered an “active ingredient” of e.g., cognitive behavioral therapy (23). On the other hand, lack of blinding of participants and clinicians can be understood as a bias in line with other unblinded persons involved in a trial. For a more detailed discussion about this dilemma, please consult Juul et al. (23).

Furthermore, I want to specifically highlight the problem with incomplete outcome data bias in BPD trials. Participants with BPD, and particularly adolescent populations, can be difficult to engage in follow-up interviews, once they are enrolled in a trial. If the degree of missing data is significant, this can seriously threaten the validity of the trial results. In systematic reviews, we generally tend to assess trials at low risk of incomplete outcome data bias, if missing data is less than 5% of all randomized participants (24). The 5% cutoff is not definitive, but it is a coarse rule of thumb. When comparing this rule with the missing data values sometimes exceeding 33% in psychotherapy trials with BPD (25), we as trialists must consider very carefully, how to better design trials at lower risks of incomplete outcome data bias. Multiple imputation, and other methods to statistically account for missing data can be used, if missing data is impossible to ignore. However, if the proportions of missing data are very large (for example, more than 40%) on important outcomes, no statistical method can solve that problem, and the trial results may then be considered hypothesis-generating only (24).

Usually, missing data will affect outcomes that require an action from the participant, e.g., if they need to fill out questionnaires or participate in a clinical interview. One solution could be to minimize the use of such outcomes to make participating in a trial less overwhelming for patients, while also considering more objective outcomes that can be retrieved from, e.g., national patient registries or medical records by (blinded) research personnel; i.e., hospitalization, suicide-related behavior, and employment status.

Standardization of PICOs

In the Cochrane review of psychotherapy for BPD, most trials included small sample sizes, with the number of participants ranging from 7 to 151 (5). Furthermore, heterogeneity was observed in the selection of interventions and outcomes, hindering the pooling of effect estimates. The small sample sizes and the inability to pool results led to imprecise effect estimates (5). Therefore, future BPD trials should preferably be larger, and aim to assess similar PICOs (participants,

interventions, comparators, and outcomes), so that pooling trials in a systematic review becomes less influenced by heterogeneity.

Participants

The need to replicate existing trials is an aim that pragmatically conflicts with the wish for more personalized psychotherapy. BPD is a heterogeneous disorder with many potential subgroups (26). The more subgroups that are identified (e.g., patients with different levels of symptom severity at baseline), the more participants are needed, either if we should start multiple trials for every subgroup, or perform subgroup analyses embedded within a trial. While we should aim to conduct trials with different types of BPD patients to cover the whole spectrum of patient characteristics, we should still be mindful of which conclusions can be drawn from subgroup analyses that are exploratory by nature (8).

Interventions and comparators

Several standardized interventions specifically developed to treat BPD and BPD-related symptoms have already been developed, including DBT, MBT, schema therapy (ST), and transference focused psychotherapy (TFP) (5). While developing new and potentially improved interventions may seem promising, more trials assessing the effects of the already existing BPD-interventions are still needed, if we want to confirm or reject intervention effects on several important outcomes (5). Furthermore, an adequate description of the assessed trial interventions and comparators is required for trialists to design replication trials and for clinicians to reliably implement interventions (27, 28). Both the experimental and the control interventions need to be described in detail (29). To improve reporting of interventions, the Template for Intervention Description and Replication checklist and guide (TIDieR) has been developed (28), which can be used as an addition to the CONSORT guideline for reporting of trials (30).

Outcomes

Existing BPD trials have used a wide range of outcome measures, which makes it difficult to synthesize data in systematic reviews (31). Furthermore, the selected outcomes do not always adequately reflect BPD patient experience (31). This calls for a discussion of which outcomes are the most patient-important. A standard set of patient-reported outcomes for the International

Classification of Diseases—11th version (ICD-11) personality disorder classification has recently been proposed by the International Consortium for Health Outcomes Measurement (ICHOM) multidisciplinary working group (32). However, these recommendations did not cover more objective outcomes. Agreeing on a core outcome set (COS) for future trials of psychotherapy for BPD is needed and will improve the development of evidence-based treatment guidelines in the future (31). In my opinion, the field could move forward by including both continuous outcomes like quality of life and symptom severity, and also dichotomous outcomes like hospitalization, self-harm, suicide or suicide-attempts and employment status. However, the development of a COS is highly needed and should involve key stakeholders including researchers and methodologists but also patient organizations, relatives, clinicians, funders, and administrators, and should follow strict development guidelines provided by the Core Outcome Measures in Effectiveness Trials (COMET) initiative (33).

Improving conflict of interest disclosures and retrieve unpublished data

Psychotherapy research has long struggled with the potential bias of trialists who believe in the superiority of one psychological intervention over another, a phenomenon typically referred to as attention bias or researcher allegiance (34). Researcher allegiance is a heterogeneous construct ranging from developing the treatment manual to advocating for it to contributing to a related disease model to, ultimately, conducting a trial showing results in favor of the new experimental intervention (35). Examples of financial conflicts of interest in psychotherapy trials are when trialists also have financial gains from e.g., professional trainings of that particular intervention, books, therapy manuals, courses, speaker's fees, paid advisory positions, grants etc. (34).

One way of assessing the potential impact of conflicts of interest is for systematic reviewers of randomized clinical trials to carefully look for signs of publication bias when performing a meta-analysis. Publication bias refers to the publication or non-publication of research findings, depending on the direction of the results (36). Trialists with a strong allegiance to an experimental intervention may decide not to publish the trial results, if the results do not comply with their expectations. Pre-registration of trials at registries such as www.clinicaltrials.gov are now required when launching a trial. Pre-registration of trials minimizes study publication bias (when trials showing negative or no effect are not published) and

selective outcome reporting bias (when trialists fail to report unfavorable data, include only a subset of data analyzed, or change or omit the outcome of interest in the interest of statistical significance) (37). However, while pre-registration of trials are a methodological safeguard, both publication bias and selective outcome reporting bias may still occur.

Assessment of publication bias can be performed by visually inspecting funnel plots (36) and by statistically testing the funnel plot asymmetry using various tests depending on the outcome of interest (36). A funnel plot is a scatter plot of the effect estimates from individual trials against some measure of each trial's size or precision (usually the standard error) (38).

In the Cochrane review of psychotherapies for BPD, the inspection of the funnel plot suggested potential bias (small asymmetry) (5). Furthermore, the authors assessed almost a third of the included trials as being influenced by "other bias" such as attention, affiliation, or adherence bias, indicating a potential conflict of interest in most included trials.

To control for researcher allegiance on a trial level, trialists should aim to implement blinding of all possible key persons involved in data collection, analysis, interpretation, and dissemination of the trial results (23). Furthermore, disclosures of financial conflicts of interest for all contributing authors of published trial reports should be improved (34), and, on a systematic review level, authors should carefully try to retrieve unpublished data. Moreover, all trials need to be transparently registered before launch and all trial data transparently registered after analysis and publication (39).

Conclusion

While we have already come a long way in both designing and implementing structured, manualized psychotherapies for BPD, and in assessing their effects in randomized clinical trials, the current evidence is still restricted by substantial methodological limitations. In this paper, I have highlighted some key methodological limitations and suggested the following recommendations: *balancing beneficial and harmful effects, minimizing systematic errors (bias), standardization of PICO, improving conflict of interest disclosures, and retrieving unpublished data*. Improving these methodological limitations can lead to us potentially identifying more evidence-based treatments, which may ultimately result in better care for BPD patients.

The reader should take into account the potential limitations of this paper. The recommendations are based on my own opinion, rather than originating from an international

consortium of experts. Thus, the recommendations are in no way exhaustive, but may serve as a stepping stone for further improvement in the field.

Evidence-based medicine is the conscientious, explicit, and judicious use of best evidence in making decisions about the care of individual patients (6). Evidently, clinicians should, by default, offer psychotherapy supported by the best available evidence. It is our job as both clinicians and trialists to continuously make that evidence as trustworthy as possible.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

The author received salaries from the Copenhagen Trial Unit, Centre for Clinical Intervention Research and a research grant from TrygFonden A/S (grant no. 123488).

References

- Grant B, Chou S, Goldstein R, Huang B, Stinson F, Saha T, et al. Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. (2008) 69:533.
- Torgersen S, Kringlen E, Cramer V. The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry*. (2001) 58:590–6.
- Cheng A, Mann A, Chan K. Personality disorder and suicide: A case-control study. *Br J Psychiatry*. (1997) 170:441–6.
- Black D, Blum N, Pfohl B, Hale N. Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *J Personal Dis*. (2004) 18:226–39.
- Storebø O, Stoffers-Winterling J, Völlm B, Kongerslev M, Mattivi J, Jørgensen M, et al. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. (2020) 2012:CD005652.
- Sackett D, Rosenberg W, Gray J, Haynes R, Richardson W. Evidence based medicine: what it is and what it isn't. *BMJ*. (1996) 321:71–2. doi: 10.1136/bmj.312.7023.71
- Garattini S, Jakobsen J, Wetterslev J, Bertelè V, Banzi R, Rath A, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med*. (2016) 32:13–21. doi: 10.1016/j.ejim.2016.03.020
- Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. (2011). Available online at: <https://handbook-5-1.cochrane.org> (Accessed November 25, 2022).
- Guyatt G, Oxman A, Vist G, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. (2008) 336:924–6. doi: 10.1136/bmj.39489.470347.AD
- Juul S, Lunn S, Poulsen S, Sørensen P, Salimi M, Jakobsen J, et al. Short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality disorder: a protocol for a randomized clinical trial. *Trials*. (2019) 20:196.
- Juul S, Simonsen S, Poulsen S, Lunn S, Sørensen P, Bateman A, et al. Detailed statistical analysis plan for the short-term versus long-term mentalization-based therapy for outpatients with subthreshold or diagnosed borderline personality

Acknowledgments

The author would like to thank Prof. Christian Gluud and Ph.D. student Caroline Kamp Jørgensen for their valuable feedback on the first version of this manuscript.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- disorder randomized clinical trial (MBT-RCT). *Trials*. (2021) 22:497. doi: 10.1186/s13063-021-05450-y
- Sokol D. "First do no harm" revisited. *BMJ*. (2013) 347:f6426. doi: 10.1136/bmj.f6426
- Dimidjian S, Hollon S. How would we know if psychotherapy were harmful? *Am Psychol*. (2010) 65:21.
- Linden M. How to define, find and classify side effects in psychotherapy: From unwanted events to adverse treatment reactions. *Clin Psychol Psychother*. (2013) 20:286–96. doi: 10.1002/cpp.1765
- Juul S, Poulsen S, Lunn S, Sørensen P, Jakobsen J, Simonsen S. Short-term versus long-term psychotherapy for adult psychiatric disorders: a protocol for a systematic review with meta-analysis and trial sequential analysis. *Syst Rev*. (2019) 8:169. doi: 10.1186/s13643-019-1099-0
- Jakobsen J, Gluud C, Kirsch I. Should antidepressants be used for major depressive disorder? *BMJ Evid Based Med*. (2019) 25:130–6. doi: 10.1136/bmjebm-2019-111238
- Altman D, Royston P. The cost of dichotomising continuous variables. *BMJ*. (2006) 332:1080. doi: 10.1136/bmj.332.7549.1080
- Kirsch I, Moncrieff J. Clinical trials and the response rate illusion. *Contemp Clin Trials*. (2007) 28:348–51.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. *ICH Harmonised Guideline: Integrated Addendum to ICH E6 (R1): Guideline for Good Clinical Practice (ICH-GCP)*. (2015). Available online at: <https://ichgcp.net/da> (accessed November 25, 2022).
- Jakobsen J, Wetterslev J, Gluud C. Considerations on the strengths and limitations of using disease-related mortality as an outcome in clinical research. *BMJ Evid Based Med*. (2019) 26:127–30. doi: 10.1136/bmjebm-2018-111154
- Wood L, Egger M, Gluud L, Schulz K, Jüni P, Altman D, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ*. (2008) 336:601–5. doi: 10.1136/bmj.39465.451748.AD

22. Savović J, Turner R, Mawdsley D, Jones H, Beynon R, Higgins J, et al. Association between risk-of-bias assessments and results of randomized trials in Cochrane reviews: the ROBES meta-epidemiologic study. *Am J Epidemiol.* (2018) 187:1113–22. doi: 10.1093/aje/kwx344
23. Juul S, Gluud C, Simonsen S, Frandsen FW, Kirsch I, Jakobsen JC. Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. *BMJ Evid Based Med.* (2021) 26:109. doi: 10.1136/bmjebm-2020-111407
24. Jakobsen J, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol.* (2017) 17:162. doi: 10.1186/s12874-017-0442-1
25. Linehan M, Korslund K, Harned M, Gallop R, Lungu A, Neacsiu A, et al. Dialectical behavior therapy for high suicide risk in individuals with borderline personality disorder: a randomized clinical trial and component analysis. *JAMA Psychiatry.* (2015) 72:475–82. doi: 10.1001/jamapsychiatry.2014.3039
26. Zimmerman M. Heterogeneity of borderline personality disorder: do the number of criteria met make a difference? *J Pers Dis.* (2007) 21:615.
27. Boutron I, Moher D, Altman D, Schulz K, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med.* (2008) 148:295–309. doi: 10.7326/0003-4819-148-4-200802190-00008
28. Hoffmann T, Glasziou P, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ.* (2014) 348:g1687. doi: 10.1136/bmj.g1687
29. Faltinsen E, Todorovic A, Bruun L, Hróbjartsson A, Gluud C, Kongerslev M, et al. Control interventions in randomised trials among people with mental health disorders. *Cochrane Database Syst Rev.* (2022) 4:MR000050.
30. Schulz K, Altman D, Moher D, Consort Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* (2010) 152:726–32.
31. National Institute for Health and Care Excellence (NICE). *Borderline personality disorder: recognition and management. Clinical guideline [CG78].* (2009). Available online at: <https://www.nice.org.uk/guidance/cg78> (accessed September 19, 2022).
32. Prevolnik Rupel V, Jagger B, Fialho L, Chadderton L, Gintner T, Arntz A, et al. Standard set of patient-reported outcomes for personality disorder. *Qual Life Res.* (2021) 30:3485–500.
33. Prinsen C, Vohra S, Rose M, King-Jones S, Ishaque S, Bhaloo Z, et al. COMET initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a “core outcome set.” *Trials.* (2014) 15:247. doi: 10.1186/1745-6215-15-247
34. Cristea I, Ioannidis J. Improving disclosure of financial conflicts of interest for research on psychosocial interventions. *JAMA Psychiatry.* (2018) 75:541–2.
35. Munder T, Flückiger C, Gerger H, Wampold B, Barth J. Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *J Couns Psychol.* (2012) 59:631. doi: 10.1037/a0029571
36. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 *Cochrane.* (2019). Available online at: <https://training.cochrane.org/handbook> (accessed November 25, 2022).
37. Mitra-Majumdar M, Kesselheim A. Reporting bias in clinical trials: Progress toward transparency and next steps. *PLoS Med.* (2022) 19:e1003894. doi: 10.1371/journal.pmed.1003894
38. Sterne J, Sutton A, Ioannidis J, Terrin N, Jones D, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ.* (2011) 343:d4002. doi: 10.1136/bmj.d4002
39. Skoog M, Saarimäki J, Gluud C, Sheinin M, Erlendsson K, Aamdal S. *Transparency and registration in clinical research in the Nordic countries.* Oslo: NordForsk (2015). p. 1–108.