

**“The Bergen 4-Day OCD Treatment – B4DT” in Hamburg:
Die Effektivität der Behandlung von Zwangsstörung mittels
hochfrequenter Exposition**

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Medizin (Dr. med.)

an der

Medizinischen Fakultät der Universität Hamburg

Betreuung durch Prof. Dr. Lena Jelinek

vorgelegt von

Anna-Marie Josefine Serve

aus

Frankfurt am Main

2025

Betreuer:in / Gutachter:in der Dissertation: Prof. Dr. Lena Jelinek

Gutachter:in der Dissertation: Prof. Dr. Bernd Löwe

Vorsitz der Prüfungskommission: Prof. Dr. Bernd Löwe

Mitglied der Prüfungskommission: PD Dr. Marion Eisele

Mitglied der Prüfungskommission: PD Dr. Sigrun Vehling

Datum der mündlichen Prüfung: 08.07.2026

Inhaltsverzeichnis

1. Zielsetzung und Arbeitshypothese.....	1
2. Einleitung	3
3. Theoretischer Hintergrund	6
3.1. Zwangsstörung	6
3.2. Behandlung von Zwangsstörungen	12
3.3. Gegenwärtige Probleme in der Therapie von Menschen mit Zwangsstörungen	14
3.4. B4DT	16
4. Methoden.....	20
4.1. Design	20
4.2. Proband:innen	21
4.3. Behandlung	22
4.4. Instrumente und Datenerhebung	23
4.5. Kriterien für klinisch signifikante Verbesserung und Klassifikation von Therapieansprechen und Remission.....	29
4.6. Hypothesen und Statistische Analyse.....	29
5. Ergebnisse	33
5.1. Stichprobe	33
5.2. Adhärenz und Drop-out.....	33
5.3. Zwangssymptomatik.....	36
5.4. Klinische Relevanz	40
5.5. Depressivität.....	41
5.6. Allgemeines Funktionsniveau.....	42
5.7. Zufriedenheit.....	42
6. Diskussion.....	45
6.1. Einordnung der Ergebnisse in den Kontext bestehender Literatur	45
6.2. Wirkmechanismen und Besonderheiten des B4DT	46

6.3.	Unterschiede zu früheren Studien und mögliche Einflussfaktoren	48
6.4.	Limitationen und Stärken der Studie	51
6.5.	Implikationen für Forschung und Praxis	52
7.	Zusammenfassung	55
8.	Literaturverzeichnis.....	58
9.	Abkürzungsverzeichnis	66
10.	Abbildungsverzeichnis	67
11.	Tabellenverzeichnis	68
12.	Vorveröffentlichung und Erklärung des Eigenanteils.....	69
13.	Eidesstattliche Versicherung	70
14.	Danksagung	71
15.	Anhang	72
15.1.	Tabelle 1	72
15.2.	Vorveröffentlichung.....	84

1. Zielsetzung und Arbeitshypothese

Ziel der vorliegenden Arbeit war die Durchführung des Bergen 4-Day Treatment (B4DT) bei Patient:innen mit Zwangsstörung und die anschließende Evaluation der Wirksamkeit und Akzeptanz der Behandlung. B4DT ist ein neuartiges Behandlungsformat zur Behandlung von Menschen mit Zwangsstörungen mittels hochfrequenter Exposition. In veröffentlichten Auswertungen der Behandlung zeigt sich eine durchschlagende Effektivität (Hansen et al. 2018, Launes et al. 2019a). Die vorliegende Arbeit sollte mit der Durchführung einer unizentrischen, prospektiven, nicht-randomisierten Pilotstudie als erste Implementierung von B4DT im deutschen Versorgungskontext einen Beitrag zu der Frage leisten, ob in Norwegen erzielte Ergebnisse in anderen Versorgungssystemen replizierbar sind und ob B4DT die Versorgung von Menschen mit Zwangsstörungen verbessern könnte.

Es wurden 33 Patient:innen in Hamburg mit dem B4DT behandelt. Primärer Outcome war die fremdbeurteilte Zwangssymptomatik eine Woche nach der Behandlung sowie drei Monate nach der Behandlung. Sekundäre Ergebnisparameter waren die Veränderungen der selbstberichteten Zwangssymptome, der depressiven Symptome, des allgemeinen Funktionsniveaus sowie die Zufriedenheit der Teilnehmenden mit der Behandlung. Arbeitshypothese der vorliegenden Studie war, dass B4DT bei Patient:innen mit Zwangsstörung im deutschen Versorgungskontext zu einer signifikanten Reduktion der Zwangssymptomatik, gemessen anhand der Yale-Brown Obsessive-Compulsive Scale (Y-BOCS, Goodman et al. (1989), deutsche Version von Hand und Büttner-Westphal (1991)) führt und Ansprech- und Remissions- sowie Drop-out-Raten erreicht, die mit den bisherigen norwegischen Studien vergleichbar sind. Darüber hinaus wird die Hypothese aufgestellt, dass das B4DT depressive Symptomatik reduziert und die globale Funktionsfähigkeit der Patient:innen bessert. Die Hypothesen werden im Methodenteil im Abschnitt 5.6. *Statistische Analyse* ausgeführt.

Die in dieser Dissertation dargestellten Daten stammen aus einer Pilotstudie zur Anwendung des B4DT in Hamburg. Ergebnisse dieser Studie wurden bereits in der Publikation „Exploring the Peaks and Potholes: Understanding positive and negative effects of concentrated exposure treatment for obsessive-compulsive disorder“ (Jelinek et al. 2024b) im Journal of Obsessive-Compulsive and Related Disorders veröffentlicht (siehe Anhang). Die vorliegende Arbeit stellt die Daten umfassend dar, ordnet sie systematisch ein und diskutiert sie vertieft im Kontext aktueller Forschungsergebnisse.

2. Einleitung

Zwang ist vielgestaltig und hat allein in der Medizin viele verschiedene Formen. In der Psychiatrie werden Menschen „zwangseingewiesen“, also stationär untergebracht, um vor Eigen- oder Fremdaggressivität zu schützen. Weitere Beispiele sind Zwangsernährung, Zwangsmedikation und zwangsweise Quarantänemaßnahmen – eine Person wird durch jemanden oder etwas gezwungen etwas zu tun und in ihren Handlungsoptionen eingeschränkt.

Zwang hat keine einheitliche Definition und in der Philosophie wird der Begriff breit diskutiert. Eine notwendige Bedingung, um von Zwang zu sprechen, lässt sich jedoch bestimmen: Der Eingriff in die Freiheit einer Person. „Ist jemand gezwungen, etwas zu tun oder zu unterlassen, ist er nicht (gänzlich) frei (...)“ (Kellerwessel 2012, S. 7).

In den genannten Beispielen kommt der Zwang von außen und überwindet den Willen der „gezwungenen“ Person mit der Intention, sie oder die Gesellschaft zu schützen und Gefahren abzuwenden – die Gefahr der Ausbreitung einer ansteckenden Erkrankung durch Quarantänemaßnahmen, den drohenden Suizid oder eine mögliche Fremdgefährdung durch eine Unterbringung in der Psychiatrie nach richterlichem Beschluss. In welchen Situationen und in welchem Ausmaß die Anwendung äußeren Zwangs durch Personen oder Institutionen rechtlich und ethisch zulässig ist, ist eine komplexe Fragestellung im Spannungsfeld zwischen individueller Freiheit und kollektiver Verantwortung.

Kommt der Zwang von innen, aus einer Person selbst, in Form von Impulsen, Gedanken oder Vorstellungen, die das eigene Verhalten und Erleben bestimmen und die subjektive Freiheit der Person einschränken, so stellt sich die ebenso komplexe Frage, wo die Grenze zwischen Normalität und dysfunktionalem Zwang liegt. Ab einem gewissen Ausmaß von Leidensdruck und funktionaler Beeinträchtigung durch Vorstellungen oder Impulse wird in der Psychiatrie und Psychologie von Zwangsstörungen gesprochen. Bei den Betroffenen von Zwangsstörungen manifestieren sich wiederkehrende, intrusive Gedanken (Obsessionen) und/oder ritualisierte Verhaltensweisen (Kompulsionen), die so viel Zeit im Alltag in Anspruch nehmen und so belastend sein können, dass es den Betroffenen durch eine hohe

funktionale Beeinträchtigung nicht möglich ist, ein normales Leben zu führen (Hollander et al. 2016, Voderholzer et al. 2022). Diesem Verlust von Freiheit durch innere Zwänge versuchen seit dem 19. Jahrhundert Therapeut:innen eine wirksame Behandlung entgegenzusetzen (Oberbeck et al. 2013). Obwohl bereits wirksame Behandlungskonzepte mithilfe kognitiver Verhaltenstherapie existieren, gelang es Gerd Kvale und Bjarne Hansen 2018 mit ihrem neuen Behandlungsformat, dem B4DT, die Therapie von Menschen mit Zwangsstörungen neu zu denken und unter die fünfzig einflussreichsten Personen für das amerikanische Gesundheitswesen des Time-Magazins zu gelangen (Oaklander 2018).

Die Innovation des Behandlungsformats liegt zum einen in einer verkürzten Behandlungsdauer von vier Tagen, die selbst nach häufig langjährigen und erfolglosen Vorbehandlungen eine signifikante Reduktion von Zwangsstörungen ermöglicht (Kvale et al. 2020, Deutsche Psychotherapeuten Vereinigung e.V. 2021, Stengler et al. 2013). Zum anderen basiert die Behandlung auf dem Prinzip der Exposition mit Reaktionsverhinderung (ERP), unterscheidet sich jedoch von traditionellen Formaten durch ihre hohe Intensität, enge therapeutische Begleitung und gruppengestützte Struktur. Über vier Behandlungstage arbeiten die Patient:innen in Kleingruppen und mit individueller therapeutischer Betreuung an Psychoedukation, Expositionsübungen und Plänen für die Integration dieser in den Alltag. Ziel ist es, das zugrunde liegende „Zwangsprojekt“, die aufrechterhaltende Intention hinter den Zwangshandlungen, zu erkennen und loszulassen. Die Teilnehmenden sollen eine veränderte Haltung gegenüber Angst und Unsicherheit entwickeln und positive Erfahrungen der Selbstwirksamkeit im aktiven Verfolgen eigener Ziele und Werte machen (Jelinek et al. 2025). Die norwegischen Auswertungen des Behandlungsformats belegen eine bemerkenswerte Effektivität der Behandlung. Die von den Entwickler:innen durchgeführten Studien ergaben, dass mehr als 90 % der Proband:innen auf die Therapie ansprachen (Kvale et al. 2018). In einem Podcast-Interview gibt Gerd Kvale, gemeinsam mit Bjarne Hansen Entwicklerin des Formats, an, dass die Ursprünge des B4DT auf der Bevölkerungsstruktur Norwegens sowie rein praktischen Überlegungen basieren (Ralph 2019). In Deutschland liegt die Bevölkerungsdichte bei 240, in Norwegen hingegen bei 15 Einwohner:innen pro Quadratkilometer (Statistisches

Bundesamt (Destatis) 2025). Dies führt dazu, dass die wöchentliche spezialisierte Behandlung für die Patient:innen in Norwegen oftmals mit einem hohen Aufwand verbunden oder gar nicht möglich ist. Die Entwickler hatten daher den einmaligen Kontakt zu den Patient:innen intensiviert und dabei positive Erfahrungen gemacht (Ralph 2019). Die vorliegende Arbeit zielte darauf ab, die Wirksamkeit und Umsetzbarkeit des Bergen 4-Day Treatment im Kontext der deutschen Versorgung zu evaluieren.

3. Theoretischer Hintergrund

3.1. Zwangsstörung

Epidemiologie

Mit Lebenszeitprävalenzraten von 1 bis 3% gehören Zwangsstörungen zu den häufigsten psychischen Erkrankungen (Fawcett et al. 2020, Kessler et al. 2012, Ruscio et al. 2010). Die 12-Monats-Prävalenz von Zwangsstörungen beträgt in Deutschland 3.6 % (Jacobi et al. 2014). Die Prävalenz für subklinisches Auftreten von Zwangsgedanken und -handlungen wird noch höher geschätzt (Ruscio et al. 2010). Das Auftreten der Symptome erfolgt oft schleichend. Viele Menschen leiden viele Jahre lang unter Zwangsvorstellungen und Zwängen bevor sie diese als psychische Erkrankung erkennen, sodass eine Diagnostik und Behandlung möglich ist (Hollander et al. 2016, Skoog und Skoog 1999, Stewart et al. 2004).

Symptomatik

Die Kernsymptome einer Zwangsstörung (englisch: Obsessive Compulsive Disorder, OCD) sind Zwangsgedanken und Zwangshandlungen. Zwangsgedanken sind wiederkehrende Gedanken oder Bilder, die sich den Betroffenen aufdrängen. Sie werden als eigene, jedoch als irrationale, störende oder übertriebene Gedanken wahrgenommen. Diese Gedanken lösen bei den Betroffenen oft Stress, Unruhe, hohe Anspannung oder Angst aus. Die Inhalte solcher Zwangsgedanken sind individuell und sehr verschieden. Zwangsgedanken beziehen sich häufig auf Verschmutzung und Krankheit, auf die Angst, Fehler zu machen oder Schaden anzurichten sowie auf Themen wie Sexualität, Gewalt, Symmetrie, Ordnung, Religion und Moral (Schäfer et al. 2024, Stein et al. 2020). Zwangsgedanken sind häufig schambesetzt (Voderholzer 2022).

Zwangshandlungen sind Verhaltensweisen, die sich wiederholen oder ritualisiert sind und häufig starren Regeln folgen. Inhaltlich sind sie oft mit den vorhandenen Zwangsgedanken verknüpft. Die Zwangshandlungen dienen dazu, den Zwangsgedanken zu „neutralisieren“, was zu einer kurzzeitigen Erleichterung führt bis der

Zwangsgedanke erneut auftritt. Die Betroffenen fühlen sich zu den Zwangshandlungen als Reaktion auf einen Zwangsgedanken gezwungen. Die Angst oder Anspannung, die durch den Zwangsgedanken ‚Ich könnte unbemerkt im Straßenverkehr jemandem geschadet haben.‘ verursacht wird, könnte beispielsweise kurzfristig reduziert werden, indem das Auto auf Kratzer kontrolliert wird. Der Gedanke einer Kontamination könnte beispielsweise kurzfristig durch die Zwangshandlung des Händewaschens neutralisiert werden. Die Zwangshandlungen sind so individuell und verschieden wie die Zwangsgedanken. Häufig treten Waschwänge (z.B. Händewaschen, Wäsche waschen, Duschen, Zähneputzen), Kontrollzwänge (z.B. die Kontrolle elektrischer Geräte, Fenster oder Türen), ein bestimmtes Ordnen von Dingen (z.B. Symmetrie am Arbeitsplatz) sowie wiederholte gedankliche Handlungen auf (z.B. Zählen oder ein gedankliches Durchgehen von Abläufen oder Situationen) (Schäfer et al. 2024). Obwohl die Zwangshandlungen kurzzeitig eine Erleichterung für die Betroffenen darstellen können, werden sie langfristig als unangenehm empfunden. Oft sind ein hoher Zeitaufwand und Leidensdruck mit den Zwangshandlungen verknüpft (Bobes et al. 2007, S. 293-303). Viele Menschen mit Zwangsstörungen zeigen außerdem Vermeidungsverhalten, um das Auftreten von Zwangsgedanken und -handlungen zu verringern. Bei einer zwanghaften Angst vor Kontamination kann beispielsweise das Berühren von Gegenständen vermieden werden. Vermeidungsverhalten führt bei Betroffenen zu starken Einschränkungen im Alltag und zu sozialer Isolation (Rosa et al. 2012).

Diagnostik

Während in der International Classification of Diseases 10th Revision (ICD-10) Zwangsstörungen gemeinsam mit phobischen Störungen und Angststörungen im Kapitel „Neurotische, Belastungs- und somatoforme Störungen“ gelistet werden, ist die Zwangsstörung in den aktuellen internationalen Diagnoseklassifikationskriterien, dem Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) und der International Classification of Diseases 11th Revision (ICD-11), erstmals als eigene Kategorie unter dem Begriff „Zwangsstörungen und verwandte Störungen“ charakterisiert (American Psychiatric Association, A. P. A. 2025, Endres et al. 2025, Stein et al. 2016). Eingeschlossen in die Gruppe der Zwangsstörungen sind in ICD-11 neben

der Zwangsstörung die körperdysmorphe Störung, pathologisches Horten, die olfaktorische Referenzstörung, die Hypochondrie sowie körperbezogene repetitive Verhaltensstörungen wie Trichotillomanie (das zwanghafte Auszupfen von Haaren) und Dermatillomanie (ein wiederholtes Kratzen, Zupfen oder Quetschen der Haut). Wiederkehrende, nicht gewollte Gedanken und repetitives Verhalten, welches mit den Gedanken in Zusammenhang steht, sind das zentrale psychopathologische Merkmal aller Störungen der Kategorie (Stein et al. 2016).

Neu in beiden aktuellen Auflagen der internationalen Klassifikationssysteme ist der Stellenwert der Einsichtsfähigkeit der Betroffenen als Kriterium zur Spezifizierung der Zwangsstörung. Im DSM-5 ist dies durch eine dreigestufte Zusatzcodierung „mit guter oder angemessener Einsicht“, „mit wenig Einsicht“ und „mit fehlender Einsicht/wahnhafte Überzeugungen“ realisiert, im ICD-11 wird mit einer zweigestuften Subkodierung zwischen „guter oder angemessene Einsicht“ und „wenig oder keiner Einsicht“ unterschieden. Neben differentialdiagnostischen Vorteilen dieser Kategorisierung wurden Unterschiede in der Wirksamkeit von Pharmakotherapie und Psychotherapie je nach vorhandener Einsicht der Betroffenen gezeigt (Backenstraß 2022), sodass die spezifischere Diagnostik unter Einbeziehung der Einsichtsfähigkeit zu einer individuelleren und zielgerichteten Therapie beitragen soll.

Verlauf

Meist manifestieren sich Zwangsstörungen bereits in der Jugend oder im frühen Erwachsenenalter und begleiten die Betroffenen viele Jahre bis Jahrzehnte (Ruscio et al. 2010). In den Vereinigten Staaten liegt das durchschnittliche Alter bei Auftreten der Zwangsstörung bei 19.5 Jahren (Ruscio et al. 2010), in einer internationalen Studie von Brakouliasa et al. (2017) bei 16.9 Jahren. Frauen sind im Erwachsenenalter etwas häufiger betroffen als Männer, während bei Männern häufiger als bei Frauen bereits im frühen Kindes- und Jugendalter die ersten Zwänge auftreten (Kessler et al. 2012, Ruscio et al. 2010). Etwa ein Viertel der Betroffenen haben die ersten Symptome vor dem 14. Lebensjahr, ein Ausbruch nach dem 35. Lebensjahr ist ungewöhnlich (Grant et al. 2007). Im Mittel vergehen 17 Jahre bevor Menschen mit Zwangsstörungen die korrekte Diagnose und Therapie erhalten (Elsouri et al. 2024). Menschen, bei denen die ersten

Symptome in einem frühen Alter auftreten (Early Onset), haben eine längere Latenzzeit bis zum Kontakt mit professioneller Hilfe als jene mit späterem Beginn der Zwangsstörung (Stengler et al. 2013). Je später der Zugang zu Behandlung gegeben ist, desto höher ist der Grad an Chronifizierung der Zwangserkrankung und die Einschränkung des sozialen Lebens der Betroffenen (Grant et al. 2007, Stengler et al. 2013). Unbehandelt verläuft die Zwangsstörung chronisch (Endres et al. 2025, Skoog und Skoog 1999). Dabei kann sich die Intensität der Symptomatik und der Inhalt der Zwänge verändern. Stress führt häufig zu einer Zunahme der Symptomatik (Voderholzer et al. 2022). Menschen mit Zwangsstörungen haben eine verkürzte Lebenserwartung und eine im Vergleich zur Allgemeinbevölkerung deutlich erhöhte Rate an Suizidversuchen (Elsouri et al. 2024).

Pathogenese

Die Pathogenese von Zwangsstörungen ist ungeklärt und Gegenstand aktueller Forschung. Es werden verschiedene Störungsmodelle mit teils kognitiv-behavioralen, psychodynamischen und systemischen Ansätzen diskutiert (Voderholzer et al. 2022). Studien beschreiben Einflüsse durch spezifische Lebensereignisse, durch die Prägung durch Erziehungsweisen der Eltern, durch Persönlichkeitsmerkmale, durch Lernen und Bewertung von Erfahrungen, durch psychologische Vulnerabilitätsfaktoren, dysfunktionale (metakognitive) Überzeugungen sowie durch klassische und operante Konditionierung (Endres et al. 2025, Stein et al. 2020).

Verwandte ersten Grades einer Person mit Zwangsstörungen weisen ein sechsfach höheres Risiko auf, eine Zwangsstörung zu entwickeln (Endres et al. 2025, Schäfer et al. 2024). Aus der neurobiologischen Perspektive konnte eine gesteigerte Aktivität in bestimmten Hirnarealen festgestellt werden, sowie ein Ungleichgewicht des serotonergen, dopaminergen und glutamatergen Systems und eine Fehlregulation der Basalganglien (Endres et al. 2025). Zusammengenommen legt dies nahe, dass molekulargenetische Prädispositionen und biologische Faktoren zu der Entstehung der Zwangsstörung beitragen (Stein et al. 2020).

Folgen

Zwangsstörungen haben für die Betroffenen erhebliche Auswirkungen auf das soziale sowie das Berufsleben. Die Metaanalyse von Bobes et al. (2007) zeigt, dass die Lebensqualität von Menschen mit Zwangsstörungen im Vergleich zur Allgemeinbevölkerung erheblich beeinträchtigt ist, vor allem in Bereichen der sozialen Funktionalität, der emotionalen Rolle und der psychischen Gesundheit. Je schwerer dabei die Zwangsstörung, insbesondere die Zwangsgedanken, desto größer die Einschränkung der Lebensqualität der Betroffenen. Die soziale Anpassung ist zusätzlich beeinträchtigt und die Lebensqualität erniedrigt, wenn die Zwangsgedanken der Betroffenen sexuell oder religiös sind, oder Hortungssymptome bestehen, oder eine der folgenden Komorbiditäten vorliegt: Depression, posttraumatische Belastungsstörung oder eine Essstörung (Rosa et al. 2012).

Bei Erkrankungsbeginn im jungen Erwachsenenalter haben Zwangsstörungen häufig große Auswirkungen auf die Ausbildung und die berufliche Entwicklung der Betroffenen (Voderholzer et al. 2022). Auch die Arbeitsfähigkeit der Menschen mit Zwangsstörungen ist häufig eingeschränkt. Dies geht mit finanziellen Belastungen für die Betroffenen und ihre Angehörigen einher (Voderholzer et al. 2022). Auch Angehörige und Pflegende von Menschen mit Zwangsstörungen sind in ihrer Lebensqualität beeinträchtigt. Die Belastung der Angehörigen ist unter anderem von der Schwere der Erkrankung, der Funktionalität der Menschen mit Zwangsstörungen und der Relation der Angehörigen zu den Betroffenen abhängig (Hollander et al. 2016).

Zwangsstörungen waren 2007 neben unipolaren Depressionen, Schizophrenie, Alkoholkonsum und bipolaren Störungen unter den fünf psychischen Störungen, die weltweit die meiste Erwerbsunfähigkeit verursachen (Bobes et al. 2007, S. 293-303). Gesamtgesellschaftlich entstehen durch Zwangsstörungen entsprechend hohe Kosten (Hollander et al. 2016). Kochar et al. (2023) modellierten die Kosten durch OCD-Patient:innen in Großbritannien und verzeichneten auch mit konservativen Analysen eine erhebliche wirtschaftliche jährliche Belastung von über fünf Milliarden Pfund. Diese entsteht einerseits durch direkte Ausgaben des

Gesundheitsdienstleisters für ambulante und stationäre Behandlungen und andererseits indirekt durch die Einbußen der Produktivität bei fehlender Erwerbsfähigkeit sowie durch Ausgaben für private Therapien (Kochar et al. 2023).

Komorbiditäten

Bei Menschen mit Zwangsstörungen besteht eine hohe Last an Komorbiditäten. Die Lebenszeitprävalenzraten mindestens einer weiteren psychiatrischen Störung zum Zeitpunkt der Diagnose der Zwangsstörungen rangieren in verschiedenen Studien zwischen 65% und 92% (Brakouliasa et al. 2017, Ruscio et al. 2010). In einer repräsentativen Umfrage unter Erwachsenen in den USA fanden Ruscio et al. (2010) als häufigste Begleiterkrankungen bei Zwangsstörungen Angststörungen (75,8 %), gefolgt von affektiven Störungen (63,3 %), Impulskontrollstörungen (55,9 %) und Substanzabhängigkeiten (38,6 %). Klinische Studien ergeben ähnliche Prävalenzen (Lochner et al. 2014). Beschrieben werden auch Assoziationen mit Essstörungen und zwanghaften Persönlichkeitsstörungen sowie mit Tic-Störungen und körperdysmorphen Störungen. Menschen mit Zwangsstörungen haben zudem ein höheres Risiko eine Schizophrenie zu entwickeln (Endres et al. 2025). Die angegebenen Prävalenzraten der Komorbiditäten unterliegen in der Literatur einer starken Varianz (Voderholzer et al. 2022). Menschen mit Zwangsstörungen können auch durch sekundäre dermatologische Symptome auffällig werden. Zum einen durch eine Komorbidität mit körperbezogenen repetitiven Verhaltensstörungen wie Trichotillomanie, Dermatotillomanie, Onychotillomanie und Onychophagie, die im DSM-5 und ICD-11 unter den Zwangsspektrumstörungen bzw. unter den verwandten Störungen einer Zwangsstörung geführt werden, zum anderen durch atopische Dermatitis oder chronisch irritative Kontaktekzeme durch exzessives Waschen der Hände oder anderer Körperteile bei Wasch- und Putzzwängen (Voderholzer et al. 2022).

Komorbiditäten können den Krankheits- und Behandlungsverlauf stark beeinflussen: Die Anzahl der komorbiden Angst- und affektiven Störungen korreliert positiv mit dem Schweregrad der Zwangsstörungssymptome und der Krankheitsdauer (Lochner et al. 2014). Die Lebensqualität ist bei Menschen mit

Zwangsstörungen und Komorbiditäten stärker beeinträchtigt als ohne Komorbiditäten (Rosa et al. 2012). Studienergebnisse zeigen Hinweise, dass depressive Symptome bei Menschen mit Zwangsstörungen den langfristigen Behandlungserfolg einschränken, unter anderem durch eine erhöhte Rate an Therapieabbrüchen (Hansen et al. 2019, Leeuwerik et al. 2019). Komorbiditäten sollten daher in der Auswahl und Gestaltung einer individuellen Behandlung Berücksichtigung finden (Stein et al. 2020).

3.2. Behandlung von Zwangsstörungen

Einführung

Die ersten deutschsprachigen Beschreibungen von Zwangsphänomenen entstanden in der zweiten Hälfte des 19. Jahrhunderts (Griesinger 1868, Oberbeck et al. 2013). Erstmals als eigenes Störungsbild beschrieben wurden Zwangsstörungen in den 1890er Jahren von Sigmund Freud, der den Ursprung der Zwangssymptome in unterdrückten und abgewehrten Vorstellungen sah, die meist erotische Fantasien oder Triebe beinhalteten (Freud 1952, Oberbeck et al. 2013, Voderholzer et al. 2022). Der Affekt und die Erregung über diese erotischen Inhalte übertrügen sich bei einem einmaligen Erlebnis, das den Patient:innen oft nicht Erinnerungswürdig sei, auf eine andere, häufig banale Vorstellung. Diese banale Vorstellung trete fortan mit dem Affekt der unterdrückten erotischen Inhalte und der eigenen Empörung über solche schambesetzten Gedanken als Zwangsvorstellung immer wieder auf und sei belastend. Sigmund Freud ordnete die Zwangsstörung 1894 als „Abwehr-Neuropsychose“ ein (Freud 1952, S.59-74). Seine Werke wurden posthum 1952 von Anna Freud gesammelt herausgegeben.

Seine Theorien führten zu psychodynamischen Krankheitsmodellen der Zwangsstörung und zu wenig wirksamen Therapien mit Antipsychotika, langfristigen Hospitalisierungen und Sedierungen (Meyer 1966, Salkovskis et al. 2018, S. 67-86). Bis in die 1960er Jahre waren die Behandlungen für Menschen mit Zwangsstörungen häufig langwierig, invasiv und frustrierend (Meyer 1966). In den 1960er Jahren folgten erste Fallberichte über die erfolgreiche Behandlung von

Menschen mit Zwangsstörungen durch behaviorale Behandlungsansätze (Meyer 1966). Auf Grundlage der Beobachtung von Tieren mit ritualisierten Verhaltensweisen entstanden neue Krankheitsmodelle, die den Zwang als eine Art erlerntes Vermeidungsverhalten durch Konditionierung verstanden. Die damaligen Ideen, in der Behandlung die Zwangshandlungen zu unterbinden und negative Erwartungen zu reduzieren, werden als Wegbereiter der heutigen kognitiven Ansätze der Verhaltenstherapie gesehen (Salkovskis et al. 2018, S. 68). Zur gleichen Zeit entstanden Behandlungsansätze mit einer Konfrontation mit angstbesetzten Situationen. Die Wirksamkeit dieser verhaltenstherapeutischen Ansätze konnte vielfach nachgewiesen und als evidenzbasierter Behandlungsstandard fortlaufend weiterentwickelt werden. Es fanden so auch kognitive Techniken Einzug in die Behandlung von Zwangsstörungen (Salkovskis et al. 2018, S. 68).

Heutige psychotherapeutische Praxis

Heute ist das psychotherapeutische Verfahren der Wahl nach der aktuellen S3-Leitlinie die kognitive Verhaltenstherapie (KVT), ein Überbegriff für verschiedene kognitive und metakognitive Techniken, die in Kombination angewandt werden (Voderholzer et al. 2022). Im Zentrum der kognitiven Verhaltenstherapie bei Zwangsstörungen steht die Exposition mit Reaktionsmanagement (ERM) (Voderholzer et al. 2022). Die Betroffenen werden dabei therapeutisch unterstützt und begleitet, sich systematisch und geplant mit gefürchteten Zwangsinhalten oder Situationen zu konfrontieren und dabei keine Zwangshandlungen als Reaktion durchzuführen. Auf der kognitiven Ebene geht es um die Konsequenzen und um dysfunktionale Überzeugungen während der Exposition. Je nach Zwangsinhalten können die Expositionen sehr verschieden aussehen.

Die Wirksamkeit der KVT mit ERM ist vielfach nachgewiesen. Metaanalysen von randomisierten kontrollierten Studien haben übereinstimmend eine signifikante Verbesserung der Zwangssymptomatik durch KVT mit großer Effektstärke (Hedges's $g = 1.39; 1.33$) gezeigt (Olatunji et al. 2013, Öst et al. 2015). Auch wenn ein Vergleich der Ergebnisse durch die große Variabilität an Kriterien erschwert wird, zeigen Ansprechraten von 62-68 % eine gute Wirksamkeit der KVT (Öst et al. 2015).

Medikamentöse Therapie

Neben der KVT mit ERM empfiehlt die S3-Leitlinie auch eine medikamentöse Therapie von Zwangsstörungen mittels Psychopharmaka (Voderholzer et al. 2022). Die kognitive Verhaltenstherapie kann allein zum Einsatz kommen, bei fehlendem Ansprechen ist die Kombination mit einer medikamentösen Therapie eine mögliche Augmentation. Bei hoher Symptomlast kann auch mit einer parallelen Therapie begonnen werden (Voderholzer et al. 2022). Eine alleinige medikamentöse Therapie wird empfohlen, wenn die Psychotherapie abgelehnt wird oder aufgrund der Schwere der Symptome nicht durchgeführt werden kann. Außerdem kommen Psychopharmaka ohne psychotherapeutische Begleitung bei mangelnden Ressourcen für Psychotherapie sowie bei langen Wartezeiten zum Tragen (Voderholzer et al. 2022). Eingesetzt werden dabei selektive Serotonin-Wiederaufnahme-Inhibitoren (SSRIs) oder Clomipramin, welches der Medikamentenklasse der Trizyklischen Antidepressiva angehört. Die Wirksamkeit ist vielfach nachgewiesen (Voderholzer et al. 2022). Skapinakis et al. (2016) zeigten in einer großen Netzwerk-Metaanalyse, dass SSRIs eine Reduktion des Y-BOCS-Scores um 3.49 (95%-Konfidenzintervall (KI) -5.12 bis -1.81) Punktwerte unter Einschluss von 37 Studien mit insgesamt 3158 Patienten erzielten. Clomipramin erreichte eine Reduktion des Y-BOCS-Scores um 4.72 (95%-KI -6.85 bis -2.60) Punktwerte in der Analyse von 13 Studien mit insgesamt 831 Proband:innen. Der Wirksamkeitsunterschied zwischen Clomipramin und der Klasse der SSRIs war dabei nicht signifikant (Mittelwertdifferenz = -1.23; (95%-KI -3.41 bis -0.94) (Skapinakis et al. 2016). Auch innerhalb der Klasse der SSRIs konnte keine Überlegenheit eines bestimmten Wirkstoffs festgestellt werden. Die Auswahl des Medikaments erfolgt individuell je nach Nebenwirkungsprofil der einzelnen Medikamente (Voderholzer et al. 2022).

3.3. Gegenwärtige Probleme in der Therapie von Menschen mit Zwangsstörungen

Mit der Studie zur Gesundheit Erwachsener in Deutschland und dem Zusatzmodul für psychische Gesundheit (DEGS1-MH) wurde im Jahr 2011 eine repräsentative nationale Kohorte (N = 5317) zur Inanspruchnahme von Behandlungen bei

bestehenden psychischen Erkrankungen befragt (Jacobi et al. 2014). Dabei ergab sich für Zwangsstörungen ein Anteil von 18.2 % der Betroffenen, die eine Behandlung in Anspruch nahmen. Eine niedrigere Behandlungsrate zeigte sich nur für substanzbezogene Störungen mit einer 12-Monats-Behandlungsprävalenz von 15.6 % (Mack et al. 2014, Deutsche Psychotherapeuten Vereinigung e.V. 2021).

Was sind Gründe für dieses Therapiedefizit und die damit verbunden hohen Belastungen für das Individuum und die Gesellschaft?

Auf der einen Seite nehmen Menschen mit Zwangsstörungen professionelle Hilfe häufig erst spät in Anspruch (Deutsche Psychotherapeuten Vereinigung e.V. 2021): Der erste Kontakt zu Hilfsangeboten findet im Mittel sieben bis zehn Jahre nach dem ersten Auftreten von Symptomen statt (Stengler et al. 2013, Voderholzer et al. 2015). Als Gründe dafür werden real erlebte Stigmatisierung und Angst vor Stigmatisierung, Scham und das fehlende Wissen über adäquate professionelle Hilfe diskutiert (Ansari et al. 2020, Stengler et al. 2013). Zudem werden die Zwangssymptome oft frühzeitig wahrgenommen, jedoch von Betroffenen und Angehörigen häufig bagatellisiert oder einer Norm zugeordnet, anstatt als Frühzeichen einer Störung erkannt zu werden (Stengler et al. 2013).

Auf der anderen Seite gibt es in Deutschland aktuell einen großen Mangel an Therapieplätzen und spezialisierten Angeboten (Voderholzer et al. 2022). Menschen mit Zwangsstörungen nehmen häufig hohe Wartezeiten auf sich, um einen Therapieplatz zu erhalten. Dabei neigen gerade lange unbehandelte Zwangsstörungen zu Chronifizierung und bringen größere funktionale Einschränkungen sowie Einbußen der Lebensqualität mit sich (Öst et al. 2015).

Dieser Behandlungslücke begegnet die gegenwärtige psychotherapeutische Praxis nur unzureichend. Die Therapien, die durchgeführt werden, entsprechen häufig nicht den Leitlinien. Obwohl KVT mit ERM die wirksamste zurzeit bekannte Therapie der Zwangsstörung ist (Skapinakis et al. 2016, Voderholzer et al. 2022), führt ein substanzieller Anteil von behandelnden Psychotherapeut:innen in Deutschland keine Expositionen mit ihren Patient:innen durch (Jelinek et al. 2024a, Külz et al. 2010, Moritz et al. 2019, Voderholzer et al. 2015). Und auch wenn Menschen mit Zwangsstörungen sich in psychotherapeutischer Therapie befinden, ist nicht

gesichert, dass diese hilft. Viele Menschen sprechen nicht vollständig auf die Therapie an oder brechen die Behandlung frühzeitig ab. Für Patient:innen ist die ERM-Therapie durch die Konfrontation mit gefürchteten Reizen schwierig, die Therapie-Abbruchquote wird auf 11 bis 19% geschätzt (Leeuwerik et al. 2019, Olatunji et al. 2013).

3.4. B4DT

Einführung

Um die Wirksamkeit von Therapien für Zwangsstörungen zu verbessern, Therapieabbrüche zu minimieren und praktikable Behandlungsoptionen für einen großen Bedarf zu schaffen wurden vermehrt Therapieformate untersucht, die in einem engeren Zeitrahmen stattfanden und Wirksamkeitsstudien durchgeführt.

Dabei verglichen Abramowitz et al. (2003) eine intensive Therapie mit 15 zweistündigen Expositionssitzungen täglich über drei Wochen mit einer weniger intensiven Therapie mit zwei zweistündigen Sitzungen wöchentlich und einem kurzen wöchentlichen Telefonkontakt über acht Wochen. Beide Behandlungsvarianten führten zu einer signifikanten Verbesserung der Zwangssymptome mit großen Effektstärken am Ende der Therapie, wobei die Effektstärke der intensiven Therapie etwas größer war als die der weniger intensiven Therapie ($d = 2.70$ versus $d = 1.80$). Eine klinisch signifikante Verbesserung zeigte sich in der intensiv behandelten Gruppe bei 85 % der Stichprobe ($N = 20$) nach der Behandlung, bei der weniger intensiv behandelten bei nur 55 % der Probandinnen ($N = 20$).

Auch die Kognitive Verhaltenstherapie mit Exposition im Gruppenformat wurde von verschiedenen Arbeitsgruppen untersucht (Voderholzer et al. 2022). Dabei konnte auch für das gruppentherapeutische Setting eine Wirksamkeit von KVT gezeigt werden. Die Effektstärken waren etwas geringer als im einzeltherapeutischen Setting, jedoch waren Therapieabbrüche seltener (Haland et al. 2010, Jaurrieta et al. 2008). Diskutiert wird als Grund hierfür eine Verringerung der Angst vor Stigmatisierung sowie gegenseitige Motivierung zur Exposition. Mit dem B4DT,

einer konzentrierten Expositionstherapie an vier aufeinanderfolgenden Tagen, konnten in norwegischen Studien vielversprechende Ergebnisse erzielt werden (Hansen et al. 2018, Hansen et al. 2019, Havnen et al. 2014, Havnen et al. 2017). Die expositionsbasierte Intervention wird in einer Gruppe von drei bis sechs Patient:innen bei einem Eins-zu-eins-Verhältnis zu den Therapeut:innen durchgeführt und soll durch die Kombination von Gruppen- und Einzeltherapie die Vorteile beider Therapiemodalitäten verbinden.

Wirksamkeit und Studienlage

Die bisher veröffentlichten Studien zur Wirksamkeit des B4DT bei Zwangsstörungen sind in Tabelle 1 (Anhang) dargestellt. Aufgeführt werden alle Veröffentlichungen, die die Auswirkung des Behandlungsformats auf die Zwangssymptomatik untersuchten. Diese wurde in allen Veröffentlichungen mithilfe der Y-BOCS erhoben. In unkontrollierten Studien in Norwegen erreichten die Entwickler des Formats bei 90 % der Patient:innen eine relevante klinische Veränderung nach der Behandlung (Kvale et al. 2018). Fast 70 % sind langfristig in Remission (Hansen et al. 2018, Hansen et al. 2019). Auch die erste randomisierte kontrollierte Studie von Launes et al. (2019a), die B4DT mit einer Selbsthilfe-Intervention sowie einer Wartelistenbedingung verglich, zeigte große Effektstärken (Cohen's $d = 3.86$ gegenüber Warteliste, $d = 2.57$ gegenüber Selbsthilfe) und konnte mit 62.5 % der Proband:innen in Remission direkt nach der Intervention und 81.3 % der Proband:innen in Remission sechs Monate nach der Intervention die Daten aus den vorangehenden unkontrollierten Studien bestätigen. Zu beachten ist, dass es sich um einen recht kleinen Stichprobenumfang ($N = 48$) mit 16 Proband:innen pro Studienarm handelte. Betrachtet man die Randomisierung der Gruppen, so zeigen sich größere Unterschiede in den bestehenden Komorbiditäten und im Alter der erstauftgetretenen Symptome. Zudem wurde den Proband:innen in den Kontrollgruppen in Aussicht gestellt, nach Abschluss der Studie an der Intervention B4DT teilzunehmen. Über einen Nachbeobachtungszeitraum von vier Jahren wurden die Studienpopulationen aus den unkontrollierten Studien Havnen et al. (2014), Havnen et al. (2017) nachbeobachtet (Hansen et al. 2019). Nach vier Jahren

zeigten über 90 % der Proband:innen ein immer noch konsistentes Ansprechen (22.1 %) oder befanden sich in Remission (72.7 %) nach internationalen Konsensus-Kriterien. Hansen et al. (2019) verglichen in der genannten Veröffentlichung die Ergebnisse mit Effektivitätsstudien nicht-konzentrierter Expositionsbehandlung und fanden eine signifikant höhere Reduktion des Y-BOCS-Scores sowie signifikant höhere Remissions-Raten. Davíðsdóttir et al. (2019) implementierten B4DT erstmals außerhalb Norwegens in Reykjavík, Island. Die Resultate waren konsistent mit den in Norwegen publizierten Daten. Auch hier ist die Aussagekraft durch eine fehlende Kontrollgruppe und ein Stichprobenumfang von 19 Proband:innen begrenzt, dennoch belegen die Ergebnisse die Durchführbarkeit und Akzeptanz des Formats außerhalb des norwegischen Gesundheitssystems unter Einbeziehung isländischer Therapeut:innen. In Finnland führten Silver et al. (2023) eine unkontrollierte Pilotstudie durch. Auch hier zeigte sich eine große Effektivität der Behandlung mit einer signifikanten Reduktion der Zwangssymptome und der depressiven Symptome sowie einer Reduktion von Schlafproblemen, Angst und Problemen in der psychosozialen Funktionsfähigkeit. Die Effektstärken in Finnland sind mit Cohen's $d = 1.91$ von T0 zu T1 und $d = 1.93$ von T0 zu T2 etwas geringer im Vergleich zu den norwegischen und isländischen Studienergebnissen (Launes et al. (2020): $d = 4.60$ von T0 zu T1, Davíðsdóttir et al. (2019): $d = 4.26$ von T0 zu T1).

Zusammenfassend lässt sich festhalten, dass das B4DT als ein hochwirksames, gut akzeptiertes und längerfristig stabiles Behandlungsformat anzusehen ist. Besonders hervorzuheben sind die kurze Behandlungsdauer und die niedrigen Abbruchraten. Die in Island und Finnland erhobenen Daten deuten darauf hin, dass das B4DT auch in unterschiedlichen gesundheitssystemischen Kontexten erfolgreich implementiert werden kann. Die Datenlage ist bisher begrenzt durch das Überwiegen unkontrollierter Studien, kleine Stichprobengrößen, einen maximalen Beobachtungszeitraum von vier Jahren sowie die Konzentration der Daten auf Norwegen, Island und Finnland.

Empirische Grundlage für den deskriptiven Vergleich

Für den deskriptiven Vergleich der Ergebnisse der vorliegenden Pilotstudie mit norwegischen Publikationen wurden die folgenden Ansprech- und Remissionsraten aus früheren Studien herangezogen. Eine Woche nach der Intervention (T1) sprachen 16-19% der Teilnehmenden auf die Intervention an und über 70% befanden sich in Remission (Hansen et al. 2018, Kvale et al. 2018, Launes et al. 2019b). Drei Monate nach der Intervention (T2) sprachen 9-16% der Teilnehmenden auf die Intervention an und etwa 70% befanden sich in Remission (Hansen et al. 2018, Kvale et al. 2018, Launes et al. 2019b).

Diese Werte dienen als Referenz und werden in **Hypothese H5** im Methodenteil aufgegriffen.

4. Methoden

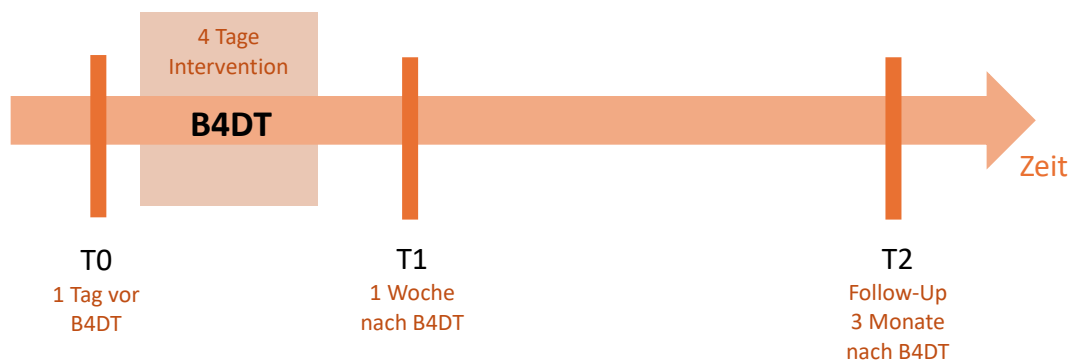
Die Dissertation basiert auf der Pilotstudie zum Bergen 4-Day Treatment in Hamburg. Teile der Ergebnisse wurden bereits in der Publikation „Exploring the Peaks and Potholes: Understanding positive and negative effects of concentrated exposure treatment for obsessive-compulsive disorder“ (Jelinek et al. 2024b) veröffentlicht (siehe Anhang). Die hier dargestellten Analysen und Auswertungen wurden eigenständig von der Doktorandin durchgeführt.

4.1. Design

An der Klinik und Poliklinik für Psychiatrie und Psychotherapie des Universitätsklinikums Hamburg-Eppendorfs wurde eine prospektive, monozentrische, einarmige Pilotstudie zur Evaluation der Durchführbarkeit und der Effektstärke des B4DT auf die Zwangssymptomatik durchgeführt. Die 33 Teilnehmenden wurden in elf Gruppen mit je drei bis vier Teilnehmenden mit dem B4DT am UKE behandelt. Die Datenerhebung erfolgte im Längsschnittdesign einen Tag vor der Intervention (T0), eine Woche nach der Intervention (T1) sowie drei Monate nach der Intervention (T2) (siehe Abbildung 1). Das Studiendesign wurde von der Lokalen Psychologischen Ethikkommission (LPEK) am Zentrum für Psychosoziale Medizin des UKE genehmigt (LPEK-0512) und im Deutschen Register Klinischer Studien (DRKS00030022) präregistriert.

Abbildung 1

Zeitpunkte der Datenerhebung



Anmerkungen. B4DT = Bergen 4-Day Treatment

4.2. Proband:innen

Einschluss- und Ausschluss-Kriterien

Einschlusskriterien der durchzuführenden prospektiv nicht-randomisierten Interventionsstudie waren ein Alter zwischen 18 und 75 Jahren und die Diagnose einer Zwangsstörung, gesichert mittels Mini-International Neuropsychiatric Interview (M.I.N.I. Sheehan et al. (1998)), Version 7 nach DSM-5. Ausschlusskriterien waren akute Suizidalität, psychotische oder manische Symptome aktuell oder in der Vorgeschichte, Substanzmissbrauch und die Einnahme von Benzodiazepinen. Die Teilnehmer:innen mussten mindestens sechs Wochen medikamentenfrei sein oder eine stabile Medikation erhalten. Ein weiteres Ausschlusskriterium waren mangelnde Deutsch- oder Englischkenntnisse. Eine Gruppe erhielt die Therapie auf englisch, alle anderen Gruppen auf deutsch.

Rekrutierung

Die Rekrutierung der Teilnehmenden erfolgte über Mitarbeitende der Station für Angst- und Zwangsstörungen am Universitätsklinikum Hamburg Eppendorf (UKE), über die psychiatrisch-psychotherapeutische Ambulanz des UKEs sowie über soziale Medien. In Vorgesprächen wurden die Ein- und Ausschlusskriterien abgesichert, die Teilnehmenden erhielten Informationen zum Ablauf der Intervention, zur Freiwilligkeit, zur Datenverarbeitung, zum Datenschutz sowie der Pseudonymisierung und Anonymisierung der Daten. Die Teilnehmenden gaben ihre informierte Zustimmung mittels einer unterzeichneten Einverständniserklärung. Die Festlegung der angestrebten Stichprobengröße von 33 Proband:innen erfolgte ohne Durchführung einer formalen Fallzahlplanung unter Berücksichtigung der statistischen Teststärke, da in den internationalen Vorläuferstudien sehr große Effektgrößen berichtet wurden, die auch bei deutlich kleineren Stichproben eine ausreichende Power zur Detektion signifikanter Effekte erwarten ließen. Die gewählte Fallzahl leitete sich daher primär aus den Stichprobengrößen anderen Studien zum B4DT ab (Havnen et al. 2014, Havnen et al. 2017, Launes et al. 2019b, Thorsen et al. 2020b).

4.3. Behandlung

Beschreibung der Intervention

Das B4DT umfasst vier aufeinanderfolgende Behandlungstage mit insgesamt etwa 28 Therapiestunden (à 50 Minuten), die von einem interdisziplinären Team in Einzel- und Gruppensitzungen durchgeführt werden. Der erste Tag des B4DT dient der Psychoedukation und dem Kennenlernen der Technik für die folgenden Expositionen sowie der Planung von Expositionen. Wichtig ist hier unter anderem die LET-Technik („LEan into The anxiety“), die den Teilnehmenden helfen soll, sich den Ängsten zu nähern bzw. sich „in die Angst hineinzulehnen“ und sich das eigene Vermeidungsverhalten bewusst zu machen (Hansen et al. 2018, Havnen et al. 2013).

Am zweiten und dritten Tag führen die Teilnehmenden, unterstützt von Therapeut:innen, individuell angepasste Expositionen durch und werden motiviert, die erlernten Techniken anzuwenden. Die Gruppe trifft sich mehrmals täglich zu festgesetzten Zeiten, beispielsweise zum gemeinsamen Mittagessen. Die individuellen Expositionen werden in der Gruppe nachbesprochen. Am Nachmittag des dritten Tages werden Angehörige und Freunde der Proband:innen zu einem psychoedukativen Treffen eingeladen.

Der vierte und letzte Tag dient der Zusammenfassung des Gelernten und der individuellen Planung weiterer Ziele und möglicher Expositionen für einen Zeitraum von drei Wochen nach der Behandlungswoche im Alltag.

Zur nachhaltigen Stabilisierung der Therapieergebnisse erfolgt drei Monate nach der Behandlung eine Auffrischungssitzung der Teilnehmenden mit einer/einem Therapeuten/Therapeutin, in der die Erfahrungen der Proband:innen besprochen werden, erlernte Strategien wiederholt und für die selbstständige Anwendung gefestigt werden (Hansen et al. 2019, Schäfer et al. 2024).

Therapeut:innen-Team

Zu Beginn des Projekts schulten die Entwickler:innen des Behandlungsformats die Therapeut:innen in Hamburg. Später fanden Schulungen untereinander statt. Drei oder

vier Therapeut:innen betreuten eine Gruppe im Eins-zu-eins-Verhältnis zu den Proband:innen.

4.4. Instrumente und Datenerhebung

Ablauf der Datenerhebung

Der Zeitraum der Rekrutierung und Datenerhebung für die vorliegende Pilotstudie erstreckte sich von Oktober 2022 bis Oktober 2023.

Nach dem Vorgespräch und informierter Zustimmung wurden die Teilnehmenden zu der Intervention an vier aufeinanderfolgenden Tagen im tagesklinischen Setting eingeladen. Eine Woche vor Beginn der Interventionswoche erhielten die Teilnehmenden einen Online-Fragebogen zur Erhebung des Ausgangswerts (T0) des Obsessive-Compulsive Inventory - Revised (OCI-R, Foa et al. 2002; deutsche Version Gönner et al. 2007), des Patient Health Questionnaire (PHQ-9, Kroenke et al. 2001), der Rosenberg-Self-Esteem-Scale (RSES, Rosenberg 1965, S. 338), des Brief Experimental Avoidance Questionnaire (BEAQ, Gamez et al. 2014) sowie die Skala zur Messung der Allgemeinen Selbstwirksamkeitserwartung (SWE, Schwarzer und Jerusalem 1999). Einen Tag vor Beginn der Intervention kamen die Teilnehmenden für das tagesklinische Aufnahme-procedere sowie für die Erhebung der Y-BOCS, des M.I.N.I. sowie des GAF (Global Assessment of Functioning, Saß et al. 2003) (T0) in die Klinik. Das M.I.N.I., die Y-BOCS und das GAF wurden zum Zeitpunkt T0 immer in persona erhoben.

Eine Woche nach der Behandlungswoche wurde eine weitere Erhebung der Y-BOCS und des GAF in der Regel in persona sowie des OCI-R, PHQ-9, RSES, BEAQ, SWE und des Fragebogens zur Messung der Patientenzufriedenheit (ZUF-8, Schmidt et al. 1989) online durchgeführt (T1).

Die Teilnehmenden wurden zu einer therapeutischen Auffrischungssitzung (Booster-Sitzung) ohne Exposition sowie zu einer weiteren Follow-Up Datenerhebung mit äquivalenten Instrumenten zu T1 drei Monate nach der Behandlungswoche eingeladen (T2). Zum Zeitpunkt T2 erhielten die Teilnehmenden einen Link zu einer Online-Befragung per Mail, die die Proband:innen selbstständig zu Hause ausfüllten.

Die Y-BOCS und das GAF zu den Zeitpunkten T1 und T2 wurden teilweise telefonisch oder via Videokonferenz erhoben, insbesondere bei Proband:innen, die nicht wohnhaft in Hamburg waren. Alle Instrumente mit Erhebungsmodalität und Erhebungszeitpunkt sind in Tabelle 2 dargestellt. Die Wahl der Zeitpunkte T1 und T2 orientierten sich an den Studien der Initiatoren des B4DT aus Norwegen, um die Vergleichbarkeit der Daten gewährleisten zu können.

Tabelle 2
Erhebungsmodalitäten und Erhebungszeitpunkte der Instrumente

Instrument	Erhebungsmodalität	Zeitpunkt		
		T0	T1	T2
M.I.N.I.	In persona	x		
Soziodemographische Daten	Online	x		
Y-BOCS	In persona oder telefonisch	x	x	x
GAF	In persona oder telefonisch	x	x	x
OCI-R	Online	x	x	x
PHQ-9	Online	x	x	x
RSES	Online	x	x	x
BEAQ	Online	x	x	x
SWE	Online	x	x	x
ZUF-8	Online		x	x

Anmerkungen. T0 vor der Intervention, T1 eine Woche nach der Intervention, T2 beim Follow-Up 3 Monate nach der Intervention. BEAQ = Brief Experimental Avoidance Questionnaire, GAF = Global Assessment of Functioning, M.I.N.I. = Mini-International Neuropsychiatric Interview, OCI-R = Obsessive-Compulsive Inventory – Revised, PHQ-9 = Patient Health Questionnaire, RSES = Rosenberg Self Esteem Scale, SWE = Skala zur Allgemeinen Selbstwirksamkeitserwartung, Y-BOCS = Yale-Brown Obsessive-Compulsive Scale, ZUF-8 = Fragebogen zur Patientenzufriedenheit.

Rater:innen

Die Rater:innen wurden für die Erhebung des M.I.N.I., der Y-BOCS und des GAF jeweils geschult und von der Studienleitung supervidiert. Die Rater:innen zu den verschiedenen Zeitpunkten konnten bei den einzelnen Proband:innen variieren, waren jedoch immer unabhängig vom Therapeut:innen-Team der jeweiligen Interventionsgruppe. Zur Bestimmung der Übereinstimmung zwischen den beiden unabhängigen Rater:innen wurde für die Y-BOCS-Gesamtpunktzahl der Intraclass-Correlation-Coefficient (ICC) berechnet.

Auch wenn im Rahmen der Pilotstudie eine Vielzahl an Instrumenten erhoben wurde (vgl. Tabelle 2), werden nachstehend ausschließlich jene Verfahren näher erläutert, deren Ergebnisse zur Prüfung der Hypothesen (s. 5.6.) dieser Dissertation herangezogen wurden.

Yale-Brown Obsessive-Compulsive Scale (Y-BOCS)

Die Erfassung der Zwangssymptomatik erfolgte zu den Zeitpunkten T0, T1, und T2 mit der Y-BOCS, dem Goldstandard zur Quantifizierung von Zwangsstörungen. Die Y-BOCS ist ein halbstrukturiertes Interview, das in zwei Teile gegliedert ist. Im ersten Teil wird der Inhalt der Zwangsgedanken und -handlungen anhand einer Checkliste mit 70 Items aus 15 thematischen Untergruppen von Zwängen erfragt. Mit Hilfe der Symptomcheckliste sollen explorativ alle Zwangsinhalte erfasst werden, die zum Zeitpunkt der Befragung und in der Vergangenheit eine Rolle gespielt haben. Anschließend wird von dem Rater/ der Raterin eine Liste der Zwangsinhalte erstellt. In dieser Liste werden in hierarchischer Reihenfolge die drei Zwangsgedanken, Zwangshandlungen und Arten von Vermeidungsverhalten notiert, die für die befragte Person zum Zeitpunkt der Befragung am meisten im Vordergrund stehen.

Im zweiten Teil wird der Schweregrad der Zwangsgedanken und -handlungen anhand von 21 Items erfasst. Die Items werden auf einer fünfstufigen Skala von 0 ("nicht vorhanden") bis 4 ("extrem") bewertet. Aus der Summe von zehn Items ergibt sich ein Gesamtscore von maximal 40.

Die ersten sechs Items (1, 1b, 2, 3, 4, 5) beziehen sich auf die Zwangsgedanken und können auch getrennt von den Zwangshandlungen (Items 6, 6b, 7, 8, 9 und 10) betrachtet werden. Die Items erfassen den Zeitaufwand für die Zwänge, die Beeinträchtigung in der alltäglichen sozialen und beruflichen Funktion, den Leidensdruck, den Widerstand gegen die Zwänge und die Kontrolle über Zwänge.

Der Cut-off-Wert für eine klinische Relevanz liegt bei einem Gesamtscore von 16, wenn sowohl Zwangsgedanken als auch Zwangshandlungen berücksichtigt werden. Werden entweder Gedanken oder Handlungen isoliert betrachtet, liegt der Cut-off-Wert bei zehn (Hand und Büttner-Westphal 1991). Werte zwischen 0 und 7 werden als subklinische, 8 bis 15 als milde, 16 bis 23 als moderate, 24 bis 31 als schwere und 32 bis

40 als extreme Zwangssymptomatik eingestuft (Marques et al. 2010, S.48). Das Interview dauert zwischen 30 und 60 Minuten.

Die mittlere interne Konsistenz (Cronbach's α) für den Y-BOCS-Gesamtwert wurde zwischen $\alpha = 0.69$ (Woody et al. 1995) und $\alpha = 0.89$ (Goodman et al. 1989) geschätzt. Die Interrater-Reliabilität wurde von Woody et al. (1995) sowohl für den Gesamtwert als auch für die Subskalen der Zwangsgedanken und Zwangshandlungen mit mindestens $r = 0.89$ als exzellent eingestuft. Bei isolierter Betrachtung der ersten zehn Items beträgt der niedrigste Wert der Interrater-Reliabilität auch bei Goodman et al. (1989) $r = 0.80$. Die mittlere Test-Retest-Reliabilität für den Gesamtscore wird in einer Metaanalyse von López-Pina et al. (2015) mit $r = 0.848$ angegeben. Die konvergente Validität mit zwei unabhängigen Skalen zur Erfassung von Zwangsstörungen wurde mit Werten von $r = 0.67$ und $r = 0.74$ als hoch eingestuft (Hand und Büttner-Westphal 1991). Die diskriminante Validität ist im Vergleich zu einer Depressions- und Angstskala eingeschränkt (Woody et al. 1995).

Obsessive-Compulsive Inventory - Revised (OCI-R)

Mit dem OCI-R nahmen die Teilnehmenden zu den Zeitpunkten T0, T1 und T2 im Rahmen einer Online-Befragung eine Selbsteinschätzung der Schwere von Zwangssymptomen vor. Das Inventar besteht aus 18 Items, von denen jeweils drei Items zu einer Subskala gehören. Die sechs Subskalen „Waschen“, „Kontrollieren“, „Ordnen“, „Zwangsvorstellungen“, „Horten“ und „Gedankliches Neutralisieren“ bilden jeweils einen typischen Symptombereich der Zwangsstörung ab. Bei jedem Item, das in Form einer Selbstaussage formuliert ist, schätzen die Teilnehmenden auf einer Skala von null („gar nicht“) bis vier („sehr stark“) ein, wie stark sie das Verhalten oder Erleben in der vergangenen Woche beeinträchtigt hat. Der Testwert jeder Subskala reicht von 0 bis 12 und der Gesamtwert von 0 bis 72. Gesamtwerte bis 15 können als leichte („mild“), Werte von 16 bis 27 als mäßige („moderate“) und Werte über 28 als schwere („severe“) Zwangssymptome interpretiert werden (Abramovitch et al. 2020). Für das Inventar konnte eine hohe interne Konsistenz (Cronbach's $\alpha = 0.85$) und eine gute Test-Retest-Reliabilität ($r = 0.74 - 0.91$) gezeigt werden (Foa et al. 2002, Gönner et

al. 2007). Die konvergente und divergente Validität wurde als ausgezeichnet beschrieben (Gönner et al. 2007).

Global Assessment of Functioning (GAF)

Das allgemeine Funktionsniveau der psychischen, sozialen und beruflichen Leistungsfähigkeit wurde mit Hilfe der Skala des Global Assessment of Functioning (GAF) zu den Zeitpunkten T0, T1 und T2 erhoben. Das GAF ist ein etabliertes und verbreitetes globales Fremdbeurteilungsverfahren, das das Funktionsniveau einer Person auf einer Skala von 1 bis 100 erfasst. Beeinträchtigungen der Funktionsfähigkeit durch körperliche oder durch das soziale Umfeld bedingte Einschränkungen werden dabei nicht berücksichtigt. Jede Dekade der Skala wird durch einen kurzen Fließtext näher beschrieben, wobei jeder Zwischenwert kodiert werden kann. Die Skala reicht von ständiger Fremd- oder Selbstgefährdung, Suizidversuch und/ oder anhaltender Unfähigkeit zur Körperpflege (1 bis 10) bis hin zu Leistungsfähigkeit in allen Bereichen, sozialer Integration und Interesse an einer Vielzahl von Aktivitäten (90 bis 100). Das GAF wurde 1987 von der American Psychological Association entwickelt und ist als Erfassung des Funktionsniveaus (Achse fünf) Teil der multiaxialen Diagnostik und Klassifikation psychischer Störungen des international anerkannten Diagnostic and Statistical Manual of Mental Disorders, vierte Auflage (DSM-IV, Hilsenroth et al. 2000, Saß et al. 2003). Insgesamt kann die Reliabilität des GAF als ausreichend, bei geschulten Beurteilenden und guten Erhebungsbedingungen als hoch eingestuft werden (Laireiter et al. 2001, Lange und Heuft 2002). Die Konstrukt- und Kriteriumsvalidität wird als hinreichend beschrieben (Jones et al. 1995), während die geringe diskriminierende Validität zwischen der GAF-Skala und der psychopathologischen Symptomatik häufig kritisiert wird (Laireiter et al. 2001, S. 210-228, Lange und Heuft 2002).

Patient Health Questionnaire – 9 (PHQ-9)

Zur Quantifizierung einer möglichen komorbiden Depression wurde der PHQ-9 als Selbstbeurteilungsinstrument im Rahmen des Online-Fragebogens zu den Zeitpunkten T0, T1 und T2 erhoben.

Der PHQ-9 ist ein Teil des PHQ, einem Multiple-Choice-Selbsteinschätzungsinstrument, das für das Screening und die Diagnose verschiedener psychischer Störungen wie Depressionen, Panik- und anderen Angststörungen, Alkoholabhängigkeit, Essstörungen und somatoforme Störungen entwickelt wurde.

Der PHQ-9 ist der Teil, der sich auf die Diagnose von Major Depression und anderen depressiven Symptomen bezieht und besteht aus neun Items, die sich jeweils auf eines der neun DSM-IV-Kriterien für Depression beziehen.

Jedes der neun Items wird auf einer Skala von 0 („überhaupt nicht“) bis 3 („beinahe jeden Tag“) bewertet. Ein zusätzliches zehntes Item fragt, wie stark die zuvor abgefragten Symptome die befragte Person belasten. Dieses geht nicht in den Gesamtwert ein und wurde hier nicht erhoben. Der Gesamtwert des PHQ-9, indikativ für den Schweregrad der Symptomatik, kann zwischen 0 und 27 liegen. Werte von 0 bis 4 bedeuten minimale depressive Symptome („minimal“), 5 bis 9 eine leichte Depression („mild“), 10 bis 14: „moderate“, 15 bis 19: „moderately severe“) und ≥ 20 eine schwere Depression („severe“).

Für den PHQ-9 konnte eine hohe interne Konsistenz (Cronbach's $\alpha = 0.86 - 0.89$) und eine exzellente Test-Retest-Reliabilität gezeigt werden. Es erwies sich als ein valides Instrument zur Bewertung des Schweregrads einer Depression (Kroenke et al. 2001).

Fragebogen zur Messung der Patientenzufriedenheit ZUF-8

Die Zufriedenheit der Teilnehmenden mit dem B4DT wurde zu den Zeitpunkten T1 und T2 mit Hilfe des ZUF-8 im Rahmen des Online-Fragebogens erhoben. Der ZUF-8 wurde in Anlehnung an den amerikanischen Client Satisfaction Questionnaire (CSQ-8, Attkisson und Zwick (1982)) als Messinstrument zur globalen, eindimensionalen Erfassung der Patientenzufriedenheit nach einer Behandlung entwickelt. Der CSQ-8

wurde für den ambulanten Mental-Health-Bereich konzipiert. Der ZUF-8 wurde so adaptiert, dass er auch für die Messung der Patientenzufriedenheit nach einem stationären Aufenthalt in einem Krankenhaus, einer Rehabilitationsklinik, einer stationären Psychotherapie oder einer psychosomatischen Rehabilitation eingesetzt werden kann. Der ZUF-8 besteht aus acht Items, die auf einer Skala von 1 („ziemlich unzufrieden“) bis 4 („sehr zufrieden“) beantwortet werden können. Es kann ein Gesamtzufriedenheitswert errechnet werden, der Werte zwischen 8 und 32 annehmen kann.

Der ZUF-8 ist nicht normiert und weist eine hohe interne Konsistenz (Cronbach's $\alpha = 0.87 - 0.93$) auf (Schmidt et al. 1989). Untersuchungen zur Validität ergaben ein hohes Maß an faktorieller, konvergenter und prognostischer Validität (Schmidt und Nübling 2002, S. 392-396).

4.5. Kriterien für klinisch signifikante Verbesserung und Klassifikation von Therapieansprechen und Remission

Das Ansprechen auf die Behandlung wurde auf der Grundlage der Konsenskriterien internationaler Experten berechnet (Mataix-Cols et al. 2016). Ein klinisch relevantes Ansprechen auf die Behandlung („Response“) wurde definiert als eine Verringerung des Y-BOCS-Scores um $\geq 35\%$. Eine Person gilt als in Remission, wenn der Y-BOCS-Score nach der Behandlung ≤ 12 Punkten betrug (Mataix-Cols et al. 2016). Zusätzlich wurde berechnet, bei wie vielen Personen der Y-BOCS-Score nach der Behandlung sowohl um $\geq 35\%$ reduziert wurde als auch einen Wert von ≤ 12 Punkten annahm.

4.6. Hypothesen und Statistische Analyse

Primärer Outcome: Reduktion der fremdbeurteilten Zwangssymptome

H1: Die Teilnahme am B4DT führt bei Patient:innen mit Zwangsstörung zu einer signifikanten Reduktion der Zwangssymptomatik von T0 zu T1 und von T0 zu T2, gemessen anhand der Y-BOCS.

Sekundärer Outcome: Selbstberichtete Zwangssymptome

H2: Die Teilnahme am B4DT führt bei Patient:innen mit Zwangsstörung zu einer signifikanten Reduktion der selbstberichteten Zwangssymptome von T0 zu T1 und T0 zu T2, gemessen anhand der OCI-R.

Sekundärer Outcome: Depressive Komorbidität

H3: Die Teilnahme am B4DT führt bei Patient:innen mit Zwangsstörung zu einer signifikanten Reduktion der depressiven Symptome von T0 zu T1 und T0 zu T2, gemessen anhand des PHQ-9.

Sekundärer Outcome: Funktionsfähigkeit

H4: Die Teilnahme am B4DT führt zu einer signifikanten Steigerung der allgemeinen Funktionsfähigkeit bei Patient:innen mit Zwangsstörung von T0 zu T1 sowie von T0 zu T2, gemessen anhand des GAF.

Für alle Analysen wurde die Software IBM SPSS 29.0 (IBM Corp., 2022) verwendet.

Für die Testung der Hypothesen **H1**, **H2**, **H3** und **H4** wurde mittels dreistufiger Varianzanalyse mit Messwiederholung (ANOVA mit Messwiederholung) geprüft, ob statistisch signifikante Unterschiede zwischen den drei abhängigen Stichproben T0, T1 und T2 bestanden.

Verglichen wurden dabei die Mittelwerte der Y-BOCS, des OCI-R, des PHQ-9 und des GAF vor der Intervention (T0), eine Woche nach der Intervention (T1) und drei Monate nach der Intervention (T2). Die statistischen Auswertungen wurden auf Basis einer Complete-Case-Analyse durchgeführt. Für die ANOVA wurden ausschließlich Datensätze von Teilnehmer:innen berücksichtigt, für die zu allen relevanten Messzeitpunkten vollständige Werte vorlagen. Fälle mit fehlenden Werten (Missing Data) in den primären Outcome-Variablen wurden von der jeweiligen Analyse ausgeschlossen. Eine Imputation fehlender Werte erfolgte nicht.

Wenn der Mauchly-Test auf Sphärizität Signifikanz annahm, wurde die Greenhouse-Geisser-Korrektur angewandt. Um bei der kleinen Stichprobenzahl trotz

durchgeführter Complete-Case-Analyse alle vorhandenen Datensätze auszuwerten, wurden zusätzlich für die Testung der oben genannten Hypothesen **H1**, **H2**, **H3** und **H4** gepaarte *t*-Tests durchgeführt. So konnten auch Personen in die Analyse einbezogen werden, für die keine Follow-Up-Daten zu T2 vorlagen. Ergebnisse mit $p \leq .05$ wurden als statistisch signifikant angesehen.

Die Effektstärken wurden mit Cohen's *d* und ηp^2 berechnet. Für die Interpretation der Effektstärken wurden die Faustregeln von Cohen (1988) angewandt, nach denen Cohen's *d* Werte unter 0.2 als sehr gering bis vernachlässigbar, Werte von 0.2 bis 0.5 als ein geringer Effekt, Werte von 0.5 bis unter 0.8 als ein mittlerer und Werte ab 0.8 als ein großer Effekt interpretiert werden. Für die Interpretation der Effektstärke ηp^2 gelten nach Cohen (1988) Werte ab .01 als kleiner Effekt, Werte ab .06 als mittlerer Effekt und Werte ab .14 als großer Effekt.

Sekundärer Outcome: Klinische Relevanz des Ansprechens

H5: Es werden in Anbetracht der Daten aus norwegischen Publikationen vergleichbare Ergebnisse für das klinische Ansprechen auf die Behandlung der Patient:innen erzielt.

H5a: Der Anteil der Patient:innen, die eine Woche nach dem B4DT die Kriterien für ein klinisches Ansprechen ($\geq 35\%$ Reduktion auf der Y-BOCS) und eine Remission ($Y-BOCS \leq 12$) zeigen, ist deskriptiv vergleichbar mit den Ergebnissen der Entwickler:innen des Behandlungsformats (Hansen et al. 2018, Kvale et al. 2018, Launes et al. 2019b).

H5b: Der Anteil der Patient:innen, die drei Monate nach der Intervention (T2) die Kriterien für ein klinisches Ansprechen ($\geq 35\%$ Reduktion auf der Y-BOCS) auf B4DT und eine Remission ($Y-BOCS \leq 12$) zeigen, ist deskriptiv vergleichbar zu den existierenden Daten aus den Studien der Entwickler:innen des Behandlungsformats (Hansen et al. 2018, Kvale et al. 2018, Launes et al. 2019b).

Um Hypothese **H5a** und **H5b** zu testen und das klinische Ansprechen mit der bestehenden Literatur zu vergleichen, wurden die Kriterien für klinisch signifikante Verbesserung und Klassifikation von Therapieansprechen und Remission angewandt (vgl. 5.5.). Es erfolgte kein inferenzstatistischer Vergleich mit früheren Studien; der Vergleich der Ansprech- und Remissionsraten erfolgte deskriptiv. Für die deskriptiven Referenzwerte siehe *Empirische Grundlage für den deskriptiven Vergleich* (S.22/23).

Sekundärer Outcome: Zufriedenheit und Machbarkeit

H6: Die Mehrheit der Patient:innen zeigt zum Zeitpunkt T1 und T2 eine hohe Zufriedenheit mit der Behandlung, gemessen anhand des ZUF-8, und empfindet die Behandlung als hilfreich.

Um Hypothese **H6** zu testen wurde die Zufriedenheit mittels ZUF-8 zum Zeitpunkt T1 und T2 erhoben und deskriptiv ausgewertet. Ein inferenzstatistischer Vergleich wird nicht vorgenommen, da für den ZUF-8 keine etablierten Norm- oder Vergleichswerte vorliegen.

H7: Die Abbruchrate der vorliegenden Studie fällt geringer aus als 11-19%, was der berichteten Abbruchrate in der bisherigen Standardtherapie KVT mit ERM entspricht (Olatunji et al. 2013, Voderholzer et al. 2022).

Um Hypothese **H7** zu testen, wurde nach Abschluss der Datenerhebung die Drop-out-Rate berechnet und deskriptiv ausgewertet.

5. Ergebnisse

5.1. Stichprobe

Soziodemographische Informationen über die 33 Personen, die an der Behandlung teilgenommen haben, sind in Tabelle 3 dargestellt.

Zu Beginn der Studie waren 58.1 % der Teilnehmenden schon mindestens einmal in (teil)-stationärer Behandlung. Definiert man ein „Early onset“ als das Erstauftreten der Symptome im Alter von 15 Jahren oder früher (Janowitz et al. 2019), so bestand bei 39.4 % der Stichprobe ein Early onset. Die Dauer der Zwangssymptome vor Behandlungsbeginn reichte von keinem ganzen Jahr bis zu 59 Jahren und betrug im Mittel 17.94 Jahre. Der Anteil der Teilnehmenden, der aktuell, früher oder rezidivierend an Depressionen litt oder leidet lag bei 78.9 % (26 Teilnehmer:innen).

5.2. Adhärenz und Drop-out

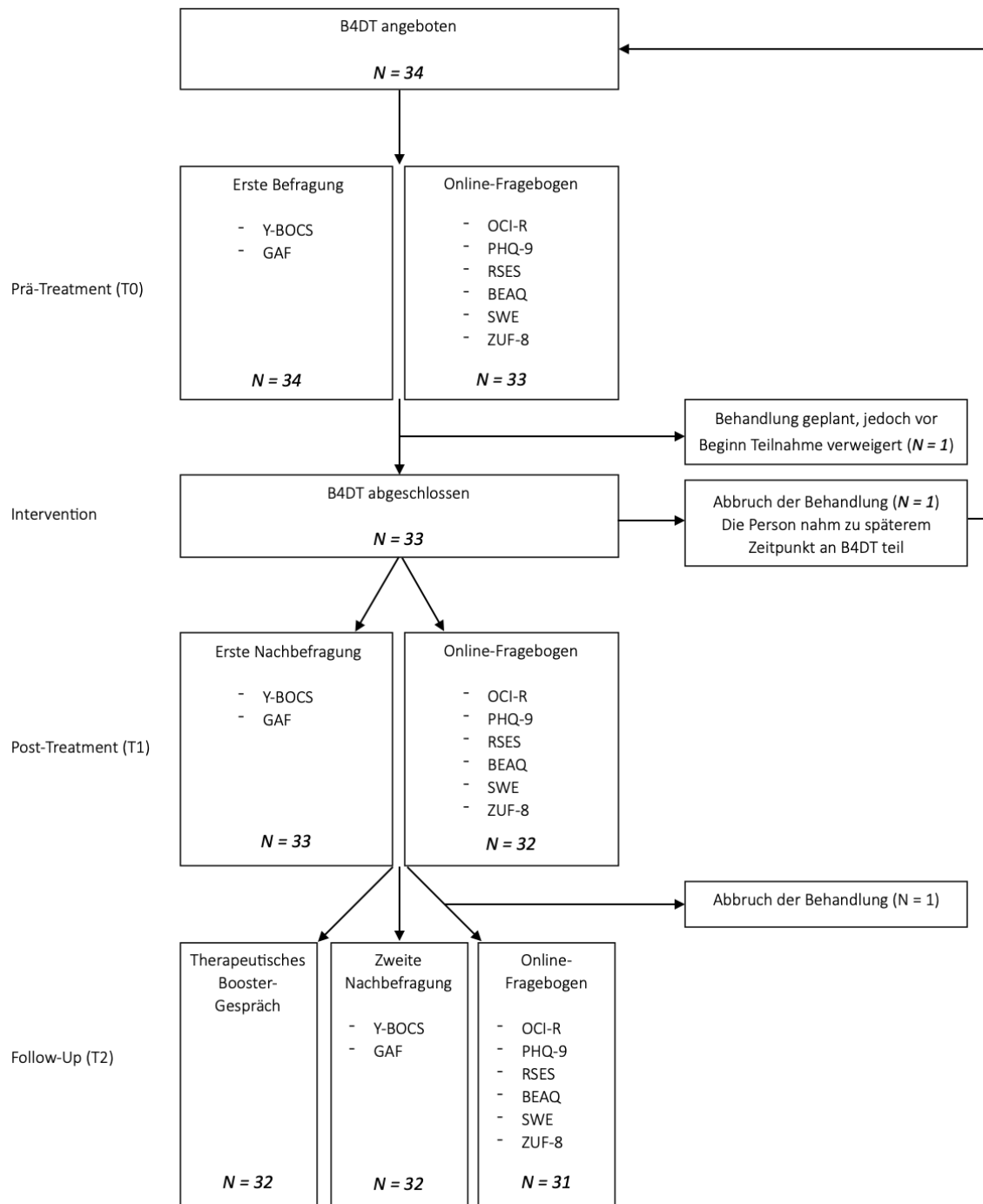
Von 34 rekrutierten und eingeplanten Personen nahmen 33 an der Behandlung teil und schlossen 32 die Studie inklusive aller Nachbefragungen vollständig ab. Eine Person war für die Intervention eingeplant, brach die Teilnahme jedoch vor dem ersten Behandlungstag ab. Von den 33 Personen, die an der Studie teilgenommen haben, brach eine Person die Behandlung vorzeitig ab, konnte jedoch zu einem späteren Zeitpunkt teilnehmen und wird in die Analysen eingeschlossen. Eine Person nahm zwar an der Intervention und den Datenerhebungen T0 und T1, jedoch nicht an T2 teil. Somit bleiben 32 Personen, die die Behandlung und alle Nachbefragungen absolviert haben und ein Drop-Out zwischen T1 und T2. Eine Person konnte aufgrund mangelnder Deutschkenntnisse nur an den in persona erhobenen, nicht jedoch an den Online-Befragungen zu den Zeitpunkten T0, T1 und T2 teilnehmen. Die entsprechende Behandlung hatte in dieser Gruppe auf Englisch stattgefunden. Die Instrumente Y-BOCS, M.I.N.I. und GAF sowie die soziodemographischen Daten wurden somit bei der Person erhoben, die Daten der Erhebungen von OCI-R, PHQ-9 und ZUF-8 fehlen zu allen Zeitpunkten. Ein Flussdiagramm der Studie ist in Abbildung 2 zu sehen. Es ergibt sich eine Drop-out-Rate von 6.1%. Hypothese **H7** konnte bestätigt werden.

Tabelle 3*Soziodemografische Daten, Komorbiditäten und Medikation der Stichprobe*

Merkmale der Stichprobe (N = 33)	N (%)	M (SD)
		Minimum / Maximum
Alter (in Jahren)		35.45 (12.94) 19/67
Geschlecht		
Weiblich	20 (60.6)	
Männlich	13 (39.4)	
Höchster Bildungsabschluss	3 (9.1)	
Mittlerer Schulabschluss	11 (33.3)	
Abitur	7 (21.2)	
Bachelor/ Äquivalent	12 (36.4)	
Master/ Äquivalent		
Beschäftigungsstatus		
Vollzeit	12 (36.4)	
Teilzeit	5 (15.2)	
Ausbildung/ Studium	7 (21.2)	
Pension/ Rente	1 (3.0)	
Arbeitsuchend	3 (9.1)	
Sonstiges	5 (15.2)	
Beziehungsstatus		
Verheiratet	9 (27.3)	
Feste Partnerschaft	10 (30.3)	
Keine feste Partnerschaft	14 (42.4)	
Schwere der Zwangsstörung T0		
Moderate Zwangsstörung (Y-BOCS-Score 18-23)	11 (33.3)	
Schwere Zwangsstörung (Y-BOCS-Score ≥ 24)	22 (66.7)	
Alter bei erstmaligem Auftreten von Zwangssymptomatik (in Jahren)		17.88 (7.34) 5/33
Dauer der Zwangsstörung (in Jahren)		27.94 (14.73) 0/59
Anzahl (teil-)stationärer Aufnahmen		1.65 (3.46) 0/19
Komorbiditäten		
Keine aktuelle Komorbidität	20 (60.6)	
Mindestens eine aktuelle Komorbidität	13 (39.4)	
Depressionen	6 (18.2)	
Panikstörung	4 (12.1)	
Soziale Phobie	3 (9.1)	
Essstörungen	2 (6.1)	
Psychotrope Medikation		
Keine	19 (57.6)	
Antidepressiva	13 (39.4)	
Antipsychotika	0 (0.0)	
Kombiniert Antidepressiva + Antipsychotika	1 (3.0)	
Benzodiazepine	0 (0.0)	

Anmerkungen. Y-BOCS = Yale-Brown Obsessive-Compulsive Scale, M = Mittelwert, SD = Standardabweichung.

Abbildung 2
Flussdiagramm der durchgeführten Pilotstudie



Anmerkungen. BEAQ = Brief Experimental Avoidance Questionnaire, GAF = Global Assessment of Functioning, OCI-R = Obsessive-Compulsive Inventory – Revised, PHQ-9 = Patient Health Questionnaire, RSES = Rosenberg Self Esteem Scale, SWE = Skala zur Allgemeinen Selbstwirksamkeitserwartung, Y-BOCS = Yale-Brown Obsessive-Compulsive Scale, ZUF-8 = Fragebogen zur Patientenzufriedenheit

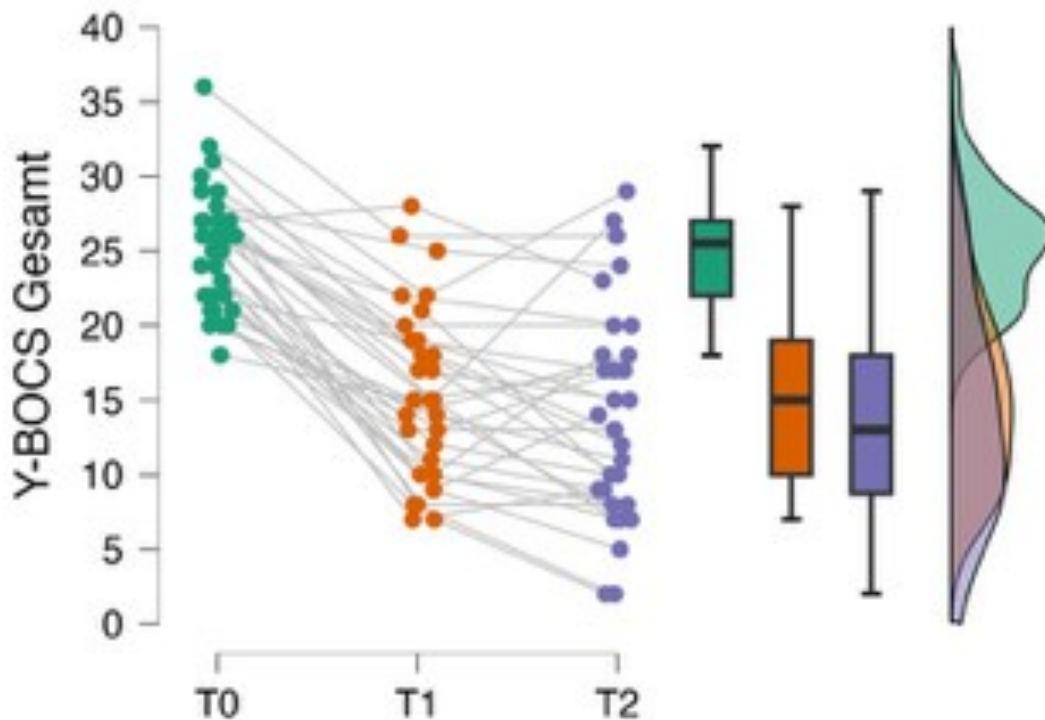
5.3. Zwangssymptomatik

Die Interrater-Reliabilität des Y-BOCS-Gesamtscores erwies sich als hoch (ICC = .871; 95%-KI .719 bis .961). Die interne Konsistenz lag ebenfalls im sehr guten Bereich (Cronbach's α = 0.894). Tabelle 4 zeigt die Ergebnisse der primären und sekundären Outcomes. Es konnte gezeigt werden, dass B4DT die Zwangssymptomatik signifikant mit großer Effektstärke reduziert. Eine ANOVA mit wiederholten Messungen wurde verwendet, um die Zwangssymptomatik, gemessen anhand der Y-BOCS, zwischen den Messzeitpunkten T0, T1 und T2 zu vergleichen. Da eine Verletzung der Voraussetzung der Sphärizität vorlag, wie der Mauchly-Test zeigte ($W = 0.81$, $p = .042$), wurde eine Greenhouse-Geisser Korrektur der Freiheitsgrade vorgenommen (GGe = .84). Wie erwartet, wurde der Unterschied der Gesamtscores der Y-BOCS zwischen den Zeitpunkten als signifikant eingestuft ($F(1.68, 52.09) = 77.80$, $p < .001$, $\eta^2 = .72$). Die gepaarten t -Tests zeigten, dass sowohl die Zwangssymptomatik von T0 zu T1 ($p < .001$), als auch von T0 zu T2 ($p < .001$) signifikant abnahm. Die dabei erzielten Effektstärken ließen sich nach Cohen (1988) jeweils als groß einschätzen (T0 zu T1 $d = 2.04$; T0 zu T2 $d = 1.69$). In Abbildung 3 sind die Y-BOCS-Gesamtscores über die Zeit (T0, T1 und T2) abgebildet.

Hypothese **H1** konnte bestätigt werden.

Abbildung 3

Individuelle Veränderungen, Boxplots und Dichteverteilungen der Y-BOCS-Gesamtwerte zu den drei Messzeitpunkten



Anmerkungen. **Links:** Punkte als Y-BOCS-Scores verbunden mit Linien repräsentieren die individuellen Verläufe **Mitte:** Mittelwerte $\pm$ Standardabweichung als Balken. **Rechts:** Verteilung der Scores als Half-Violin-Plot. Alle Darstellungen zeigen die deutliche Abnahme der Symptomschwere nach Behandlung und die Stabilität bis zum Follow-Up

Betrachtet man isoliert die Subskalen der Y-BOCS für Zwangsgedanken und Zwangshandlungen, so konnte auch hier für beide Skalen eine signifikante Reduktion mit großer Effektstärke durch B4DT gezeigt werden.

In der durchgeführten ANOVA mit wiederholten Messungen zeigte sich in beiden Subskalen ein signifikanter Unterschied über die Zeit, entsprechend einer großen Effektstärke (siehe Tabelle 4). Für die Subskala der Zwangsgedanken wurde die Annahme der Sphärizität verletzt, wie der Mauchly-Test zeigte ($W = 0.69$, $p = .004$). Unter Verwendung einer Greenhouse-Geisser-Korrektur ($GGe = .76$) war der Gruppenunterschied zwischen den Zeitpunkten statistisch signifikant ($p < .001$). Die t -Tests der gepaarten Messwerte T0 und T1 sowie T0 und T2 zeigten sowohl für die

Zwangsgedanken als auch die Zwangshandlungen jeweils eine signifikante Reduktion der Zwangssymptome mit großer Effektstärke (siehe Tabelle 4).

Auch die Testscores des Selbsteinschätzungsinstruments OCI-R wurden mittels ANOVA mit Messwiederholung und gepaarten t -Tests analysiert. Es zeigte sich auch hier über die Zeit eine signifikante Veränderung der Schwere der selbstberichteten Zwangssymptome mit großer Effektstärke ($p < .001$, $\eta p^2 = .58$, siehe Tabelle 4). Da eine Verletzung der Voraussetzung der Sphärizität vorlag, wie der Mauchly-Test zeigte ($W = 0.77$, $p = .021$), wurde eine Greenhouse-Geisser Korrektur der Freiheitsgrade vorgenommen ($GGe = .81$). t -Tests mit gepaarten Stichproben zeigten eine signifikante Reduktion der Gesamtscores des OCI-R, sowohl von T0 zu T1 ($p < .001$) als auch von T0 zu T2 ($p < .001$). Beide Effekte zeigten große Effektstärken (T0 zu T1 $d = 1.32$; T0 zu T2 $d = 1.19$, siehe Tabelle 4).

Hypothese **H2** konnte bestätigt werden.

Tabelle 4*Ergebnisse: Primäre und sekundäre Outcomes*

	T0 N = 33	T1 N = 33 M (SD)	T2 N = 32	ANOVA	T0 vs. T1	T0 vs. T2
Yale-Brown Obsessive Compulsive Scale (Y-BOCS)						
<i>Gesamtscore</i>	25.18 (3.96)	15.24 (5.63)	13.84 (7.08)	$F(1.68, 52.09) = 77.80$, $p < .001$, $\eta^2 = .72$	$t(32) = 11.72$, $p < .001$, $d = 2.04$ KI _{95%} [1.43, 2.64]	$t(31) = 9.54$, $p < .001$, $d = 1.69$ KI _{95%} [1.14, 2.22]
<i>Score für Zwangsgedanken</i>	12.48 (2.31)	8.06 (2.66)	7.03 (3.69)	$F(1.52, 47.21) = 59.89$, $p < .001$, $\eta^2 = .66$	$t(32) = 10.90$, $p < .001$, $d = 1.90$ KI _{95%} [1.32, 2.47]	$t(31) = 8.32$, $p < .001$, $d = 1.47$ KI _{95%} [0.96, 1.97]
<i>Score für Zwangshandlungen</i>	12.70 (1.94)	7.18 (3.39)	6.81 (3.80)	$F(2,62) = 63.72$, $p < .001$, $\eta^2 = .67$	$t(32) = 10.21$, $p < .001$, $d = 1.78$ KI _{95%} [1.22, 2.32]	$t(31) = 8.97$, $p < .001$, $d = 1.59$ KI _{95%} [1.06, 2.10]
Obsessive Compulsive Inventory-Revised (OCI-R)	24.67 (9.19)	11.79 (5.64)	13.25 (6.92)	$F(1.63, 50.50) = 42.68$, $p < .001$, $\eta^2 = .58$	$t(32) = 7.58$, $p < .001$, $d = 1.32$ KI _{95%} [0.85, 1.78]	$t(31) = 6.75$, $p < .001$, $d = 1.19$ KI _{95%} [0.73, 1.64]
Global Assessment of Functioning (GAF)	64.64 (13.51)	75.82 (12.87)	78.87 (13.19)	$F(2,62) = 44.25$, $p < .001$, $\eta^2 = .59$	$t(32) = 8.08$, $p < .001$, $d = -1.41$ KI _{95%} [- 1.89, - 0.91]	$t(31) = 7.98$, $p < .001$, $d = -1.41$ KI _{95%} [- 1.90, - 0.91]
Patient Health Questionnaire 9 (PHQ-9)	9.91 (5.36)	6.39 (5.30)	6.22 (4.19)	$F(2,62) = 16.86$, $p < .001$, $\eta^2 = .35$	$t(32) = 4.78$, $p < .001$, $d = 0.83$ KI _{95%} [0.43, 1.22]	$t(31) = 5.40$, $p < .001$, $d = 0.95$ KI _{95%} [0.53, 1.37]

Anmerkungen. T0 vor der Intervention, T1 eine Woche nach der Intervention, T2 beim Follow-Up 3 Monate nach der Intervention. ANOVA = Varianzanalyse mit Messwiederholung, KI = Konfidenzintervall.

5.4. Klinische Relevanz

Nach internationalen Konsenskriterien (Mataix-Cols et al. 2016) zeigten eine Woche nach der Behandlung (T1) 6 Teilnehmer:innen (18.2 %) ein klinisch relevantes Ansprechen auf die Behandlung („Response“) und 11 Personen (33.3 %) waren in Remission, während 16 Personen (48.5 %) nicht klinisch relevant auf die Intervention ansprachen (siehe Abbildung 4). Alle 11 Personen, die durch einen Y-BOCS-Score von ≤ 12 Punkten als remittiert eingestuft wurden, erfüllten gleichzeitig das Kriterium der Reduktion des Y-BOCS-Scores um ≥ 35 %. Im Vergleich zu den Ergebnissen, die aus Norwegen berichtet wurden, zeigt sich somit ein geringeres klinisches Ansprechen auf B4DT in unserer Pilotstudie zum Zeitpunkt T1.

Hypothese **H5a** kann nicht bestätigt werden.

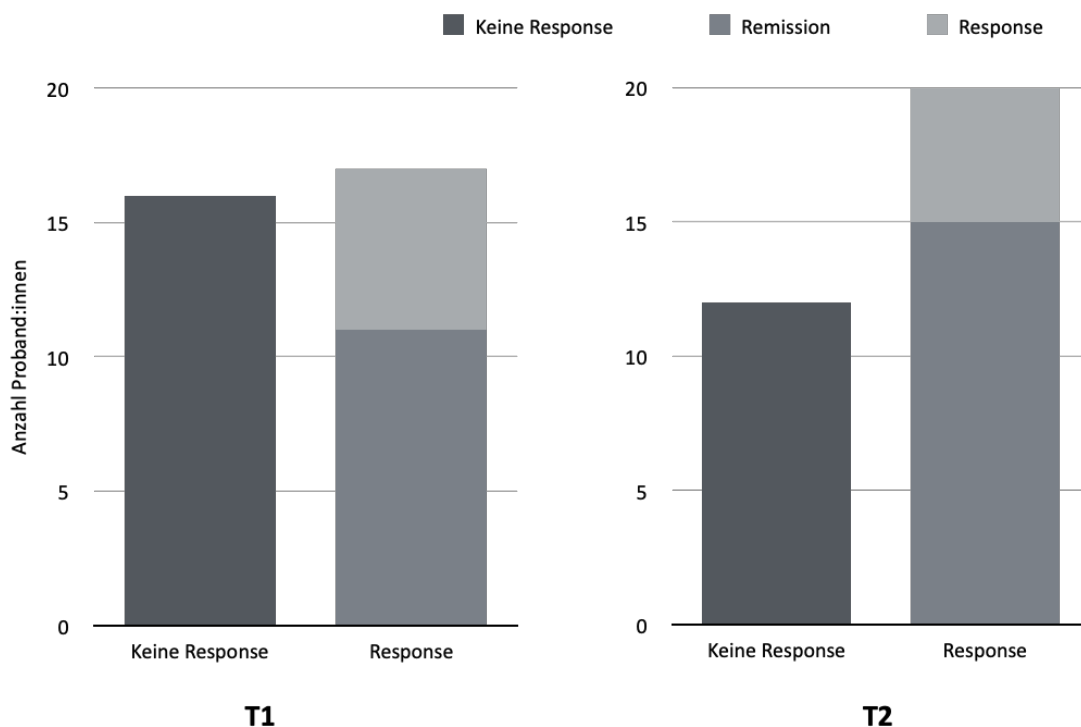
Bei dem Drei-Monats-Follow-Up (T2), zeigten 5 Personen (15.6 %) eine Response, 15 Personen (46.9 %) waren in Remission, 12 Personen (37.5 %) sprachen nicht klinisch relevant auf die Behandlung an (siehe Abbildung 4). Alle 15 Personen, die durch einen Y-BOCS-Score von ≤ 12 Punkten als remittiert eingestuft wurden, erfüllten gleichzeitig das Kriterium der Reduktion des Y-BOCS-Scores um ≥ 35 %. Auch zum Zeitpunkt T2 zeigte sich somit im Vergleich zu den norwegischen Ergebnissen ein geringeres klinisches Ansprechen auf B4DT.

Hypothese **H5b** kann nicht bestätigt werden.

Von den elf Patient:innen, die sich zum Zeitpunkt T1 in Remission befanden, blieben neun Personen auch nach drei Monaten in Remission. Von den sechs Patient:innen, die eine Woche nach der Intervention als Responder eingestuft wurden, gingen zwei Personen in Remission zum Zeitpunkt T2, zwei Personen verblieben Responder und zwei Personen stiegen in die Kategorie keiner Veränderung ab. Von den 15 Teilnehmer:innen, die keine Veränderung zum Zeitpunkt T1 zeigten, waren zum Zeitpunkt T2 vier Personen in Remission, zwei Personen zeigten ein Ansprechen und neun Personen blieben unverändert.

Abbildung 4

Klinisch signifikante Veränderungen – Therapieansprechen (Response) und Remission



Anmerkungen. Response = Verringerung des Y-BOCS-Scores um $\geq 35\%$; Remission = Y-BOCS-Score ≤ 12 Punkte

5.5. Depressivität

Die depressive Symptomatik der Proband:innen, gemessen über den PHQ-9, veränderte sich von vor bis nach der Behandlung signifikant ($p < .001$), entsprechend einer großen Effektstärke ($\eta p^2 = .35$, siehe Tabelle 4).

Die gepaarten t -Tests zeigten, dass sowohl die depressive Symptomatik von T0 zu T1 ($p < .001$), als auch von T0 zu T2 ($p < .001$) signifikant abnahm. Die dabei erzielten Effektstärken ließen sich nach Cohen (1988) jeweils als groß einschätzen (T0 zu T1 $d = 0.83$, T0 zu T2 $d = 0.95$). Während 16 Teilnehmer:innen (48.48 %) zum Zeitpunkt T0 minimale oder leichte depressive Symptome zeigten und 17 Teilnehmer:innen (51.51 %) mittelschwere oder schwere depressive Symptome, so zeigten 24 Teilnehmer:innen (75 %) zum Zeitpunkt T2 eine minimale oder leichte Depression und acht Teilnehmer:innen (25%) eine mittelschwere Depression. Schwere depressive Symptomatik zeigte keine:r der Proband:innen zum Zeitpunkt T2.

Hypothese **H3** konnte bestätigt werden.

5.6. Allgemeines Funktionsniveau

Das allgemeine Funktionsniveau der Proband:innen besserte sich durch B4DT. Die Unterschiede im Funktionsniveau der Teilnehmer:innen zu den Zeitpunkten T0, T1 und T2 nahmen Signifikanz an ($p < .001$), entsprechend einer großen Effektstärke ($\eta p^2 = .59$, siehe Tabelle 4).

Die gepaarten t -Tests zeigten, dass sowohl das Funktionsniveau von T0 zu T1 ($p < .001$), als auch von T0 zu T2 ($p < .001$) signifikant zunahm. Die dabei erzielten Effektstärken lassen sich nach Cohen (1988) jeweils als groß einschätzen (siehe Tabelle 4). Der Mittelwert lag vor der Intervention bei $M = 64.64$ und zum Zeitpunkt T2 bei $M = 78.87$, was „höchstens leichten Beeinträchtigung der sozialen beruflichen und schulischen Leistungsfähigkeit“ bedeutet (Saß et al. 2003).

Hypothese **H4** konnte bestätigt werden.

5.7. Zufriedenheit

Die Teilnehmenden bewerteten die Zufriedenheit mit dem intensiven Behandlungsformat im ZUF-8 im Mittel mit $M = 29.30$ ($SD = 3.06$) zum Zeitpunkt T1 und mit $M = 28.41$ ($SD = 4.41$) zum Zeitpunkt T2. Höchstwert der Skala des ZUF-8 ist 32, sodass erzielte Mittelwerte von $M = 29.30$ beziehungsweise $M = 28.41$ eine hohe Zufriedenheit indizieren, auch wenn keine Normierung für die Interpretation des ZUF-8 vorliegt.

Der ZUF-8 wurde als deutschsprachige Version des CSQ-8 entworfen, der ebenfalls eine Skala von 8 bis 32 verwendet (Schmidt et al. 1989). In den norwegischen Studien zum B4DT wurde bei CSQ-8-Mittelwerten zwischen 28 und 31 von hoher Patient:innenzufriedenheit berichtet (Hansen et al. 2018, Iversen et al. 2022, Launes et al. 2019a, Launes et al. 2020). Dies deutet darauf hin, dass die Behandlung auch in der vorliegenden Stichprobe als sehr zufriedenstellend wahrgenommen wurde.

Die allgemeine Zufriedenheit mit der Behandlung drei Monate nach der Intervention wurde von 22 Teilnehmer:innen (68.8 %) als „sehr zufrieden“ beschrieben, von sieben Teilnehmer:innen (21.9 %) als „weitgehend zufrieden“, zwei Personen (6.3 %) waren „leicht unzufrieden“ und eine Person (3.1 %) „ziemlich unzufrieden“. Alle 32 Teilnehmer:innen beantworteten die Frage, ob sie das B4DT an einen Freund empfehlen würden, wenn er/sie ähnliche Hilfe benötige, mit „Eindeutig ja“ (84.4 %) oder mit „Ich glaube ja“ (15.6 %). Die Qualität der Behandlung wurde von 16 Teilnehmer:innen (50 %) als „ausgezeichnet“, von 14 als „gut“ (43.8 %) und von zwei Personen (6.3 %) als „weniger gut“ bezeichnet.

Hypothese **H6** konnte bestätigt werden.

Die Teilnehmer:innen hatten drei Monate nach der Intervention die Möglichkeit anzumerken, was an dem B4DT besonders gefallen hat, und was sie für verbesserungswürdig hielten. Positiv empfanden Teilnehmer:innen den Austausch mit den anderen Teilnehmer:innen in kleiner Gruppengröße und das gemeinsame Mittagessen. Es sei die „motivierende Stimmung“ entstanden, „gemeinsam an einem Strang zu ziehen“ und eine „wohlwollende Atmosphäre“. Die Betreuung durch die Therapeut:innen, wurde als „intensiv“, „sehr gut“, „empathisch“, „kompetent“ und „motivierend“ beschrieben. Inhaltlich wurde positiv hervorgehoben, dass der Fokus während der Behandlungswoche nicht auf der Ursachenforschung und der individuellen Vergangenheit mit der Erkrankung lag, sondern auf dem „Arbeiten am Symptom“, dem „Loslassen“ und der persönlichen Motivation. Die theoretische Grundlage durch den ersten Tag habe geholfen, Zusammenhänge und Prozesse zu verstehen. Der schnelle Übergang zu den Expositionsübungen und die Verbindung dieser Zusammenhänge gaben mehrere Teilnehmer:innen als positiv an.

Als verbesserungswürdig wurden neben organisatorischen Themen, wie den nötigen Untersuchungen vor tagesklinischer Aufnahme oder dem Wunsch nach mehr Informationsmaterial vor der Intervention vor allem die Zeit und Betreuung nach der Intervention bewertet. Teilnehmer:innen fühlten sich verunsichert, oder verspürten „Ungewissheit“ in Bezug auf das richtige Anwenden der Technik nach der Intervention. Daher merkten einige Teilnehmer:innen an, dass sie sich mehrere therapeutische Treffen kürzer nach der Behandlungswoche gewünscht hätten, eine Begleitung über

einen längeren Zeitraum, oder gar die Aufteilung der Expositionsübungen auf mehrere Wochen. Auf die Frage, ob die Teilnehmenden eine Behandlung über 15 Wochen wöchentlich oder über vier Tage intensiv bevorzugen würden, antworteten 24 Personen (75 %), sie würden die intensive Behandlung über vier Tage vorziehen.

6. Diskussion

6.1. Einordnung der Ergebnisse in den Kontext bestehender Literatur

Die vorliegende Pilotstudie stellt die erste Evaluation des Bergen 4-Day Treatment (B4DT) im deutschen Versorgungskontext dar und untersucht die Durchführbarkeit, Wirksamkeit und Akzeptanz des konzentrierten Behandlungsformats außerhalb Norwegens. Das Behandlungsformat konnte erfolgreich im tagesklinischen Setting des Universitätsklinikums Hamburg Eppendorf implementiert werden. In Übereinstimmung mit früheren Forschungsergebnissen konnten wir bestätigen, dass das B4DT die Zwangssymptomatik, gemessen mittels Y-BOCS (primärer Outcome), signifikant reduziert. Es zeigten sich hierbei große Effektstärken mit einem Cohen's d von bis zu 2.04 (KI-95 % 1.43 bis 2.64) für den Gesamtscore. Auch für die selbstberichtete Zwangssymptomatik, gemessen mittels OCI-R (sekundärer Outcome) zeigte sich eine signifikante Reduktion mit großer Effektstärke (Cohen's $d = 1.19$, KI-95% 0.73 bis 1.64). Die Abnahme selbstberichteter Zwangssymptome (OCI-R) stützt zusätzlich die Validität der Ergebnisse und deutet auf eine generalisierte Symptomverbesserung aus Patient:innenperspektive hin.

Die globale Funktionsfähigkeit (GAF) der Proband:innen nahm durch die Behandlungswoche zu und auch die depressiven Symptome, gemessen mittels PHQ-9 konnten signifikant mit großer Effektstärke reduziert werden.

Ein Vergleich der Effektgrößen mit den bestehenden Daten aus Norwegen und der Implementierung in Island zeigt, dass die vorliegenden Ergebnisse als gemäßiger zu bewerten sind. Kvale et al. (2018) berichten beispielsweise eine große Effektstärke von Cohen's $d = 2.42$ von T0 zu T1, Launes et al. (2020) von Cohen's $d = 4.60$ von T0 zu T1, Davíðsdóttir et al. (2019) von Cohen's $d = 4.26$ von T0 zu T1, während wir eine immer noch als groß zu interpretierende, jedoch gemäßigtere Effektstärke von Cohen's $d = 2.04$ über den gleichen Zeitraum beobachten. Vergleicht man die Effektstärken über den Zeitraum von T0 zum Follow-Up (T2), so bleiben auch hier die erzielten Effekte hinter den norwegischen und isländischen Ergebnissen zurück.

Äquivalent dazu sind die Ansprech- und Remissionsraten geringer als in den norwegischen und isländischen Ergebnissen angegeben. Während zum Zeitpunkt T1 in

der vorliegenden Studie 51.5% auf die Behandlung ansprachen und sich die 33.3% der Proband:innen in Remission befanden, erzielten die Vergleichsstudien Response-Raten von bis zu über 90% und Remissionsraten von etwa 70% (Daviðsdóttir et al. 2019, Kvale et al. 2018, Launes et al. 2019b). Zum Zeitpunkt T2 sprachen in der vorliegenden Studie 62.5% der Proband:innen auf die Behandlung an, die Remissionsrate lag bei 46.9%. Auch zu diesem Zeitpunkt lagen die Ansprechraten mit bis zu 80% und die Remissionsraten mit etwa 70% in den Vergleichsstudien deutlich höher (Daviðsdóttir et al. 2019, Hansen et al. 2018, Holm et al. 2019b).

Die signifikante Abnahme der depressiven Symptome, die in dieser Studie beobachtet wurde, steht im Einklang mit vorangehenden Vergleichsstudien aus Norwegen (vgl. Hansen et al., 2018). Im Nachbeobachtungszeitraum manifestierte sich eine anhaltende Reduktion der depressiven Symptome. Diese Beobachtung legt die Interpretation nahe, dass die depressiven Symptome sekundär bei der Zwangsstörung auftreten und die effektive Behandlung der Primärerkrankung einen positiven Einfluss auf die Komorbidität hat.

Die signifikante Steigerung des allgemeinen Funktionsniveaus, das mittels GAF gemessen wurde, zeigt über die Symptomreduktion hinaus eine positive Auswirkung auf die Alltagsbewältigung und soziale Teilhabe der Patient:innen. Dies lässt darauf schließen, dass das B4DT nicht nur Zwangssymptome adressiert, sondern auch eine nachhaltige Verbesserung der Lebensqualität fördern kann.

6.2. Wirkmechanismen und Besonderheiten des B4DT

Warum die konzentrierte Expositionsbehandlung so effektiv ist, ist nicht mit Bestimmtheit zu sagen. Miegel et al. (2025) untersuchten verschiedene Faktoren, um die Wirkmechanismen der vier Behandlungstage besser zu verstehen und das Behandlungsformat so transparenter und besser optimierbar zu machen. Die Ergebnisse zeigen, dass das Sich-Einlassen der Teilnehmenden auf die Angst – „Leaning into anxiety“ – und die Therapieadhärenz während der Expositionsübungen das Outcome der Reduktion der Zwangssymptome durch die Behandlung beeinflussten. Aus technischer Perspektive scheint also ein Schlüsselfaktor für die Wirksamkeit von B4DT in der neuartigen Herangehensweise zu liegen, dass der Fokus darauf liegt, wie

die Teilnehmenden die Expositionen durchführen, ohne gleichzeitig subtile Vermeidungsstrategien anzuwenden. Das Konzept der LET-Technik beinhaltet die bewusste Annäherung an die Angst und das Unbehagen, die bei einer Exposition empfunden werden. Angst und Unbehagen werden als Gelegenheit herausgearbeitet, nachhaltige Veränderung zu erzielen, statt diese als Bedrohung zu sehen. Dies spiegelte sich auch in den Rückmeldungen der Teilnehmenden wider, die den inhaltlichen Fokus auf das „Arbeiten am Symptom“, das „Loslassen“ und das Stärken der persönlichen Motivation positiv hervorhoben.

Neben den Ergebnissen von Miegel et al. (2025) zeigen auch die Ergebnisse von Tjelle et al. (2021) und Tjelle et al. (2024), dass eine hohe Therapieadhärenz signifikant mit besseren Behandlungsergebnissen assoziiert ist und ein Prädiktor für die Symptomreduktion bei Zwangsstörungen ist. Patient:innen mit höherer Adhärenz, also konsequenterem Umsetzen der Expositionen mit Reaktionsmanagement, profitieren stärker und zeigen nachhaltigere Veränderungen der Zwangssymptomatik unabhängig von der Symptomschwere zu Beginn der Behandlung. Einige Eigenschaften des Behandlungsformats könnten eine höhere Adhärenz der Teilnehmenden begünstigen. Die therapeutische Intervention in der Gruppe mit einer simultanen Einzelbetreuung durch einen Therapeuten oder eine Therapeutin, erlaubt die Durchführung von Expositionen, die präzise auf die individuellen Bedürfnisse der Teilnehmer:innen abgestimmt sind. Den Teilnehmenden wird zugleich die Möglichkeit geboten, die Verhaltensmuster und Veränderungsprozesse anderer Personen, die unter Zwängen leiden, zu beobachten. Teilnehmende berichteten von einer „motivierenden Stimmung“ in der Gruppe und erlebten die therapeutische Betreuung als „intensiv“, „kompetent“ und „motivierend“.

Es könnte die Hypothese aufgestellt werden, dass die Konzentration des Formats auf einen Zeitraum von vier Tagen, einschließlich Psychoedukation und Planung der Integration von eigenständigen Expositionen, die Übertragung des therapeutischen Settings auf das Alltagsleben in den Mittelpunkt rückt. Teilnehmer:innen der vorliegenden Studie gaben das Feedback, dass der schnelle Übergang von Psychoedukation und Technik zu den Expositionsübungen und die Verbindung dieser Zusammenhänge ihnen gefiel. Darüber hinaus ist zu berücksichtigen, dass die

Patient:innen bei der Kürze und Intensität der Behandlung häufig eine hohe Therapiemotivation und den Glauben an die Wirksamkeit der Behandlung mitbringen. Obwohl B4DT als ein hoch wirksames Behandlungsformat einzustufen ist, ist die Identifikation spezifischer Komponenten, die maßgeblich für die hohe Wirksamkeit verantwortlich sind, also Gegenstand aktueller Forschung. Ein profundes Verständnis dieser Wirkmechanismen ist von entscheidender Relevanz, um die Implementierung in anderen Gesundheitssystemen erfolgreich zu gestalten und die Übertragbarkeit auf unterschiedliche Therapeut:innen-Teams und Versorgungsstrukturen sicherzustellen.

6.3. Unterschiede zu früheren Studien und mögliche Einflussfaktoren

Dies ist insbesondere vor dem Hintergrund relevant, dass die in der vorliegenden Pilotstudie beobachteten Effekte geringer ausfielen als in den norwegischen Originalarbeiten.

Der Grund hierfür kann in verschiedenen Bereichen liegen. Zum einen stammen die Vergleichsstudien zum Großteil von den Entwickler:innen selbst. Die Behandlung wurde somit von erfahreneren Teams durchgeführt, die eng mit dem Entwicklungskontext verbunden waren, was die Gefahr des „Allegiance Bias“ birgt, einer Verzerrung der Studienergebnisse durch eine besondere Präferenz oder Verbundenheit der Untersuchenden zu der untersuchten Intervention. Die Therapeut:innen in Hamburg wurden von dem Team der Entwickler:innen klinisch geschult und supervidiert, sodass auch hier ein Allegiance Bias zu diskutieren ist, jedoch weniger wahrscheinlich ist als bei der direkten Durchführung der Behandlung durch die Entwickler:innen (Jelinek et al. 2024b). Die Therapeut:innen in Hamburg hatten weniger Routine in der konzentrierten Expositionsbegleitung, was die Wirksamkeit beeinflussen könnte. Dagegen sprechen die Daten von Davíðsdóttir et al. (2019), die bei der ersten Implementation von B4DT in Island auch mit neu geschulten isländischen Therapeut:innen arbeiteten.

Unterschiede im Setting und Gesundheitssystem könnten eine Rolle gespielt haben. Durch die Pilotimplementierung im tagesklinischen Setting der Station PA7 am Universitätskrankenhaus Hamburg Eppendorf, auf welcher parallel andere

Versorgungsaufgaben bewältigt wurden, ergab sich ein logistischer Mehraufwand. Dagegen sprechen jedoch erneut die Daten aus Island sowie die ersten Daten aus Norwegen, die auch in zentralen, spezialisierten Zentren mit parallelen Versorgungsaufgaben durchgeführt wurden und direkt ein höheres Therapieansprechen erzielten (Daviðsdóttir et al. 2019, Havnen et al. 2013). In der bestehenden Literatur berichten Kvale et al. (2020) in einer randomisierten, placebokontrollierten Studie mit schwer behandelbaren Patient:innen von Ansprech- und Remissionsraten von 27.3 % respektive 56.5 % (siehe Tabelle 1), die mit unseren Daten vergleichbar sind. Dies wirft die Frage auf, ob die in der vorliegenden Studie beobachteten geringeren Effektstärken teilweise durch Unterschiede in den Stichprobencharakteristika erklärbar sind. Aus einer Prädiktoranalyse von Hansen et al. (2019) geht hervor, dass die Ausprägung einer depressiven Symptomatik vor der Behandlung mit den Chancen auf eine langfristige Remission zusammenhing, indem eine erhöhte Depressivität mit einer geringeren Wahrscheinlichkeit einer erfolgreichen Behandlung assoziiert war. Zudem konnte gezeigt werden, dass ein höherer initialer Schweregrad der Zwangssymptomatik prädiktiv für ein ungünstigeres Therapieansprechen war (Hansen et al. 2019, Leeuwerik et al. 2019). Der Vergleich der Ausgangswerte (T0) zeigt jedoch, dass der mittlere Y-BOCS-Wert in unserer Stichprobe ($M = 25.18$; $SD = 3.96$) in einem ähnlichen Bereich liegt wie in früheren Studien (von $M = 23,5$ bei Havnen et al. (2013) bis $M = 28.8$ bei Daviðsdóttir et al. (2019)). Auch der Anteil an Patient:innen mit komorbider Depression (18.2 %) und der mittlere PHQ-9-Wert ($M = 9.91$; $SD = 5.36$) lagen im Bereich der Vergleichsstudien, beispielsweise bei Hansen et al. (2018) mit 26.2 % und einem PHQ-9-Mittelwert von $M = 11.54$. Diese Daten deuten darauf hin, dass die Ausgangscharakteristika der Stichprobe die beobachteten Unterschiede in der Effektstärke nicht vollständig erklären.

Ein schwer messbarer Unterschied der Charakteristika könnte im Kontext der COVID-19-Pandemie entstanden sein. Die Pandemie brachte Einschränkungen in der psychotherapeutischen Therapie von Menschen mit Zwangsstörungen mit sich (Davide et al. 2020, Tandt et al. 2022), sodass in den Jahren der Pandemie mehr Menschen therapienaiv blieben. Gleichzeitig stellte die Pandemie für viele Menschen mit psychischen Erkrankungen eine große Mehrbelastung dar (Pfefferbaum und North

2020). Menschen mit Zwangsstörungen, insbesondere mit Waschzwängen, gelten dabei als besonders vulnerable Gruppe (Davide et al. 2020, Jelinek et al. 2021, Knowles und Olatunji 2021). Der Effekt einer konzentrierten Expositionstherapie in einer Gruppe nach einer Zeit ohne Therapie und mit gesellschaftlichen Einschränkungen könnte somit einen größeren Effekt haben als in größerem zeitlichem Abstand zur Pandemie. Zu den Auswirkungen der Pandemie auf den Krankheits- und Therapieverlauf wurden die Teilnehmenden sowohl in der vorliegenden als auch den norwegischen Studien jedoch nicht befragt.

Silver et al. (2023) verglichen Teilnehmer:innen, bei denen B4DT gut wirksam war mit den Proband:innen, bei denen die Behandlungswoche am wenigsten wirksam war. Der Vergleich ergab, dass Teilnehmer:innen, deren Funktionalität in verschiedenen Lebensbereichen durch die Zwangsstörung vor der Behandlung stärker eingeschränkt war, ein geringeres Ansprechen auf die Behandlung zeigten. Die Aussagekraft der Ergebnisse ist dabei begrenzt, da es sich um sehr kleine Stichprobengrößen handelte. Nichtsdestotrotz wirft dies die Frage auf, ob die Selektion der Proband:innen in der Screening-Phase sowie die Behandlungsvorbereitung Unterschiede in der Wirksamkeit ausmachen und auch ein Grund für die Divergenz verschiedener Ergebnisse in der Effektivität der Behandlung sein könnten. Die Screening-Phase sowie die Vorbereitung der Patient:innen sind nicht Bestandteil des Behandlungsmanuals und könnten sich von Standort zu Standort erheblich unterscheiden. Es sind weitere Untersuchungen erforderlich, um die Effektivität von B4DT in verschiedenen Patient:innengruppen und -situationen zu ermitteln und so das individuell wirksamste Therapieformat zu bestimmen. Dafür könnte sinnvoll sein, Screening-Kriterien sowie die Vorgespräche und Vorbereitungsphase als Teil des Behandlungsformats zu manualisieren.

Auch die Motivation für die Behandlung sowie die Bereitschaft, Zwangssymptome zu eliminieren und eine Veränderung zu erzielen, werden als potenzielle prädiktive Faktoren für den Behandlungserfolg diskutiert (Jelinek et al. 2024b, Silver et al. 2023). Der experimentelle Charakter einer Pilotstudie sowie der intensive Charakter der konzentrierten Behandlungswoche könnten einerseits sehr motivierte Patient:innen anziehen mit einem möglichen Einfluss von Erwartungseffekten. Andererseits könnten stark betroffene Patient:innen oder Patient:innen mit vielen vorangegangenen

Therapieversuchen die Therapie als letzte Behandlungsmöglichkeit sehen und ein höheres Risiko für eine größere Enttäuschung bei ausbleibender Effektivität mitbringen.

6.4. Limitationen und Stärken der Studie

Eine mögliche Selektivität der Stichprobe kann als Limitation der vorliegenden Studie angesehen werden. Ein möglicher Rekrutierungsbias schränkt die externe Validität der Ergebnisse ein.

Eine wesentliche Limitation der vorliegenden Studie ist das Fehlen einer unbehandelten Kontrollgruppe, um spontane Symptomverbesserungen, unspezifische Effekte (Erwartungshaltungen, erhöhte therapeutische Aufmerksamkeit, Einfluss von Zeit) ausschließen oder von spezifischen Behandlungseffekten unterscheiden zu können. In der randomisierten Kontrollstudie von Launes et al. (2019a) wurde B4DT im Vergleich zu einer Wartelistenkontrolle untersucht. Die unbehandelte Kontrollgruppe (Warteliste) zeigte keine Verbesserung der Zwangssymptome, während die Y-BOCS-Scores der Teilnehmenden an der Behandlung signifikant reduziert werden konnten. Die Metaanalyse von Öst et al. (2015) zeigt, dass in Studien zu kognitiver Verhaltenstherapie die Effekte gegenüber unbehandelten Kontrollgruppen konsistent größer ausfallen. Es gibt also keinen Anhalt für eine Symptomreduktion nur aufgrund verstrichener Zeit oder Erwartungseffekten, nichtsdestotrotz ist die Durchführung zukünftiger randomisiert-kontrollierter Studien im deutschen Versorgungskontext essenziell, um die Wirksamkeit des B4DT gegenüber Standardbehandlungen evidenzbasiert zu überprüfen.

Eine weitere Limitation ist die kleine Stichprobengröße, durch die die statistische Power eingeschränkt ist. Zudem erstreckt sich der Beobachtungszeitraum der hier ausgewerteten Daten der Pilotstudie auf drei Monate. Durch die hier fehlenden Langzeitdaten lässt sich noch keine Aussage über die weitergehende Stabilität der erzielten Effekte treffen.

Eine methodische Einschränkung der Studie betrifft die Verwendung von Complete-Case-Analysen. Von den ursprünglich eingeschlossenen Patient:innen konnten 6.1% ($N = 2$) aufgrund fehlender Nachuntersuchungsdaten nicht berücksichtigt werden. Obwohl

die geringe Drop-out-Rate auf eine gute Akzeptanz des Behandlungsformats hinweist, kann der Ausschluss dieser Fälle ein Bias-Risiko bergen. Complete-Case-Analysen setzen voraus, dass die fehlenden Daten vollständig zufällig sind und mit keiner anderen Variable, wie beispielsweise der Schwere der Zwangsstörung, systematisch in Zusammenhang stehen. Sollte dies nicht der Fall sein, ist eine selektive Verzerrung und leichte Überschätzung der Effektstärken möglich.

Stärken der Studie liegen in der Erhebung der Daten durch unabhängige Rater:innen und die Verwendung der Y-BOCS mit ausgezeichneter Interrater-Reliabilität. Die Orientierung am originalen B4DT-Protokoll gewährleistet eine hohe Vergleichbarkeit mit den bisherigen Studien. Die Erhebung multipler Outcomes (klinisch, selbstberichtete Symptome und Funktionsniveau) ermöglicht eine differenzierte Betrachtung des Therapieeffekts. Die niedrige Rate an Therapieabbrüchen spricht für eine praktische Umsetzbarkeit, Akzeptanz bei Patient:innen und gute Machbarkeit im klinischen Alltag. Dass 68.8 % der Teilnehmenden „sehr zufrieden“ und 21.9 % „weitgehend zufrieden“ waren, unterstreicht die große Akzeptanz der Behandlung. Die Qualität der Behandlung wurde von der Hälfte der Teilnehmenden als „ausgezeichnet“ und von 43.8 % als „gut“ bezeichnet.

Eine weitere Stärke der Studie liegt in der ersten Untersuchung des B4DT im deutschsprachigen Raum. Dadurch erweitert die Studie die bisher vornehmlich aus dem skandinavischen Raum stammende Evidenz und liefert erste Hinweise zur Anwendbarkeit und Wirksamkeit des B4DT in einem neuen kulturellen und gesundheitssystemischen Kontext. Dies stellt zum einen wichtigen Schritt in der externen Validierung des Behandlungsformats dar. Zum anderen birgt dies Potential vor dem Hintergrund der bestehenden Versorgungsdefizite bei Zwangsstörungen in Deutschland.

6.5. Implikationen für Forschung und Praxis

Die Implementierung des B4DT im deutschen Versorgungssystem wäre mit strukturellen Anpassungen verbunden: Das Format erfordert eine kurzfristige Koordination mehrerer Therapeut:innen sowie eine logistische Organisation intensiver Settings. Diese Hürde einer Implementierung des neuen Behandlungsformats zu

überwinden ist jedoch mit einem großen Mehrwert für die Behandlung von Menschen mit Zwangsstörungen in Deutschland verbunden. Aktuell führen ein signifikanter Anteil an Psychotherapeut:innen in Deutschland keine Expositionen durch oder setzen diese nur unzureichend um (Külz et al. 2010, Moritz et al. 2019, Voderholzer et al. 2015). Das standardisierte und manualisierte Vorgehen in der Expositionsbehandlung begegnet diesem Problem und könnte somit zu einer Qualitätssteigerung der Behandlung von Menschen mit Zwangsstörungen beitragen. Das B4DT stellt in diesem Kontext ein klar strukturiertes, intensiv betreutes Interventionsformat dar, das gut in Weiterbildungsprozesse integriert werden könnte, äquivalent zu der Schulung der Therapeut:innen für die vorliegende Pilotstudie.

So könnte also der Charakter des Behandlungsformats zu einer gesichert expositionsbasierten, standardisierten Behandlung der Menschen mit Zwangsstörungen in Deutschland führen und die spezifischen Weiterbildungsprozesse für Therapeut:innen verbessern.

Der Vergleich mit einer Kontrollgruppe von stationären Patient:innen bei Jelinek et al. (2024b) indiziert die Überlegenheit von B4DT gegenüber der herkömmlichen stationären Therapie. Als wirksamere und schnellere Therapie könnte B4DT die bisher bestehende Versorgungslücke (siehe 3.3) schließen, Wartezeiten verkürzen und zur Entlastung ambulanter Versorgungsstrukturen beitragen. Abgesehen von der strukturellen und gesundheitssystemischen Ebene ist die Bedeutung einer wirksamen und praktikablen Therapie von Zwangsstörungen für die Betroffenen hervorzuheben. Gegenüber langen Wartezeiten und frustranen, für die Betroffenen anstrengenden Therapieversuchen mit hohen Abbruchraten kann B4DT als schnelles und wirksames Behandlungsformat Menschen motivieren, sich Hilfe zu suchen. Die schnelle Symptomreduktion sowie der intensive therapeutische Fokus in kurzer Zeit könnten eine Entstigmatisierung und Förderung der Motivation zur Behandlung bewirken, insbesondere bei Menschen mit Zwangsstörungen, die aus Scham oder Unwissenheit lange keine Hilfe gesucht haben. Die Aussicht auf eine kurzfristige und effektive Behandlung kann hier als Türöffner fungieren für den großen Anteil von Menschen mit Zwangsstörungen, die in Deutschland aktuell unbehandelt bleiben (Jacobi et al. 2014). Über die individuellen Effekte hinaus kann das Bergen 4-Day Treatment auch eine

wichtige relationale Entlastung für Angehörige von Menschen mit Zwangsstörungen bewirken. Zwangsstörungen beeinflussen häufig das gesamte familiäre System, indem sie zu finanzieller, beruflicher, sozialer und familiärer Belastung, und so zu Mitbeteiligung und emotionaler Erschöpfung der Angehörigen führen (El-Slamon et al. 2022). Eine rasch wirksame Behandlung kann daher nicht nur Symptome, sondern auch Beziehungsmuster verändern – und damit Hoffnung und Entlastung in ein Umfeld bringen, das oft über Jahre von der Erkrankung mitbetroffen war.

Trotz der Wirksamkeit und des großen Potenzials von B4DT für die Behandlung von Menschen mit Zwangsstörungen bedarf es weiterer Untersuchungen mittels randomisierter und kontrollierter Studien mit größeren Stichproben und längeren Follow-Up-Zeiträumen, um die Nachhaltigkeit der Effekte sowie die Übertragbarkeit auf unterschiedliche Versorgungssettings zu evaluieren. Die Identifikation von Prädiktoren für den Therapieerfolg könnte zu einer individualisierten Behandlung beitragen und die Effizienz aus gesundheitsökonomischer Sicht steigern sowie die Behandlung für das Individuum effektiver gestalten. Zudem könnten Prädiktoren für eine erfolgreiche Behandlung helfen, B4DT weiterzuentwickeln und zu optimieren, um das Potential von B4DT auszuschöpfen, Menschen mit Zwangsstörungen in Deutschland zu entlasten und ihre Lebensqualität zu steigern.

7. Zusammenfassung

Deutsche Version

Zwangsstörungen zählen zu den häufigsten und zugleich behandlungsintensivsten psychischen Erkrankungen mit einer Lebenszeitprävalenz von etwa 1–3 %. Trotz der hohen Evidenz für die Wirksamkeit der kognitiven Verhaltenstherapie mit Exposition und Reaktionsmanagement bestehen in der klinischen Versorgung weiterhin erhebliche Herausforderungen, darunter eingeschränkte Verfügbarkeit spezialisierter Angebote, lange Wartezeiten und hohe Therapieabbruchraten. Das in Norwegen entwickelte Bergen 4-Day Treatment ist ein konzentriertes, intensiviertes Behandlungsformat, das über vier Tage in Gruppen von drei bis sechs Patient:innen mit therapeutischer Eins-zu-eins-Betreuung durchgeführt wird. Es zeigte in mehreren Studien hohe Effektivität und Akzeptanz. Daten zur Übertragbarkeit in andere Gesundheitssysteme sind bislang begrenzt. Die vorliegende Pilotstudie markiert die erste Implementation des Bergen 4-Day Treatment in Deutschland und liefert somit einen wesentlichen Beitrag zur internationalen Evaluierung dieses innovativen Kurzzeitformats. Im tagesklinischen Setting des Universitätsklinikums Hamburg-Eppendorf wurden 33 Patient:innen mit Zwangsstörungen nach dem B4DT-Manual behandelt. Erhebungen fanden vor der Behandlung (T0), eine Woche danach (T1) und drei Monate nach Abschluss (T2) statt. Die Ergebnisse zeigten signifikante Reduktion der Zwangssymptome mit großen Effektstärken ($\eta^2 = .72$) über alle Messzeitpunkte. Rund ein Drittel der Teilnehmenden erreichte eine Remission. Die Behandlung führte zu einer signifikanten Verbesserung depressiver Symptome und der globalen Funktionsfähigkeit.

Die hohe Patient:innenzufriedenheit von über 90% und die niedrige Drop-out-Rate unterstreichen die Akzeptanz und Praktikabilität des Formats.

Die Effektstärken waren im Vergleich zu den norwegischen Originalstudien und der isländischen Replikation etwas geringer. Die Frage, warum die Effektstärken hinter den Resultaten der Originalstudien zurückbleiben sowie die Identifikation individueller prädiktiver Faktoren für den Therapieerfolg, stellt weiterhin ein zentrales Forschungsthema dar und erfordert weiterführende Untersuchungen. Insgesamt zeigen

die Ergebnisse, dass das B4DT im deutschen Versorgungssystem sicher, machbar und wirksam durchgeführt werden kann. Die Studie liefert damit einen wichtigen Schritt zur Implementierung und Anpassung evidenzbasierter Kurzzeitinterventionen für Zwangsstörungen in Deutschland und legt die Grundlage für größere, kontrollierte Wirksamkeitsstudien.

Englische Version

Obsessive-compulsive disorder is among the most common and treatment-intensive mental illnesses, with a lifetime prevalence of approximately 1–3%. Despite strong evidence for the efficacy of cognitive behavioral therapy with exposure and response management, significant challenges persist in clinical practice, including the limited availability of specialized treatments, prolonged waiting periods, and high treatment dropout rates. The Bergen 4-Day Treatment is a concentrated, intensive treatment format, developed in Norway. It is a group-based treatment, with each group consisting of three to six patients, and is accompanied by one-to-one therapeutic support. The effectiveness and acceptance of B4DT have been demonstrated in several studies. Data on its transferability to other health care systems are limited so far. This pilot study marks the first implementation of Bergen 4-Day Treatment in Germany, thereby providing a substantial contribution to the international evaluation of this innovative short-term format. In the day clinic setting of the University Medical Centre Hamburg-Eppendorf, 33 patients diagnosed with obsessive-compulsive disorder were treated according to the B4DT manual. Patients were assessed at three distinct timepoints: before treatment (T0), one week after treatment (T1) and three months after treatment (T2). The results demonstrated a significant reduction in obsessive-compulsive symptoms with large effect sizes ($\eta^2 = .72$) across all measurement points. Approximately one third of participants achieved remission. The treatment resulted in a significant improvement in depressive symptoms and overall functioning. The high patient satisfaction rate of over 90% and the low drop-out rate are indicative of the acceptance and practicability of the format.

The effect sizes were found to be somewhat lower than in the original Norwegian studies and the Icelandic replication. The question of why the effect sizes are below the

results of the original studies, as well as the identification of individual predictive factors for therapeutic success, remains a central research topic and requires further investigation. The results demonstrate that B4DT can be implemented safely, feasibly and effectively within the German healthcare system. The study therefore signifies an important step towards the implementation and adaptation of evidence-based short-term interventions for obsessive-compulsive disorder in Germany, paving the way for larger, controlled efficacy studies.

8. Literaturverzeichnis

- Abramovitch A, Abramowitz JS, Riemann BC, McKay D (2020) Severity benchmarks and contemporary clinical norms for the Obsessive-Compulsive Inventory-Revised (OCI-R). *Journal of Obsessive-Compulsive and Related Disorders*. 27.
- American Psychiatric Association, A. P. A. (2025) Diagnostisches und Statistisches Manual Psychischer Störungen – Textrevision – DSM-5-TR® Deutsche Ausgabe. Falkai P, Wittchen H-U, Döpfner M, Gaebel W, Maier W, Rief W, Saß H, Zaudig M (Hrsg.) Hogrefe Verlag GmbH & Co. KG, Göttingen.
- Ansari E, Mishra S, Tripathi A, Kar SK, Dalal PK (2020) Cross-sectional study of internalised stigma and medication adherence in patients with obsessive compulsive disorder. *General Psychiatry*. 33.
- Attkisson CC, Zwick R (1982) The Client Satisfaction Questionnaire: Psychometric Properties and Correlations with Service Utilization and Psychotherapy Outcome. *Evaluation and Program Planning*. 5, 233-237.
- Backenstraß M (2022) Zwangsstörungen und Verwandte Störungen in DSM-5 und ICD-11. *Verhaltenstherapie & Verhaltensmedizin*. 43 (1) 9-20.
- Bobes J, García-Portilla M-P, Bascaran M-T, Sáiz P-A, Bobes-Bascarán M-T, Bousoño M (2007) Quality Of Life In Obsessive-Compulsive Disorder. In: *Quality of Life Impairment in Schizophrenia, Mood and Anxiety Disorders – New Perspectives on Research and Treatment*. Ritsner MS, Awad AG (Hrsg.) Springer, Dordrecht, Niederlande, 293-303.
- Brakouliasa V, Starcevic V, Bellochb A, Brownc C, Ferrao YA, Fontenelle LF, Lochnerg C, Marazziti D, Matsunagai H, Miguelj EC, Reddyk YCJ, do Rosariol MC, Shavittj RG, Shyam Sundark A, Steing DJ, Torresm AR, Viswasam K (2017) Comorbidity, age of onset and suicidality in obsessive-compulsive disorder (OCD): An international collaboration. *Comprehensive Psychiatry*. 76, 79-86.
- Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2. Edition., Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Davide P, Andrea P, Martina O, Andrea E, Davide D, Mario A (2020) The impact of the COVID-19 pandemic on patients with OCD: Effects of contamination symptoms and remission state before the quarantine in a preliminary naturalistic study. *Psychiatry Research*. 291, 113213.
- Davíðsdóttir SD, Sigurjónsdóttir O, Ludvigsdóttir SJ, Hansen B, Laukvik IL, Hagen K, Björgvinsson T, Kvale G (2019) Implementation of the Bergen 4-day treatment for obsessive compulsive disorder in Iceland. *Clinical Neuropsychiatry*. 16 (1) 33-38.
- El-Slalom M, Al-Moteri M, Plummer V, Alkarani AS, Ahmed MG (2022) Coping Strategies and Burden Dimensions of Family Caregivers for People Diagnosed with Obsessive-Compulsive Disorder. *Healthcare (Basel)*. 10 (3).
- Elsouri KN, Heiser SE, Cabrera D, Alqurneh S, Hawat J, Demory ML (2024) Management and Treatment of Obsessive-Compulsive Disorder (OCD): A Literature Review. *The Cureus Journal of Medical Science*. 16 (5) e60496.
- Endres D, Schiele MA, von Zedtwitz K, Dressle RJ, Maier A, Hohagen F, Baldermann JC, Coenen VA, Jelinek L, Domschke K, Voderholzer U (2025) Obsessive-compulsive

- disorder - A state-of-the-art review. *Neuroscience & Biobehavioral Reviews*. 177, 106320.
- Fawcett EJ, Power H, Fawcett JM (2020) Women Are at Greater Risk of OCD Than Men: A Meta-Analytic Review of OCD Prevalence Worldwide. *The Journal of Clinical Psychiatry*. 23;81 (4).
- Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, Salkovskis PM (2002) The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment* 14 (4) 485-495.
- Freud, S (1952) Die Abwehr-Neuropsychosen. Versuch einer psychologischen Theorie der akquirierten Hysterie, vieler Phobien und Zwangsvorstellungen und gewisser halluzinatorischer Psychosen. In: Erster Band. Werke aus den Jahren 1892-1899. Freud A, Bibring E, Hoffer W, Kris E, Isakower O (Hrsg.) Imago Publishing Co., London UK, 59-74.
- Gamez W, Chmielewski M, Kotov R, Ruggero C, Suzuki N, Watson D (2014) The brief experiential avoidance questionnaire: development and initial validation. *European Journal of Psychological Assessment*. 26 (1) 35-45.
- Gönner S, Leonhart R, Ecker W (2007) Das Zwangsinventar OCI-R – die deutsche Version des Obsessive-Compulsive Inventory-Revised: Ein kurzes Selbstbeurteilungsinstrument zur mehrdimensionalen Messung von Zwangssymptomen. *Psychotherapie Psychosomatik Medizinische Psychologie*. 57, 395-404.
- Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS (1989) The Yale-Brown Obsessive Compulsive Scale: Development, Use, and Reliability. *Archives of General Psychiatry*. 46, 1006-1011.
- Grant JE, Mancebo MC, Pinto A, Williams KA, Eisen JL, Rasmussen SA (2007) Late-onset obsessive compulsive disorder: Clinical characteristics and psychiatric comorbidity. *Psychiatry Research*. 152, 21-27.
- Griesinger W (1868) Über einen wenig bekannten psychopathischen Zustand. *Archiv für Psychiatrie und Nervenkrankheiten*. 1, 626-635.
- Haland AT, Vogel PA, Lie B, Launes G, Pripp AH, Himle JA (2010) Behavioural group therapy for obsessive-compulsive disorder in Norway. An open community-based trial. *Behaviour Research and Therapy*. 48 (6) 547-554.
- Hand I, Büttner-Westphal H (1991) Die Yale-Brown Obsessive Compulsive Scale (Y-BOCS): Ein halbstrukturiertes Interview zur Beurteilung des Schweregrades von Denk- und Handlungszwängen. *Verhaltenstherapie*. 1, 223-225.
- Hansen B, Hagen K, Ost LG, Solem S, Kvale G (2018) The Bergen 4-Day OCD Treatment Delivered in a Group Setting: 12-Month Follow-Up. *Frontiers in Psychology*. 9, 639.
- Hansen B, Kvale G, Hagen K, Havnen A, Ost LG (2019) The Bergen 4-day treatment for OCD: four years follow-up of concentrated ERP in a clinical mental health setting. *Cognitive Behaviour Therapy*. 48 (2) 89-105.
- Havnen A, Hansen B, Haug ET, Prescott P, Kvale G (2013) Intensive Group Treatment of Obsessive-Compulsive-Disorder: A Pilot Study. *Clinical Neuropsychiatry*. 10 (1) 48-55.

- Havnen A, Hansen B, Öst L-G, Kvale G (2014) Concentrated ERP delivered in a group setting: An effectiveness study. *Journal of Obsessive-Compulsive and Related Disorders*. 3 (4) 319-324.
- Havnen A, Hansen B, Ost LG, Kvale G (2017) Concentrated ERP Delivered in a Group Setting: A Replication Study. *Behavioural and Cognitive Psychotherapy*. 45 (5) 530-536.
- Hilsenroth MJ, Ackerman SJ, Blagys MD, Baumann BD, Baity MR, Smith SR, Price JL, Smith CL, Heindselman TL, Mount MK, Holdwick DJ (2000) Reliability and Validity of DSM-IV Axis V. *American Journal of Psychiatry*. 157 (11) 1858-1863.
- Hollander E, Doernberg E, Shavitt R, Waterman RJ, Soreni N, Veltman DJ, Sahakian BJ, Fineberg NA (2016) The cost and impact of compulsivity: A research perspective. *European Neuropsychopharmacology*. 26 (5) 800-809.
- Holm SEH, Hansen B, Kvale G, Eilertsen T, Hagen K, Solem S (2019a) Is sensory processing sensitivity related to treatment outcome in concentrated exposure and response prevention treatment for obsessive-compulsive disorder?. *Journal of Obsessive-Compulsive and Related Disorders*. 23, 100486.
- Holm SEH, Hansen B, Kvale G, Eilertsen T, Johnsen BH, Hystad SW, Solem S (2019b) Dispositional resilience in treatment-seeking patients with obsessive-compulsive disorder and its association with treatment outcome. *Scandinavian Journal of Psychology*. 60, 243-251.
- Iversen HM, Eide TO, Harvold M, Solem S, Kvale G, Hansen B, Hagen K (2022) The Bergen 4-day treatment for panic disorder: replication and implementation in a new clinic. *BMC Psychiatry*. 22 (1) 728.
- Jacobi F, Höfler M, Strehle J, Mack S, Gerschler A, Scholl L, Busch MA, Maske U, Hapke U, Gaebel W, Maier W, Wagner M, Zielasek J, Wittchen H-U (2014) Psychische Störungen in der Allgemeinbevölkerung. Studie zur Gesundheit Erwachsener in Deutschland und ihr Zusatzmodul Psychische Gesundheit (DEGS1-MH). *Nervenarzt*. 85, 77-87.
- Jacobson NS, Truax P (1991) Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology*. 59, 12-19.
- Janowitz D, Grabe HJ, Ruhrmann S, Ettelt S, Buhtz F, Hochrein A, Schulze-Rauschenbach S, Meyer K, Kraft S, Ferber C, Pukrop R, Freyberger HJ, Klosterkötter J, Falkai P, John U, Maier W, Wagner M (2019) Early Onset of Obsessive Compulsive Disorder and Associated Comorbidity. *Depression And Anxiety*. 26, 1012-1017.
- Jaurrieta N, Jimenez-Murcia S, Menchón JM, Del Pino Alonso M, Segalas C, Álvarez-Moya EM, Labad J, Granero R, Vallejo J (2008) Individual versus group cognitive-behavioral treatment for obsessive-compulsive disorder: a controlled pilot study. *Psychotherapy Research*. 18 (5) 604-614.
- Jelinek L, Moritz S, Miegel F, Voderholzer U (2021) Obsessive-compulsive disorder during COVID-19: Turning a problem into an opportunity? *Journal of Anxiety Disorders*. 77, 102329.
- Jelinek L, Balzar A, Moritz S, Liebherz S, Yassari AH (2024a) Exposure therapy in patients with harm-related obsessive-compulsive disorder: The theory-practice gap and its relation to experiential avoidance and negative beliefs about exposure. *Journal of Obsessive-Compulsive and Related Disorders*. 40.

- Jelinek L, Serve A, Pampuch S, Scheunemann J, Schultz J, Miegel F, Hansen B, Hagen K, Bohnsack F, Gallinat J, Yassari AH (2024b) Exploring the Peaks and Potholes: Understanding positive and negative effects of concentrated exposure treatment for obsessive-compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders*. 43, 100913.
- Jelinek L, Hufenbach F, Miegel F, Scheunemann J, Yassari AH (2025) Konzentrierte Expositionstherapie bei Zwangsstörungen: das Bergen 4-Day Treatment (B4DT). *PiD – Psychotherapie im Dialog*. 26, 78-82.
- Jones SH, Thornicroft G, Coffey M, Dunn G (1995) A brief mental health outcome scale-reliability and validity of the Global Assessment of Functioning (GAF). *The British Journal of Psychiatry*. 166 (5) 654-659.
- Kellerwessel W (2012) Philosophische Betrachtungen zum Thema 'Zwang' - sprachphilosophische, moralphilosophische und medizinethische Perspektiven. In: *Medizin, Zwang, Gesellschaft*. Ernst J-P, Kehl M, Thal M, Groß D (Hrsg.) Medizinisch Wissenschaftliche Verlagsgesellschaft, Aachen, 3-13.
- Kessler RC, Petukhova M, Sampson NA, Zaslavsky A, Wittchen H-U (2012) Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*. 21 (3) 169-184.
- Knowles KA, Olatunji BO (2021) Anxiety and safety behavior usage during the COVID-19 pandemic: The prospective role of contamination fear. *Journal of Anxiety Disorders*. 77, 102323.
- Kochar N, Ip S, Vardanega V, Sireau NT, Fineberg NA (2023) A cost-of-illness analysis of the economic burden of obsessive-compulsive disorder in the United Kingdom. *Comprehensive Psychiatry*. 127, 152422.
- Kroenke K, Spitzer RL, Williams JBW (2001) The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*. 16, 606-613.
- Külz AK, Hassenpflug K, Riemann D, Linster HW, Dornberg M, Voderholzer U (2010) Ambulante psychotherapeutische Versorgung bei Zwangserkrankungen – Ergebnisse einer anonymen Therapeutenbefragung. *Psychotherapie Psychosomatik Medizinische Psychologie*. 60 (6) 194-201.
- Kvale G, Hansen B, Bjorgvinsson T, Bortveit T, Hagen K, Haseth S, Kristensen UB, Launes G, Ressler KJ, Solem S, Strand A, van den Heuvel OA, Ost LG (2018) Successfully treating 90 patients with obsessive compulsive disorder in eight days: the Bergen 4-day treatment. *BMC Psychiatry*. 18 (1) 323.
- Kvale G, Hansen B, Hagen K, Abramowitz JS, Bortveit T, Craske MG, Franklin ME, Haseth S, Himle JA, Hystad S, Kristensen UB, Launes G, Lund A, Solem S, Ost LG (2020) Effect of D-Cycloserine on the Effect of Concentrated Exposure and Response Prevention in Difficult-to-Treat Obsessive-Compulsive Disorder: A Randomized Clinical Trial. *JAMA Netw Open*. 3 (8) e2013249.
- Laireiter A-R, Baumann U, Stieglitz R-D (2001) Soziodiagnostik: Soziales Netzwerk, Soziale Unterstützung und Soziale Anpassung. In: *Psychodiagnostik in Klinischer Psychologie, Psychiatrie, Psychotherapie* (2. Ed.). Stieglitz R-D, Baumann U, Freyberger HJ (Hrsg.) Georg Thieme Verlag, Stuttgart, 210-228.
- Lange C, Heuft G (2002) Die Beeinträchtigungsschwere in der psychosomatischen und psychiatrischen Qualitätssicherung: Global Assessment of Functioning Scale

- (GAF) vs. Beeinträchtigungs-Schwere-Score (BSS). Zeitschrift für Psychosomatische Medizin und Psychotherapie. 48 (3) 256-269.
- Launes G, Hagen K, Sunde T, Ost LG, Klovning I, Laukvik IL, Himle JA, Solem S, Hystad SW, Hansen B, Kvale G (2019a) A Randomized Controlled Trial of Concentrated ERP, Self-Help and Waiting List for Obsessive- Compulsive Disorder: The Bergen 4-Day Treatment. *Frontiers in Psychology*. 10, 2500.
- Launes G, Laukvik IL, Sunde T, Klovning I, Hagen K, Solem S, Öst L-G, Hansen B, Kvale G (2019b) The Bergen 4-Day Treatment for Obsessive-Compulsive Disorder: Does It Work in a New Clinical Setting? *Frontiers in Psychology*. 10.
- Launes G, Hagen K, Ost LG, Solem S, Hansen B, Kvale G (2020) The Bergen 4-Day Treatment (B4DT) for Obsessive-Compulsive Disorder: Outcomes for Patients Treated After Initial Waiting List or Self-Help Intervention. *Frontiers in Psychology*. 11, 982.
- Leeuwerik T, Cavanagh K, Strauss C (2019) Patient adherence to cognitive behavioural therapy for obsessive-compulsive disorder: A systematic review and meta-analysis. *Journal of Anxiety Disorders*. 68, 102135.
- Lochner C, Fineberg NA, Zohar J, van Ameringen M, Juven-Wetzler A, Altamura AC, Cuzen NL, Hollander E, Denys D, Nicolini H, Dell’Osso B, Pallanti S, Stein DJ (2014) Comorbidity in obsessive–compulsive disorder (OCD): A report from the International College of Obsessive–Compulsive Spectrum Disorders (ICOCS). *Comprehensive Psychiatry*. 55, 1513-1519.
- Mack S, Jacobi F, Gerschler A, Strehle J, Höfler M, Busch MA, Maske UE, Hapke U, Seiffert I, Gaebel W, Zielasek J, Maier W, Wittchen H-U (2014) Self-reported utilization of mental health services in the adult German population – evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). *International Journal of Methods in Psychiatric Research*. 23 (3) 289-303.
- Marques L, Chosak A, Simon NM, Phan D-M, Wilhelm S, Pollack M (2010) Rating Scales for Anxiety Disorders. In: *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health*. Baer L, Blais MA (Hrsg.) Humana Press c/o Springer Science + Business Media, LLC, New York, USA, 37-72.
- Mataix-Cols D, Fernandez de la Cruz L, Nordsletten AE, Lenhard F, Isomura K, Simpson HB (2016) Towards an international expert consensus for defining treatment response, remission, recovery and relapse in obsessive-compulsive disorder. *World Psychiatry*. 15 (1) 80-81.
- Meyer V (1966) Modification of expectations in cases with obsessional rituals. *Behaviour Research and Therapy*. 4 (4) 273-280.
- Miegel F, Scheunemann J, Pampuch S, Schultz J, Hansen B, Hagen K, Gallinat J, Zapf A, Yassari AH, Jelinek L (2025) Exploring the Black Box: What Happens during Brief Concentrated Exposure and Response Prevention for Obsessive-Compulsive Disorder?. *Psychotherapy and Psychosomatics*. 94 (3) 162-174.
- Moritz S, Kulz A, Voderholzer U, Hillebrand T, McKay D, Jelinek L (2019) "Phobie a deux" and other reasons why clinicians do not apply exposure with response prevention in patients with obsessive-compulsive disorder. *Cognitive Behaviour Therapy*. 48 (2) 162-176.
- Oaklander M (2018) Bjarne Hansen and Gerd Kvale – Speeding Up Therapy. *The Health Care 50 – Fifty people transforming healthcare*. TIME Magazine, New York, USA [Online im Internet] URL: <https://time.com/collection/health-care-50>

- 50/5425089/gerd-kvale-and-bjarne-hansen/ [Stand: 16.August 2025, 13:54 Uhr].
- Oberbeck A, Stengler K, Steinberg H (2013) Die Geschichte der Zwangserkrankung: Ihre Stellung im Wandel der psychiatrischen Formenlehre bis Anfang des 20. Jahrhunderts. *Fortschritte der Neurologischen Psychiatrie*. 81 (12) 706-714.
- Olatunji BO, Davis ML, Powers MB, Smits JA (2013) Cognitive-behavioral therapy for obsessive-compulsive disorder: a meta-analysis of treatment outcome and moderators. *Journal of Psychiatric Research*. 47 (1) 33-41.
- Öst LG, Havnen A, Hansen B, Kvale G (2015) Cognitive behavioral treatments of obsessive-compulsive disorder. A systematic review and meta-analysis of studies published 1993-2014. *Clinical Psychology Review*. 40, 156-169.
- Pfefferbaum B, North CS (2020) Mental Health and the Covid-19 Pandemic. *New England Journal of Medicine*. 383 (6) 510-512.
- Deutsche Psychotherapeuten Vereinigung e.V. (2021) Report Psychotherapie 2021, 2. Auflage. Berlin [Online im Internet]. URL: https://www.dptv.de/fileadmin/Redaktion/Bilder_und_Dokumente/Wissensdatenbank_oeffentlich/Report_Psychotherapie/DPTV_Report_Psychotherapie_2021.pdf [Stand: 16. August.2025, 18:12].
- Ralph S (2019) Gerd Kvale and Bjarne Hansen – The Bergen OCD Format. The OCD-Stories, Episode 169 [Podcast, Online im Internet] URL: <https://open.spotify.com/show/13YuzynkMsgjtR3K3Brhle> [Stand: 16.August 2025, 16:19 Uhr].
- Rosa AC, Diniz JB, Fossaluza V, Torres AR, Fontenelle LF, De Mathis AS, da Conceição Rosário M, Miguel EC, Shavitt RG (2012) Clinical correlates of social adjustment in patients with obsessive-compulsive disorder. *Journal of Psychiatric Research*. 46, 1286-1292.
- Rosenberg M (1965) *Society and the Adolescent Self-Image*. Princeton University Press, Princeton, New Jersey, 338.
- Ruscio AM, Stein DJ, Chiu WT, Kessler RC (2010) The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry*. 15 (1) 53-63.
- Salkovskis PM, Ertle A, Kirk J (2018) Zwangsstörung. In: *Lehrbuch der Verhaltenstherapie, Band 2* (4. Auflage). Margraf J, Schneider S (Hrsg.) Springer Berlin, Heidelberg, 67-86.
- Saß H, Wittchen H-U, Zaudig M, Houben I (2003) Diagnostische Kriterien des Diagnostischen und Statistischen Manuals Psychischer Störungen DSM-IV. Hogrefe, Göttingen.
- Schäfer M, Yassari AH, Hansen BKA, Jelinek L (2024) Zwangsstörungen: Expositionsbehandlung und Besonderheiten der Kompaktbehandlung. *PSYCH up2date*. 18 (1) 17-35.
- Schmidt J, Lamprecht F, Wittmann WW (1989) Zufriedenheit mit der stationären Versorgung: Entwicklung eines Fragebogens und erste Validitätsuntersuchungen. *Psychotherapie Psychosomatik Medizinische Psychologie*. 39, 248-255.
- Schmidt J, Nübling R (2002) ZUF-8: Fragebogen zur Messung der Patientenzufriedenheit. In: *Diagnostische Verfahren in der Psychotherapie*. Brähler E, Schumacher J, Strauß B (Hrsg.) Hogrefe, Göttingen, 392-396.

- Schwarzer R, Jerusalem M (1999) Skalen zur Erfassung von Lehrer- und Schülermerkmalen – Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Freie Universität Berlin, Berlin.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*. 59 (20).
- Silver E, Isometsä E, Baryshnikov I, Heino H, Stenberg J-H, Saarni SE (2023) Bergen 4-Day Treatment (B4DT) for Obsessive-Compulsive Disorder – An observational pilot study of a treatment protocol in Finland. *Psychiatria Fennica*. 54, 198-217.
- Skapinakis P, Caldwell DM, Hollingworth W, Bryden P, Fineberg NA, Salkovskis P, Welton NJ, Baxter H, Kessler D, Churchill R, Lewis G (2016) Pharmacological and psychotherapeutic interventions for management of obsessive-compulsive disorder in adults: a systematic review and network meta-analysis. *Lancet Psychiatry*. 3 (8) 730-739.
- Skoog G, Skoog I (1999) A 40-Year Follow-up of Patients With Obsessive-compulsive Disorder. *Archives Of General Psychiatry*. 56, 121-127.
- Statistisches Bundesamt (Destatis) (2025) Basistabelle Bevölkerungsdichte nach Ländern. Destatis, Wiesbaden [Online im Internet] URL: https://www.destatis.de/DE/Themen/Laender-Regionen/Internationales/Thema/Tabellen/Basistabelle_Bevoelkerungsdichte.html [Stand: 16. August.2025, 14:12 Uhr].
- Stein DJ, Kogan CS, Atmaca M, Fineberg NA, Fontenelle LF, Grant JE, Matsunaga H, Reddy YCJ, Simpson HB, Thomsen PH, van den Heuvel OA, Veale D, Woods DW, Reed GM (2016) The classification of Obsessive-Compulsive and Related Disorders in the ICD-11. *Journal of Affective Disorders*. 190, 663-674.
- Stein DJ, Costa DLC, Lochner C, Miguel EC, Reddy YCJ, Shavitt RG, van den Heuvel OA, Simpson HB (2020) Obsessive–compulsive disorder. *Nature Reviews Disease Primers*. 5 (1) 52.
- Stengler K, Olbrich S, Heider D, Dietrich S, Riedel-Heller S, Jahn I (2013) Mental health treatment seeking among patients with OCD: impact of age of onset. *Social Psychiatry and Psychiatric Epidemiology*. 48, 813-819.
- Stewart SE, Geller DA, Jenike M, Pauls D, Shaw D, Mullin B, Faraone SV (2004) Long-term outcome of pediatric obsessive-compulsive disorder: a meta-analysis and qualitative review of the literature. *Acta Psychiatrica Scandinavica*. 110, 4-13.
- Tandt HLN, Van Parys H, Leyman L, Purdon C, Lemmens GMD (2022) How are OCD patients and their families coping with the COVID-19 pandemic? A qualitative study. *Current Psychology*. 41, 505-515.
- Thorsen AL, de Wit SJ, Hagland P, Ousdal OT, Hansen B, Hagen K, Kvale G, van den Heuvel OA (2020a) Stable inhibition-related inferior frontal hypoactivation and fronto-limbic hyperconnectivity in obsessive-compulsive disorder after concentrated exposure therapy. *Neuroimage: Clinical*. 28, 102460.
- Thorsen AL, Vriend C, de Wit SJ, Ousdal OT, Hagen K, Hansen B, Kvale G, van den Heuvel OA (2020b) Effects of Bergen 4-Day Treatment on Resting - State Graph Features

- in Obsessive-Compulsive Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. 6, 973-982.
- Tjelle K, Opstad HB, Solem S, Launes G, Hansen B, Kvale G, Hagen K (2021) Treatment Adherence as Predictor of Outcome in Concentrated Exposure Treatment for Obsessive-Compulsive Disorder. *Frontiers in Psychology*. 12.
- Tjelle K, Opstad HB, Solem S, Kvale G, Wheaton MG, Björgvinsson T, Hansen B, Hagen K (2024) Patient adherence as a predictor of acute and long-term outcomes in concentrated exposure treatment for difficult-to-treat obsessive-compulsive disorder. *BMC Psychiatry*. 24 (327).
- Voderholzer U, Schlegl S, Diedrich A, Külz AK, Thiel N, Hertenstein E, Schwartz C, Rufer M, Herbst N, Nissen C, Hillebrand T, Osen B, Stengler K, Jelinek L, Moritz S (2015) Versorgung Zwangserkrankter mit kognitiver Verhaltenstherapie als Behandlungsmethode erster Wahl. *Verhaltenstherapie*. 25 (3) 183-190.
- Voderholzer U (2022) Zwangsstörungen und Zwangsspektrumsstörungen. *Nervenarzt*. 93 (7) 659-660.
- Voderholzer U, Favreau M, Rubart A, Staniloiu A, Wahl-Kordon A, Zurowski B, Kathman N (2022) Therapie der Zwangsstörungen: Empfehlungen der revidierten S3-Leitlinie Zwangsstörungen. *Nervenarzt*. 93, 678-687.
- Woody SR, Steketee G, Chambless DL (1995) Reliability and Validity of the Yale-Brown Obsessive-Compulsive Scale. *Behaviour Research and Therapy*. 33 (5) 597-605.

9. Abkürzungsverzeichnis

Abkürzung	Definition
ANOVA	„analysis of variance“ / Varianzanalyse
BEAQ	Brief Experimental Avoidance Questionnaire
B4DT	Bergen 4-Day Treatment
CSQ-8	Client Satisfaction Questionnaire
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders Forth Edition
DSM-5	Diagnostic and Statistical Manual of Mental Disorders Fifth Edition
ERM	Exposition und Reaktionsmanagement
GAF	Global Assessment of Functioning
ICD-10 / ICD-11	International Classification of Diseases 10th Revision / 11th Revision
KVT	Kognitive Verhaltenstherapie
LET	„LEan into the anxiety“-Technik
M.I.N.I.	Mini-International Neuropsychiatric Interview
OCD	Obsessive-Compulsive Disorder
OCI-R	Obsessive-Compulsive Inventory – Revised
PHQ-9	Patient Health Questionnaire
RSES	Rosenberg-Self-Esteem-Scale
SSRIs	Selektive Serotonin-Wiederaufnahme-Inhibitoren
SWE	Skala zur Allgemeinen Selbstwirksamkeitserwartung
UKE	Universitätsklinikum Hamburg Eppendorf
Y-BOCS	Yale-Brown Obsessive-Compulsive Scale
ZUF-8	Fragebogen zur Messung der Patientenzufriedenheit

10. Abbildungsverzeichnis

Abbildung 1

Zeitpunkte der Datenerhebung20

Abbildung 2

Flussdiagramm der durchgeführten Pilotstudie35

Abbildung 3

Individuelle Veränderungen, Boxplots und Dichteverteilungen
der Y-BOCS-Gesamtwerte zu den drei Messzeitpunkten37

Abbildung 4

Klinisch signifikante Veränderungen –
Therapieansprechen (Response) und Remission.....41

11. Tabellenverzeichnis

Tabelle 1

Literaturübersicht zur Effektivität
des Bergen 4-Day Treatments..... Anhang 1

Tabelle 2

Erhebungsmodalitäten und Erhebungszeitpunkte der Instrumente.....24

Tabelle 3

Soziodemografische Daten, Komorbiditäten
und Medikation der Stichprobe34

Tabelle 4

Ergebnisse: Primäre und sekundäre Outcomes.....39

12. Vorveröffentlichung und Erklärung des Eigenanteils

Die vorliegende Dissertation entstand im Rahmen eines Projekts, das die Durchführung und wissenschaftliche Begleitung einer Pilotstudie zum Bergen 4-Day Treatment in Deutschland umfasste. Diese Pilotstudie wurde inzwischen veröffentlicht:

Jelinek L, Serve A, Pampuch S, Scheunemann J, Schultz J, Miegel F, Hansen B, Hagen K, Bohnsack F, Gallinat J, Yassari AH (2024b) Exploring the Peaks and Potholes: Understanding positive and negative effects of concentrated exposure treatment for obsessive-compulsive disorder. Journal of Obsessive-Compulsive and Related Disorders. 43, 100913.

Die entsprechende Publikation, an der ich als Mitautorin beteiligt bin, ist dieser Arbeit im Anhang beigelegt. Die Dissertation stellt eine vertiefte Auswertung und Einordnung der Ergebnisse dieser Pilotstudie dar.

Ich war maßgeblich an der Organisation und Durchführung der Studie beteiligt. Dazu gehörten die Koordination und Durchführung der Datenerhebung (inklusive Online-Erhebung und Klinischer Interviews zum Follow-Up), die Kommunikation mit den teilnehmenden Proband:innen sowie die Qualitätssicherung der Datensätze. Die statistischen Analysen der erhobenen Daten habe ich eigenständig mit SPSS durchgeführt.

Hamburg, 13.07.2026
Ort, Datum


Unterschrift

13. Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe, insbesondere ohne entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten, verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Das gilt insbesondere auch für alle Informationen aus Internetquellen.

Soweit beim Verfassen der Dissertation KI-basierte Tools („Chatbots“) verwendet wurden, versichere ich ausdrücklich, den daraus generierten Anteil deutlich kenntlich gemacht zu haben. Die „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“ aus September 2023 wurde dabei beachtet.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich damit einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Hamburg, 13.07.2026
Ort, Datum


Unterschrift

14. Danksagung

An erster Stelle möchte ich meiner Betreuerin und Doktormutter, Prof. Dr. phil. Lena Jelinek, für die wirklich gute, verlässliche Betreuung und flexible Unterstützung danken. Ihre fachliche Expertise und die wertvollen Ratschläge, sowie ihre Geduld haben maßgeblich zum Gelingen dieser Arbeit beigetragen.

Mein Dank gilt ebenfalls der gesamten Arbeitsgruppe und allen Beteiligten am Pilotprojekt des Bergen-4-Day-Treatments in Hamburg sowie allen Mitarbeiter:innen der Station PA7 für die stets nette Zusammenarbeit und die angenehme Arbeitsatmosphäre.

Ein besonderer Dank geht an meine Familie, an Henry und an Lea, die mich auf meinem Weg bestärkt, motiviert und begleitet haben.

Nicht zuletzt danke ich allen Menschen, die an B4DT teilgenommen haben – ohne ihre Bereitschaft wäre diese Forschung nicht durchführbar gewesen.

15. Anhang

15.1. Tabelle 1

Tabelle 1

Literaturübersicht zur Effektivität des Bergen 4-Day Treatments

Autor (Jahr) Standort der Behandlung	Stichproben- größe (N)	Kurzbeschreibung	Instrumente	Zeitpunkte der Datenerhebung	Kriterien Responder/ Remission	Mittlere Y-BOCS- Werte	Ergebnisse/ Wirksamkeit
Havnen et al. 2013 Bergen, Norwegen	7	Erste Pilotstudie: konzentrierte ERP im Gruppenformat	BDI CEQ Y-BOCS M.I.N.I. Overall satisfaction	T0: prä Intervention T1: post Intervention T2: 1 Monat post Intervention T3: 3 Monate post Intervention T4: 6 Monate post Intervention	/	T0: 23.5 T1: 5.7 T2: 6.5 T3: 5.8 T4: 6.3	Reduktion der Zwangssymptome durch B4DT mit hoher Effektstärke bei hoher Akzeptanz der Behandlung
Havnen et al. 2014 Bergen, Norwegen	35	Erweiterung der Pilotstudie von 2013: neue, größere Proband:innen- Gruppe, Evaluation der Akzeptanz der konzentrierten ERP- Behandlung	BDI CSQ-8 M.I.N.I. Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention T3: 6 Monate post Intervention	e)	T0: 26.1 T1: 9.0 T2: 10.6 T3: 10.3	T1 77.1% Recovered 11.4 % Improved 11.4 % Keine Veränderung T2 74% Recovered 14% Improved 11% Keine Veränderung T3 77.1 % Recovered 14.3 % Improved 8.6 % Keine Veränderung

						Reduktion von Zwangssymptomen mit hoher Effektstärke ($d = 2.60$) und depressiven Symptomen mit hoher Effektstärke ($d = 0.89$) nach der Behandlung, und auch nach 6 Monaten ($d = 2.37$ für Zwangssymptome, $d = 1.02$ für depressive Symptome); B4DT als praktikables Format für ERP; hohe Zufriedenheit der Teilnehmer:innen mit der Behandlung und niedrige Drop-out-/Abbruchrate
Havnen et al. 2017 Bergen, Norwegen	42	Unkontrollierte Studie mit dem Ziel der Replikation der Ergebnisse Havnen et al. (2014) in der gleichen Klinik mit einer neuen Stichprobe und verschiedenen Therapeut:innen unter der Aufsicht der Behandlungsentwickler	CSQ-8 WSAS Y-BOCS	T0: prä Intervention T1: post Intervention T2: 6 Monate post Intervention	e)	T0: 25.7 T1 T1: 10.8 73.8% Recovered T2: 12.2 9.5% Improved 16.7% Keine Veränderung T2 59.5% Recovered 16.7% Improved 23.8% Keine Veränderung Erfolgreiche Replikation der Studienergebnisse von 2014 mit neuer Stichprobe und einem neuen Therapeut:innen-Team; Starke Reduktion der Zwangssymptome mit anhaltendem Erfolg nach 6 Monaten (keine Nennung der Effektstärke); hohe Gesamtzufriedenheit der Teilnehmer:innen
Hansen et al. 2018 Bergen, Norwegen	65	Untersuchung der längerfristigen Effekte von B4DT mittels 12-Monats-Follow-Up; Vergleich der Ergebnisse mit der Meta-Analyse von Öst et al. (2015)	CSQ-8 GAD-7 M.I.N.I. PHQ-9 Y-BOCS	T0: prä Intervention T1: innerhalb einer Woche post Intervention T2: 3 Monate post Intervention T3: 12 Monate post Intervention	b)	T0: 25.8 T1 T1: 10.2 16.9 % Responders T2: 10.5 76.9 % Remission T3: 10.6 6.2 % Keine Veränderung T2 9.2 % Responders 72.3 % Remission 18.5 % Keine Veränderung T3 15.4 % Responders 67.7% Remission 16.9 % Keine Veränderung
					d)	T1 93.8 % Responders 67.7 % Remission T3

						86.2 % Responders 64.6 % Remission
						Reduktion der Zwangssymptome mit hoher Effektstärke ($d = 3.35$ von prä- zu post-Intervention); Effekte der Behandlung bleiben nach 12 Monaten stabil ($d = 3.27$); Anstieg der Zahl asymptomatischer Teilnehmer:innen ($Y\text{-BOCS} \leq 7$); keine Drop-Outs; Kosteneffizienz und gute Machbarkeit im gewöhnlichen Ambulanzsetting; signifikante Reduktion depressiver Symptome und Ängste mit kleiner Effektstärke; hohe Zufriedenheit der Teilnehmer:innen ($CSQ-8$, $M = 29.91$).
						Im Vergleich zur Standard-ERP (Öst et al. 2015) zeigten sich durch das B4DT eine signifikant höhere Reduktion der Zwangssymptome ($d = 3.35$ vs. 2.38 zum Zeitpunkt T1) sowie höherer Anteil an klinisch relevanter Verbesserung/ Remission
Kvale et al. 2018 Oslo, Norwegen	90	Systematische Replikation: größere Anzahl an Proband:innen (Behandlung von 90 Patient:innen innerhalb von 8 Tagen); Rekrutierung der Therapeut:innen aus ganz Norwegen	GAD-7 M.I.N.I. PHQ-9 Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention	b)	T0: 26.2 T1: 10.5 T2: 10.7 T1 18.9 % Responders 72.2 % Remission 8.9 % Keine Veränderung T2 16.7 % Responders 67.8 % Remission 15.6 % Keine Veränderung
						Erfolgreiche Behandlung der 90 Patient:innen innerhalb von 8 Tagen; Replikation der Ergebnisse aus bisherigen Studien mit kleineren Stichproben; Hohe Effektstärken des B4DT sowohl zum Zeitpunkt T1 als auch T2 (beides $d = 4.6$); Integration von B4DT in die Standardversorgung in Oslo (OUH)
Nordahl et al. 2018 Bergen, Norwegen	36	Untersuchung Prävalenz und des Schweregrads selbstberichteter Schlafstörungen bei Menschen mit Zwangsstörungen, Erfassung der Veränderung von Zwangsstörungssympto	BDI BIS M.I.N.I. PHQ-9 Y-BOCS	T0: prä Intervention T1: post Intervention T2: 6 Monate post Intervention	/	T0: 25.1 T1: 9.8 T2: 10.9 Mehrheit der Stichprobe mit Zwangsstörungen leidet unter Schlafstörungen; signifikante Reduktion der Zwangssymptome mit hoher Effektstärke durch B4DT ($d = 3.52$ von prä- zu post-Intervention) sowie signifikante Reduktion von Schlafstörungen durch B4DT

		men nach B4DT und eines möglichen Zusammenhangs zwischen dem Behandlungsergebnis und Schlafstörungen.				
Davíðsdóttir et al. 2019 Reykjavík, Island	19	Implementation von B4DT in Island (erstmalig außerhalb Norwegens)	CSQ-8 GAD-7 M.I.N.I. PHQ-9 Y-BOCS	T0: prä Intervention T1: 1 Woche post Intervention T2: 3 Monate post Intervention	a)	T0: 28.8 T1 T1: 10.0 94.7 % Responders T2: 11.1 73.7 % Remission T2 78.9 % Responders 63.2 % Remission Erfolgreiche Implementation des Bergen 4-Day Treatments außerhalb von Norwegen; Große Effektstärke der Behandlung auf die Zwangssymptome sowohl post-Intervention ($d = 4.26$), als auch 3 Monate nach der Intervention ($d = 4.01$); Signifikante Reduktion der depressiven Symptome mit hoher Effektstärke ($d = 0.94$); vergleichbare Ergebnisse mit den vorangegangenen norwegischen Studien; hohe Zufriedenheit der Teilnehmer:innen und der Therapeut:innen mit der Behandlung
Hansen et al. 2019 Bergen, Norwegen	77	Vierjährige Nachbeobachtung der Studienpopulationen aus Havnen et al. (2014) und Havnen et al (2017), s.o.	BDI Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention T3: 6 Monate post Intervention T4: 4 Jahre post Intervention	b)	T0: 25.9 T1 T1: 10.0 22.1 % Responders T2: 10.7 72.7% Remission T3: 11.3 5.2 % Keine Veränderung T4: 9.9 T4 11.7 % Responders 68.8 % Remission 19.5 % Keine Veränderung Reduktion der Zwangssymptome mit hoher Effektstärke ($d = 3.65$ von prä- zu post-Intervention); Wirkung des Bergen 4-Day Treatments auch nach 4 Jahren erhalten; Erhöhte Berufstätigkeit/Arbeitsfähigkeit der Teilnehmer:innen nach 4 Jahren
Holm et al. 2019a Bergen, Norwegen	104	Untersuchung des Zusammenhangs zwischen Hochsensibilität als	GAD-7 HSPS M.I.N.I. OCI-R	T0: prä Intervention T1: post Intervention	/	T0: 26.0 Kein signifikanter Zusammenhang zwischen Sensibilität, gemessen anhand der Highly Sensitive Person Scale (HSPS) und dem Therapieansprechen auf B4DT; Patient:innen mit hoher sensorischer Sensibilität profitierten ebenso stark von der

		Persönlichkeitsmerkmal von Patient:innen und Zwangssymptomen bzw. Therapieansprechen	PHQ-9 Y-BOCS	T2: 3 Monate post Intervention T3: 6 Monate post Intervention		Behandlung wie die anderen Patient:innen; Ausgangswerte der Y-BOCS als einziger signifikanter Prädiktor für das Behandlungsergebnis bei der Nachuntersuchung
Holm et al. 2019b Bergen, Norwegen	89	Untersuchung von Unterschieden der Resilienz zwischen Menschen mit Zwangsstörungen und zwei Referenzgruppen. Untersuchung der Zusammenhänge zwischen Resilienz und Symptomschwere sowie zwischen Resilienz und dem Ansprechen auf das Bergen-4-Day-Treatment.	DRS-15-R M.I.N.I. Y-BOCS	T0: prä Intervention T1: post Intervention T2: mindestens 3 Monate post Intervention	a)	T0: 26.0 T1 T1: 10.1 89.9 % Responders T2: 11.7 73 % Remission T2 78.7 % Responders 59.6 % Remission Reduktion der Zwangssymptome mit hoher Effektstärke von prä- zu post-Intervention ($d = 2.46$) sowie von prä-Intervention zum Follow-Up ($d = 1.80$); Geringere Resilienz-Werte, gemessen anhand der Dispositional Resilience Scale (DSR-15-R), in der Gruppe der Menschen mit Zwangsstörungen, Signifikante Verbesserung der Resilienz von T0 zu T1 mit mittlerer Effektstärke ($d = 0.65$); Resilienz hat keinen signifikanten Effekt auf das Therapieansprechen auf B4DT; schwache Korrelation zwischen der Resilienz und der Remission zum Zeitpunkt T2
Launes et al. 2019a Kristiansand Norwegen	48 16 B4DT 16 Selbsthilfe (SH) 16 Warteliste (WL)	Erste randomisierte kontrollierte Studie für die Untersuchung des B4DT; drei randomisierte Gruppen: B4DT, SH, WL	CSQ-8 CSSRI DOCS-SF GAD-7 M.I.N.I. OCI-R PHQ-9 SCID-5-RV WSAS Y-BOCS	T0: prä Intervention T1: in der ersten Woche post Intervention T2: 3 Monate post Intervention T3: 6 Monate post Intervention	b)	<i>Interventionsgruppe</i> <i>Interventionsgruppe</i> T1 T0: 26.8 31.3 % Responders T1: 10.9 62.5 % Remission T2: 8.6 6.3 % Keine Veränderung T3: 9.2 T3 6.3 % Responders 81.3 % Remission 12.5 % Keine Veränderung Signifikant bessere Ergebnisse durch B4DT hinsichtlich der Schwere der Zwangsstörung als SH und WL ($d = 3.75$ vs. $d = 0.77$ (SH) und $d = -0.11$ (WL) von prä- zu post-Intervention); signifikant höhere Effektstärken der Behandlung auf depressive Symptome ($d = 0.95$ vs. $d = 0.65$ (SH) vs. $d = 0.01$ (WL)), generalisierte Angst sowie höhere soziale und arbeitsbezogene Funktionsfähigkeit der Interventionsgruppe im Vergleich zu den beiden Kontrollgruppen;

						Bestätigung der Ergebnisse aus vorangehenden unkontrollierten Effektivitätsstudien	
Launes et al. 2019b Kristiansand Norwegen	35	Erste Replikationsstudie der Ergebnisse des B4DT in öffentlicher psychiatrischer Fachversorgung außerhalb der Klinik der Initiatoren	CSQ-8 DOCS-SF GAD-7 M.I.N.I. OCI-R PHQ-9 SCID-5-RV Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention	b)	T0: 26.7 T1: 10.9 T2: 10.0	T1 20.0 % Responders 74.3 % Remission 5.7 % Keine Veränderung T2 11.4 % Responders 68.6 % Remission 20.0 % Keine Veränderung
						Reduktion der Zwangssymptome mit hoher Effektstärke (prä-Intervention zu post-Intervention $d = 4.50$); Erfolgreiche Umsetzung des B4DT in einer neuen Klinik; Ergebnisse der Initiatoren auch an neuem Standort zu reproduzieren; hohe Zufriedenheit der Teilnehmer:innen mit der Behandlung (CSQ-8 $M = 29.5$)	
Launes et al. 2020 Kristiansand Norwegen	42 16 B4DT (s. Launes et al. 2019) 26 SH/ WL, die im Nachgang an B4DT inklusive T2 teilgenommen haben	B4DT mit Proband:innen aus SH- und WL-Gruppen (Launes et al. 2019), Analyse der Daten inklusive der ersten Interventionsgruppe	CSQ-8 DOCS-SF GAD-7 M.I.N.I. OCI-R PHQ-9 Y-BOCS	T0: prä Intervention T1: in der ersten Woche post Intervention T2: 3 Monate post Intervention T3: 6 Monate post Intervention	b)	T0: 27.1 T1: 11.8 T2: 10.5	T1 31 % Responders 59.5% Remission 9.5 % Keine Veränderung T2 19 % Responders 71.4% Remission 9.5 % Keine Veränderung
						Patient:innen , die zuvor in SH- und WL-Gruppe waren profitieren genauso stark von B4DT wie die ursprüngliche Interventionsgruppe; hohe Effektstärken von B4DT auf die Zwangssymptome ($d = 4.60$ von T0 zu T1 sowie $d = 4.98$ von T0 zu T2)	
Kvale et al. 2020 Bergen, Norwegen	163 67 B4DT + 250mg D-Cycloserin (DCS) 65 B4DT + 100mg 31 B4DT + Placebo	Randomisierte, placebo-kontrollierte Studie mit Proband:innen mit schwer zu behandelnden Zwangsstörungen zur Untersuchung, ob die Gabe von DCS zu	CSQ-8 DOCS-SF GAD-7 M.I.N.I. OCI-R PHQ-9 SCID-5-RV WEMWBS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention T3: 12 Monate post Intervention	b)	Gesamte Stichprobe T0: 27.1 T1: 12.7 T2: 14.3 T3: 14.3	Gesamte Stichprobe T1 27.3 % Responders 56.5 % Remission 16.2 % Keine Veränderung T3 15.1 % Responders 47.9 % Remission

		Wirkungsverstärkung von B4DT führt	Y-BOCS			37.0 % Keine Veränderung
						In allen Gruppen Reduktion der Zwangssymptome mit großer Effektstärke ($d = 3.01-3.49$); kein signifikanter Zusatznutzen von DCS gegenüber Placebo (post-Intervention $\chi^2 = 3.21$; $p = .07$); Reduktion von depressiven Symptomen und Angst sowie Steigerung des Wohlbefindens in allen Gruppen mit kleinen Effektgrößen
Thorsen et al. 2020a Bergen, Norwegen	31	Untersuchung von Veränderungen in der Leistung, der aufgabenbezogenen Hirnaktivierung und der Konnektivität während des und nach dem Bergen-4-Day-Treatments mittels MRT	BRIEF GAD-7 PHQ-9 SCID-I Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention	a)	T0: 26.8 T1 T1: 10.3 25 % Responders T2: 10.3 67 % Remission 8 % Keine Veränderung T2 8 % Responders 19 % Remission 13 % Keine Veränderung
						Menschen mit Zwangsstörungen zeigen weniger Aktivierung des Gyrus frontalis inferior bei Reaktionsmanagement; keine Veränderung der Aktivierung des Gyrus frontalis inferior und der frontolimbischen Konnektivität während des Reaktionsmanagements bei Menschen mit Zwangsstörungen nach konzentrierter Expositionstherapie; Keine Evidenz für einen Effekt der Hirnaktivierung und Konnektivität auf das Therapieansprechen
Thorsen et al. 2020b Bergen, Norwegen	34	Untersuchung der Auswirkungen von B4DT auf statische und dynamische funktionelle neuronale Verbindungen bei Patient:innen mit Zwangsstörungen mit Hilfe funktioneller Magnetresonanztomographie	GAD-7 OCI-R PHQ-9 SCID-I Y-BOCS	T0: prä Intervention T1: post Intervention	a)	T0: 26.7 T1 T1: 11.0 25 % Responders 61 % Remission 14% Keine Veränderung
						B4DT führte bei den Proband:innen mit Zwangsstörungen zu einer verringerten Konnektivität zwischen den frontoparietalen und limbischen Subnetzwerken und zu einer verringerten Flexibilität der Konnektivität des anterioren cingulären Cortex, was auf eine unabhängigere und stabilere Netzwerktopologie hindeutet. Dies kann als Hinweis auf einen geringeren limbischen Einfluss auf die Subnetzwerke der kognitiven Kontrolle nach der Behandlung interpretiert werden
Brecke et al.	32	Untersuchung der	GAD-7	T0: prä	a)	T0: 27.1 T2

2021 Bergen, Norwegen		Integrität der weißen Substanz bei Menschen mit Zwangsstörungen und Kontrollpersonen sowie der Veränderungen nach B4DT mittels diffusionsgewichteter Magnetresonanztomographie (Diffusion-Tensor-Imaging, DT-MRT)	PHQ-9 SCID-I Y-BOCS	Intervention T1: post Intervention T2: 3 Monate post Intervention		T1: 11.4 T2: 10.6	11.5 % Responders 76.9 % Remission 11.5 % Keine Veränderung Reduktion der Zwangssymptome durch B4DT mit hoher Effektstärke ($d = 3.92$); keine signifikanten (Langzeit-) Veränderungen der Mikrostruktur der weißen Substanz im DT-MRT
Hagen et al. 2021 Kristiansand Norwegen	42 Proband:innen aus Launes et al 2020 (s.o.)	Evaluation des Effekts von B4DT auf Schlafstörungen/ Schlaflosigkeit bei Patient:innen mit Zwangsstörungen sowie den Zusammenhang zwischen depressiven Symptomen und Schlafstörungen bei den B4DT-Proband:innen	BIS PHQ-9 Y-BOCS	Äquivalent zu Launes et al. 2020 (s.o.)	b)	Äquivalent Responder-/ Remissionsraten s. Launes et al. 2020 zu Launes et al. 2020 B4DT wirksam, unabhängig von komorbider Schlaflosigkeit/ Schlafstörungen; Schlafstörungen nach der Behandlung mit mittlerer Effektstärke reduziert ($d = 0.53$); Die Verbesserung des Schlafs (BIS) korrelierte moderat bis stark mit der Veränderung depressiver Symptome (PHQ-9; $r = .61$), während die Zusammenhänge zwischen Veränderungen im Schlaf und Zwangssymptomen (Y-BOCS; $r = .28$) sowie zwischen PHQ-9 und Y-BOCS ($r = .35$) schwächer und nicht signifikant waren	
Hagland et al. 2021 Bergen, Norwegen	29	Untersuchung von longitudinalen intra- und interpersonellen Zusammenhängen zwischen Zwangssymptomen, Leistungsfähigkeit, der Aktivierung des rechten Gyrus frontalis inferior und der Konnektivität zwischen der rechten Amygdala	GAD-7 PHQ-9 Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention	a)	T0: 26.8 T1: 10.5 T2: 10.5	T1 24.1 % Responders 62.1 % Remission 13.8 % Keine Veränderung Signifikante positive Korrelation zwischen der Symptomschwere von Zwangs-, Angststörungen und Depressivität und der Reaktionszeit; Keine signifikante Korrelation zwischen Symptomschwere und Aktivierung des rechten Gyrus frontalis inferior sowie der Konnektivität

		und dem rechten supplementär motorischen Kortex vor und nach B4DT mittels Magnetresonanztomog raphie				
Tjelle et al. 2021 Kristiansand Norwegen	42 Proband:innen aus Launes et al 2020 (s.o.)	Untersuchung der Behandlungsdhärenz als Prädiktor für den individuellen Therapieerfolg bei Konzentrierter Expositionstherapie (B4DT)	CSQ-8 GAD-7 PEAS PHQ-9 WSAS WEMWBS Y-BOCS	T0: prä Intervention T1: in der ersten Woche post Intervention T2: 3 Monate post Intervention T3: 6 Monate post Intervention	b)	Äquivalent Responder-/ Remissionsraten s. Launes et al. 2020 zu Launes et al. 2020 B4DT führte zu einer starken Symptomreduktion ($d = 3.56$) sowie (s.o.) deutlichen Verbesserungen in Angst, Depression und Funktionsniveau; hohe Therapieadhärenz, gemessen anhand der Patient EX/RP Adherence Scale (PEAS), war signifikant mit besseren Behandlungsergebnissen assoziiert (z. B. $r = -0.59$ mit Y-BOCS post- Intervention) und sagte den Behandlungserfolg über die Ausgangswerte hinaus voraus; Patienten mit niedriger Adhärenz zeigten schlechtere Outcomes und geringeres Wohlbefinden beim Follow-up
De Jooode et al. 2022 Bergen, Norwegen	33	Untersuchung der unmittelbaren und anhaltenden Auswirkungen von B4DT auf die Konzentrationen der Neurometabolite im anterioren cingulären Cortex mittels Magnetresonanztomog oskopie (MRS)	GAD-7 OCI-R PHQ-9 SCID-I Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention	b)	T0: 26.0 T1 T1: 10.7 27 % Responders T2: 10.0 60 % Remission 13 % Keine Veränderung T2 13 % Responders 70 % Remission 17 % Keine Veränderung Es konnten keine Unterschiede in der Konzentration der Neurometabolite Glutamat, Glutamin und N-Acetylaspartat zwischen Menschen mit Zwangsstörungen und der Kontrollgruppe vor der Intervention und über die Zeit festgestellt werden; Unterschiede in Neurometabolit-Konzentrationen schienen abhängig von Komorbiditäten und dem Schweregrad der Zwangsstörung
Silver et al. 2023 Helsinki, Finnland	46	Implementation von B4DT in Finnland Untersuchung der Durchführbarkeit und	BIS CEQ CSQ-8 DOCS-SF	T0: prä Intervention T1: 10 Tage post Intervention	c)	T0: 27.9 T1 T1: 17.1 56 % Responders T2: 17.0 28 % Remission T2

		Effektivität Behandlung finnischen Versorgungskontext	der im	GAD-7 PHQ-9 WEMWBS WSAS Y-BOCS-SR	T2: 3 Monate post Intervention		58 % Responders 30 % Remission	Erfolgreiche Implementation des Bergen 4-Day Treatments in Finnland; hohe Effektstärke der Behandlung ($d = 1.93$ von prä- zu post-Intervention) sowie signifikante Reduktion von Angst und depressiven Symptomen, von Schlafproblemen und Problemen in der psychosozialen Funktionsfähigkeit. Geringere Effektstärken im Vergleich zu vorangegangenen norwegischen Studien; hohe Zufriedenheit der Teilnehmer:innen und der Therapeut:innen mit der Behandlung. Analyse und Diskussion möglicher individuelle Einflussfaktoren auf die individuelle Effektivität der Behandlung
Björgvinsson et al. 2024 Houston, Texas, U.S.A.	48	Pilotstudie Evaluation Durchführbarkeit und Wirksamkeit von B4DT in den Vereinigten Staaten	zur der und	CSQ-8 DOCS-SF GAD-7 M.I.N.I. PHQ-9 WSAS Y-BOCS	T0: prä Intervention T1: 10 Tage post Intervention T2: 3 Monate post Intervention T3: 6 Monate post Intervention	f)	T0: 27.04 T1 T1: 11.67 91.7 % Signifikante Veränderung T2: 12.73 81.3 % Responders T3: 13.67 T2 85.1 % Signifikante Veränderung 75.0 % Responders T3 81.6 % Signifikante Veränderung 68.6 % Responders Bestätigung der in den nordischen Ländern beobachteten hohen Effektstärke von B4DT in der Reduktion des Y-BOCS-Werts auch im amerikanischen Gesundheitssystem (von prä- zu post-Intervention $d = 2.90$); zudem deutliche Reduktion von Angst (GAD7; von prä- zu post-Intervention $d = 1.38$) und depressiven Symptomen (PHQ9; von prä- zu post-Intervention $d = 0.92$). Steigerung der sozialen/beruflichen/psychosozialen Funktionsfähigkeit. Hohe Zufriedenheit der Patient:innen und Behandler:innen. Anhalten der Verbesserung der Symptome auch bis sechs Monate nach der Behandlung	
Holth Landrø et al. 2024 Bergen, Norwegen	163 Proband:innen aus Kvale et al 2020 (s.o.)	Sekundäranalyse randomisiert- kontrollierten Studie von Kvale et al. 2020 (s.o.): Untersuchung der Auswirkungen des B4DT auf	der Studie	BIS M.I.N.I. PHQ-9 Y-BOCS	Äquivalent zu Kvale et al. 2020 (s.o.)	b)	Äquivalent Responder-/ Remissionsraten s. Kvale et al. 2020 zu Kvale et al. 2020 (s.o.) Geringer Effekt von B4DT auf die Symptome der Schlaflosigkeit ($d = 0.37$), anhaltend bis T3; keine signifikanten Unterschiede zwischen den Gruppen, die DCS oder Placebo erhielten; Schlaflosigkeit hatte keinen signifikanten Einfluss auf das Therapieansprechen	

Schlafstörungen

Tjelle et al. 2024 Bergen, Norwegen	163 Proband:innen aus Kvale et al 2020 (s.o.)	Sekundäranalyse der randomisiert- kontrollierten Studie von Kvale et al. 2020 (s.o.): Untersuchung des Zusammenhangs zwischen Therapieadhärenz während der Expositionsaufgaben und Ansprechen auf B4DT	GAD-7 PEAS-P PHQ-9 WSAS Y-BOCS	Äquivalent zu Kvale et al. 2020 (s.o.)	b)	Äquivalent Responder-/ Remissionsraten s. Kvale et al. 2020 zu Kvale et al. 2020 Die Therapieadhärenz, gemessen mittels PEAS-P, war kurz- und (s.o.) langfristig mit dem besseren Outcome von geringeren Zwangssymptomen und besserer Funktionsfähigkeit verbunden, unabhängig von der Symptomschwere zu Beginn der Behandlung
Wheaton et al. 2024 Bergen, Norwegen	163 Proband:innen aus Kvale et al 2020 (s.o.)	Sekundäranalyse der randomisiert- kontrollierten Studie von Kvale et al. 2020 (s.o.): Untersuchung von Vermeidungsverhalten als möglichen Einflussfaktor auf das Ansprechen auf B4DT	WSAS Y-BOCS	Äquivalent zu Kvale et al. 2020 (s.o.)	b)	T0: 27.0 Responder-/ Remissionsraten s. Kvale et al. 2020 T1: 12.3 Die Ausgangsausprägung des Vermeidungsverhaltens, gemessen T2: 13.9 mittels WSAS, sagte das Therapieansprechen auf B4DT nicht vorher; T3: 14.4 Patient:innen mit hoher Vermeidung zeigten kein schlechteres Ansprechen nach der Intervention, hatten jedoch höhere Y-BOCS- Werte nach 12 Monaten, was auf ein erhöhtes Rückfallrisiko hindeutet
Berg et al. 2025 Bergen, Norwegen	163 Proband:innen aus Kvale et al 2020 (s.o.)	Sekundäranalyse der randomisiert- kontrollierten Studie von Kvale et al. 2020 (s.o.): Untersuchung von der von Patient:innen wahrgenommenen Glaubwürdigkeit sowie der Therapieerwartungen als Prädiktoren für den	CEQ GAD-7 M.I.N.I. PHQ-9 WSAS Y-BOCS	Äquivalent zu Kvale et al. 2020 (s.o.)	b)	Äquivalent Responder-/ Remissionsraten s. Kvale et al. 2020 zu Kvale et al. 2020 Höhere Behandlungserwartungen und -glaubwürdigkeit, gemessen (s.o.) mittels CEQ, waren mit niedrigeren Y-BOCS-Werten post- Interventionem und nach 12 Monaten sowie mit besserer sozialer und beruflicher Funktionsfähigkeit assoziiert; ein CEQ-Cut-off von 92,5 identifizierte Patienten mit hoher Erwartung, von denen 87 % positiv auf die Behandlung ansprachen

		Behandlungserfolg von B4DT						
Davíðsdóttir et al. 2025 Reykjavík, Island	86	Auswertung von B4DT für eine größere Anzahl von Teilnehmer:innen in Island. Evaluation von Durchführbarkeit und Wirksamkeit.	CSQ-8 GAD-7 M.I.N.I. PHQ-9 WSAS Y-BOCS	T0: prä Intervention T1: post Intervention T2: 3 Monate post Intervention	b)	T0: 30.49	T1	
						T1: 10.62 T2: 10.87	26.9 % Responders 68.0 % Remission 5.1 % Keine Veränderung	
								T2
								23.5 % Responders 67.9 % Remission 8.6 % Keine Veränderung
								Starke Reduktion der Zwangssymptome (Y-BOCS $\Delta \approx 19,7$ Punkte, $p < .001$) durch B4DT in Island mit hoher Effektstärke ($d = 4.83$), unabhängig von Komorbiditäten oder vorherigen Therapieversuchen; vergleichbare Ergebnisse mit norwegischen Studien. B4DT wirkte sowohl bei moderat als auch stark beeinträchtigten Patient:innen und zeigte zusätzlich Verbesserungen bei Angst und Depression; hohe Zufriedenheit der Teilnehmer:innen mit der Behandlung ohne Therapieabbrüche.

Anmerkungen. Ausführung der verschiedenen Kriterien Responder/Remission a) Konsenskriterien von internationalen Experten (Mataix-Cols et al. 2016): Response = Reduktion des individuellen T0-Y-BOCS um $\geq 35\%$, Remission = Y-BOCS-Score von ≤ 12 b) Modifikation der Konsenskriterien von internationalen Experten (Mataix-Cols et al. 2016): Response = Reduktion des individuellen Y-BOCS-Scores zum Zeitpunkt T0 um $\geq 35\%$, Remission = Erfüllung des Response-Kriteriums und Y-BOCS-Score von ≤ 12 , Verschlechterung = Anstieg des individuellen T0-Y-BOCS um $\geq 35\%$, Keine Veränderung = Weder Response noch Verschlechterung c) Modifikation der Konsenskriterien von internationalen Experten (Mataix-Cols et al. 2016): Response = Reduktion des individuellen Y-BOCS-SR-Scores zum Zeitpunkt T0 um $\geq 35\%$, Remission = Y-BOCS-SR-Score von ≤ 13 d) Klinisch signifikante Veränderungen nach Jacobson und Truax (1991)/ Fisher and Wells (2005) „Reliable Change Index“ (RCI): Response = RCI über den 95%-Konfidenzintervallgrenzen ± 1.96 , Remission = Y-BOCS-Score unterhalb eines Cut-Off Wertes (hier ein Wert von 14) berechnet nach Fisher and Wells (2005) e) Modifikation der klinisch signifikanten Veränderungen nach Jacobson und Truax (1991)/ Fisher and Wells (2005) „Reliable Change Index“ (RCI): Recovered/ Response = RCI über den 95%-Konfidenzintervallgrenzen ± 1.96 , Remission = Y-BOCS-Score unterhalb eines Cut-Off Wertes (hier ein Wert von 14) berechnet nach Fisher and Wells (2005), Improved = Erfüllte Kriterien der Recovery, jedoch einen Y-BOCS-Score über dem Cut-Off-Wert für eine Remission, Keine Veränderung = Unabhängig von dem Y-BOCS-Score keine Reduktion des Y-BOCS-Scores entsprechend des RCI f) Klinisch signifikante Veränderungen nach Jacobson und Truax (1991)/ Response nach den Konsenskriterien von internationalen Experten (Mataix-Cols et al. 2016), Klinisch signifikante Veränderung = RCI über den 95%-Konfidenzintervallgrenzen ± 1.96 , Response = Reduktion des individuellen Y-BOCS-SR-Scores zum Zeitpunkt T0 um $\geq 35\%$. **Abkürzungen** BDI = Beck-Depressions-Inventar, BIS = Bergen Insomnia Scale, BRIEF = Behavior Rating Inventory of Executive Function, CEQ = Credibility and Expectancy Questionnaire, CSQ-8 = Client Satisfaction Questionnaire, CSSRI = The Client Socio-Demographic and Service Receipt Inventory, DOCS-SF = Dimensional Obsessive-Compulsive Scale Short-Form, DRS-15-R = Dispositional Resilience Scale 15-Revised, GAD-7 = Generalized Anxiety Disorder Scale, HSPS = Highly Sensitive Person Scale, M.I.N.I. = Mini-international neuropsychiatric interview, OCI-R = Obsessive-Compulsive Inventory-Revised, PEAS = The Patient EX/RP Adherence Scale, PEAS-P = The Patient EX/RP Adherence Scale, patient-rated, PHQ-9 = Patient Health Questionnaire-9
SCID = Structured Clinical Interview for DSM Disorders (SCID-I: nach DSM-IV für Achse-I-Störungen; SCID-5-RV: Research-Version für DSM 5), SCID-5-RV = Structured Clinical Interview for DSM Disorders, WEMWBS = Warwick- Edinburgh Mental Wellbeing Scale, WSAS = The Work and Social Adjustment Scale, Y-BOCS = Yale-Brown Obsessive Compulsive Scale, Y-BOCS-SR = Yale-Brown Obsessive Compulsive Scale, self-report version.

15.2. Vorveröffentlichung

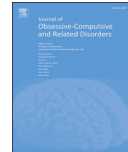
Journal of Obsessive-Compulsive and Related Disorders 43 (2024) 100913



Contents lists available at ScienceDirect

Journal of Obsessive-Compulsive and Related Disorders

journal homepage: www.elsevier.com/locate/jocrd



Exploring the Peaks and Potholes: Understanding positive and negative effects of concentrated exposure treatment for obsessive-compulsive disorder

Lena Jelinek^{a,*}, Anna Serve^a, Saskia Pampuch^a, Jakob Scheunemann^a, Josephine Schultz^a, Franziska Miegel^a, Bjarne Hansen^b, Kristen Hagen^{b,c}, Frances Bohnsack^a, Jürgen Gallinat^a, Amir H. Yassari^a

^a Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

^b Center for Brain Plasticity, Haukeland University Hospital, Bergen, Norway

^c More and Romsdal Hospital Trust, Division of Psychiatry, Molde, Norway

ARTICLE INFO

Keywords:

Intensive treatment
Obsessive-compulsive disorder
ERP
Exposure
CBT
Concentrated treatment
B4DT
Inpatient care

ABSTRACT

The Bergen 4-Day Treatment offers brief concentrated exposure with response prevention (cERP) for obsessive-compulsive disorder (OCD). To date, this intervention has primarily been studied in Norway, and no study has been done on its side effects. We tested the safety, feasibility, and effectiveness of cERP in Germany and compared cERP to a historical inpatient control group.

Thirty-three patients with OCD were treated with cERP. We assessed severity of OCD (primary outcome: Y-BOCS), depression, global functioning, self-esteem, self-efficacy, experiential avoidance, and quality of life at baseline (t0), two weeks after t0 (t1), and three months after t1 (t2). Side effects were assessed at t1 and t2. The changes in OCD were compared to a matched historical inpatient control group ($n = 33$) treated at the same site.

The cERP group improved over time, with a large effect size in OCD symptoms and other outcome measures. Improvement of OCD symptoms over treatment was superior in the cERP group compared to the historical control group, with a medium effect size for OCD symptoms. 53–56% of the participants reported at least one side effect (e.g., exhaustion). cERP is safe and effective in the treatment of OCD and can be successfully implemented in Germany.

1. Introduction

With an estimated lifetime prevalence of 1–3% (Fawcett et al., 2020; Kessler et al., 2012), obsessive-compulsive disorder (OCD) is a common mental disorder. It affects not only the people diagnosed with OCD (e.g., their quality of life, Coluccia et al., 2016; Hauschildt et al., 2010; Macy et al., 2013) but also their relatives (Cicek et al., 2013). Finally, OCD has high social costs (Hollander et al., 2016), with a recent estimate of £4.7 bn for the United Kingdom alone (Kochar et al., 2023). Effective treatment of OCD is, therefore, of great importance, both for affected individuals and for society.

According to international guidelines (e.g., National Institute for Health and Care Excellence (NICE), 2005; Voderholzer et al., 2022), the treatment of choice for OCD is cognitive-behavioral therapy (CBT) with exposure and response prevention (ERP). During ERP, therapists support

their patients in systematically exposing themselves to feared situations or thoughts while refraining from compulsive behavior that is aimed at reducing discomfort momentarily (Foa & Kozak, 1997). Studies have shown that ERP can be successfully implemented in different settings and formats (for a meta-analysis, see Olatunji et al., 2013).

However, not all patients benefit sufficiently from ERP treatment; some patients drop out, and some are reluctant to undergo ERP or only partially respond to ERP treatment (Leeuwerik et al., 2019; Olatunji et al., 2010). While early dropping out or reluctance has been associated with comorbid depression or lack of motivation (Leeuwerik et al., 2019; Moritz et al., 2019), many people with psychological problems also do not seek treatment due to fear of stigmatization (Belloch et al., 2009; Glazier & Wetterneck, 2015). Fear of stigmatization may particularly apply to treatments of long duration as patients may have to explain their absence during treatments. To narrow the treatment gap, it is

* Corresponding author. University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246, Hamburg, Germany.
E-mail address: l.jelinek@uke.de (L. Jelinek).

<https://doi.org/10.1016/j.jocrd.2024.100913>

Received 22 June 2024; Received in revised form 8 October 2024; Accepted 13 October 2024

Available online 15 October 2024

2211-3649/© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

necessary to explore alternative treatment options and formats for people with OCD.

One alternative format for ERP-based treatment is “high-intensity” treatment, referring to a high frequency of therapy/ERP sessions. In one of the first studies to explore the effect of high-frequency ERP, Abramowitz et al. (2003) compared an intensive (15 daily 2 h exposure sessions over 3 weeks) with a “twice-weekly” variant (eight weeks of twice-weekly 2 h exposure sessions plus one telephone contact). Both treatment variants led to a significant improvement in obsessive-compulsive symptoms (OCS) with high effect sizes at the end of therapy. For more studies on intensive CBT for youth or adults, see the systematic review and meta-analysis by Jónsson et al. (2015).

Researchers in Norway have developed a high-frequency or “concentrated” brief ERP (cERP) format, the Bergen-4-day treatment (B4DT). Exposure-based treatment is offered in a concentrated format over four consecutive days by a team of therapists (1:1 ratio of patients to therapists). It represents a combination of group and individual therapy as the individually tailored exposure tasks are conducted in individual sessions and patients reflect upon their experiences several times a day in the group. Learning how other people successfully confront their fears and compulsions, as well as the short treatment duration, motivates patients to engage in exposures, counteracts therapy drop-outs, and reduces fear of stigmatization. For a detailed description of the intervention, see Hansen et al. (2018, 2019) and Havnen et al. (2013, 2014).

Regarding effectiveness, several uncontrolled studies from the treatment developers have shown strong effects and high response and remission rates, with up to 70% of the patients being categorized as remitted (≤ 12 points on the Yale-Brown Obsessive Compulsive Scale [Y-BOCS]) one to four years after treatment (Hansen et al., 2018, 2019). Moreover, treatment effects have been shown to be superior (Hansen et al., 2018) in comparison to the effects of CBT reported in meta-analyses (Öst et al., 2015). Effectiveness has also been confirmed in difficult-to-treat patients, that is, patients with relapse or non-response after treatment (Kvale et al., 2020), and there is first evidence that treatment response is associated with treatment adherence (Tjelle et al., 2021). The only randomized controlled trial (RCT) on the intervention, by Launes et al. (2020), compared the B4DT with a self-help intervention and a waitlist condition. Although B4DT was superior in a decrease of OCS compared to the waitlist and self-help condition, with large between-group effect sizes (Cohen's $d = 2.57$ – 3.86) from pre- to post-treatment, the trial has several limitations. First, the developers conducted the intervention, potentially producing allegiance effects (Dragioti et al., 2015). Second, each group's sample size was small (i.e., $n = 16$). Third, randomization was not optimal, with numerical discrepancies in comorbidity (e.g., depressive disorders) and age at OCD onset. Fourth, the control groups were waitlist and self-help, which are not the most effective treatments for OCD. Thus, more studies are needed to further evaluate this promising treatment, particularly its effects compared to other effective treatment formats. However, RCTs are resource intensive and are associated with several other challenges, such as reduced generalizability and external validity, as patients willing to be randomized and/or undergo experimental treatment may favor participating in RCTs (Wasmann et al., 2019).

Besides efficacy, the systematic assessment of side effects (SE) in psychotherapy research has been advocated by the Lancet Psychiatry Commission (Holmes et al., 2018) and others. SE in psychotherapy refer to unwanted treatment reactions (e.g., symptom deterioration, occurrence of new symptoms, negative effects on relationships with others) caused by correctly applied treatment and can be distinguished from malpractice and unethical behavior caused by incorrectly applied treatment. Generally, the assessment of SE in psychotherapy is scarce, especially in the psychological treatment of OCD (Moritz et al., 2015). One of the few exceptions is an online study by Moritz et al. (2015), in which 93% of the participants with OCD reported at least one SE after psychotherapy. SE may be more frequent in concentrated formats of

exposure treatment. As discussed by Silver et al. (2023), B4DT may be considered a last-resort intervention for many patients with OCD, and it carries a significant risk of dissatisfaction as an adverse outcome, particularly if patients' high expectations for rapid improvement are not entirely met.

Up until now, however, the SE of concentrated exposure treatment have only been investigated in people with anxiety disorders, with 42% of the patients experiencing SE as measured by the Inventory for the Balanced Assessment of Negative Effects of Psychotherapy (INEP) in a recent study by Heinig et al. (2023). Exploring and transparently communicating negative SE of treatments might reduce therapists' concerns (Heinig et al., 2023) and work against the “public relation problem” that has been associated with ERP (Olatunji et al., 2009). Negative beliefs about exposure treatment in OCD are rather common among therapists and contribute to its underuse (e.g., Jelinek et al., 2024).

The aim of the current study was to test the safety, feasibility, and effectiveness of brief cERP. We investigated whether the results for effectiveness of cERP could be (1) replicated at a German site over a follow-up period of three months and (2) extended by comparing cERP treatment results to treatment results of a historical inpatient control group. Based on previous uncontrolled studies (Davíðsdóttir et al., 2019; Hansen et al., 2018; Kvale et al., 2018; Launes et al., 2019), we hypothesized large within-group effect sizes for the decrease in OCD symptomatology (Y-BOCS, primary outcome) from baseline ($T_{0\text{cERP}}$) to post-assessment ($T_{1\text{cERP}}$) and from baseline ($T_{0\text{cERP}}$) to three months follow-up assessment ($T_{2\text{cERP}}$). Secondary outcome parameters were the changes in self-reported obsessive-compulsive symptoms, depression, self-esteem, quality of life, and experience avoidance/psychological flexibility as well as general level of functioning. All measured variables have been shown to improve across treatment for OCD. We planned to compare the changes in OCD symptoms and treatment satisfaction from the concentrated treatment with those of a historical inpatient control group treated in the same hospital. Based on previous results (Hansen et al., 2018), we expected that the effects of cERP at post-treatment would be superior to the results of the historical inpatient controls.

As safety measures, we assessed SE and symptom deterioration (according to the reliable change index [RCI]) after cERP. There is a lack of evidence for negative SE (including symptom deterioration) after cERP in OCD, and we hypothesized that cERP in the study would be safe, that is, that SE (as measured by the Short Inventory of the Assessment of Negative Effects (SIAN), Dietrichkeit et al., 2021) and clinically significant symptom deterioration (as measured by the RCI) would be marginal. We expected these results to be comparable to exposure therapy in anxiety disorders (Heinig et al., 2023). Regarding feasibility, we measured patient satisfaction (CSQ), expecting positive subjective appraisals of brief cERP.

2. Materials and methods

2.1. Study design and procedure for the intervention group

Individuals with OCD were recruited via the Department for Psychiatry and Psychotherapy of the University Medical Center Hamburg-Eppendorf between September 5, 2022, and June 6, 2023, to participate in brief cERP. The inclusion criteria were (a) diagnosis of OCD (according to the Mini-International Neuropsychiatric Interview (MINI), Sheehan et al., 1998), (b) age between 18 and 75, (c) no history of a psychotic disorder (including mania), (d) no current acute suicidality or substance dependency, (e) no pharmacological treatment or stable medication for at least six weeks, and (f) provision of informed consent.

In total, 34 patients fulfilled the inclusion criteria and were scheduled for treatment with cERP in Hamburg, Germany. Two patients were not able to start scheduled treatment due to their contamination-related OCD. These two patients did not live in Hamburg and were overstrained

with external accommodations. One of these two patients was re-scheduled and was able to participate three months later, and the other was considered a drop-out, leading to a total sample of 33 patients.

Patients were between 19 and 67 years old ($M = 35.45$, $SD = 12.94$). The majority were female (60.6 %, $n = 20$). In total, 36.4% ($n = 12$) worked full-time (part-time: $n = 5$; 15.2%, in training or studying: $n = 7$, 21.1%; unemployed: $n = 3$, 9.1%; retired: $n = 1$, 3%). All patients had at least a secondary school certificate, and 19 (57.6%) had at least a bachelor's degree. In total, 57.6% of participants were married or in a stable partnership. The age of onset of OCD varied between 5 and 33 years ($M = 17.88$ years, $SD = 7.43$). A minority of the sample had an early onset of OCD (39.4%; i.e., before the age of 15). The duration of OCD was approx. 18 years ($M = 17.94$). Comorbid depression, based on the MINI, was frequent, with 26 patients (78.8%) fulfilling the criteria for a current ($n = 6$) or lifetime depressive disorder. Regarding medication, 19 participants (57.6%) were not taking psychotropic medication, 13 (39.4%) were taking antidepressants (primarily selective serotonin reuptake inhibitors), and one person was taking antidepressants in combination with an antipsychotic agent.

Participants were assessed at baseline ($T0_{\text{cERP}}$), at post ($T1_{\text{cERP}}$, two weeks after $T0_{\text{cERP}}$), and at follow-up three months after $T1_{\text{cERP}}$ ($T2_{\text{cERP}}$). The study was conducted in compliance with the Declaration of Helsinki (World Medical Association, 2013), approved by the local ethics committee (#LPEK-0512), and preregistered with the German clinical trials registry (#DRKS00030022).

2.2. Historical control group (HC)

Thirty-three people diagnosed with OCD (according to the MINI, Sheehan et al., 1998) who underwent inpatient treatment at the OCD and anxiety ward of the University Medical Center Hamburg-Eppendorf and were assessed during other trials (Miegel et al., under review) were selected based on sociodemographic criteria (i.e., age, gender, baseline Y-BOCS score, OCD content) to match the intervention group. Inpatients were assessed after admission (baseline) to the ward and eight weeks later (post-assessment) with the Y-BOCS, the Obsessive-Compulsive Inventory-Revised (OCI-R), the Rosenberg Self-Esteem Scale (RSE), and the World Health Organization Quality of Life Questionnaire (WHOQOL-BREF). The duration of OCD was approx. 16 years ($M = 15.78$). Comorbid depression, based on the MINI, was frequent, with all 33 patients fulfilling the criteria for a current ($n = 25$) or lifetime depressive disorder. Regarding medication, 27 participants (81.8%) were taking psychotropic medication.

2.3. Intervention

2.3.1. Brief concentrated exposure treatment (cERP): Bergen-4-day treatment (B4DT)

The B4DT was used as the brief cERP. It is described in detail elsewhere (Havnen et al., 2013, 2014). The B4DT is an exposure-based treatment conducted on four consecutive days in a group of patients by a matching number of therapists. The first day in the group setting is dedicated to manualized psychoeducation and planning of exposure sessions. On the second and third days, patients engage in exposures specifically tailored to their needs, with each patient receiving individual support from a therapist across a variety of OCD-provoking situations and continuing to self-expose themselves at home in the mornings and evenings of these days (a total exposure of approximately 20h). During these exposure days, the group gathers to share their progress and the challenges of exposure training. On the afternoon of the third day, family members and caregivers are invited to a psychoeducation session on understanding OCD and how they can best support their relatives' efforts. On the fourth day, the patients are taught strategies for maintaining their behavioral changes, and further self-guided exposures are planned for the following 30 days. A booster session with a therapist is offered three months after treatment. All therapists were

trained by BH and KH.

2.3.2. Inpatient treatment of the historical inpatient control group

During their stay on the specialized ward for anxiety and OCD, patients received a standardized intensive CBT-based inpatient (or residential) treatment program. This consisted of individual CBT therapy sessions (including ERP with therapist-assisted exposure sessions), medication, one virtual reality (VR) exposure treatment session to teach the exposure rationale, various specialized groups (e.g., the MCT-OCD, Jelinek et al., 2018), and non-disorder-specific groups such as social skills training and occupational therapy. ERP was a key element of the inpatient treatment, with inpatients performing in total between 2 and 220 h of ERP during their inpatient stay (median = 21 h, $M = 42.33$).

2.4. Measures

2.4.1. Primary outcome

We used the German version of the Y-BOCS (Goodman et al., 1986; Hand & Büttner-Westphal, 1991) to measure the severity of OCD. It consists of 10 items that are scored from 0 to 4. Internal consistency (Goodman et al., 1989) was confirmed with Cronbach's $\alpha = 0.80$ for the German version (Jacobsen et al., 2003). All Y-BOCS interviews were conducted by trained and supervised staff not involved in the cERP treatment. Interrater reliability was good, with ICC = 0.871, $CI_{95\%}$ [0.719–0.961], and Cronbach's $\alpha = .894$.

2.4.2. Secondary outcomes

We used the German version of the Obsessive-Compulsive Inventory-Revised (OCI-R, Foa et al., 2002; German version: Gönner et al., 2007) as a secondary measure for the severity of OCD. This 18-item scale's total score ranges from 0 to 72, with higher scores indicating greater severity. It is sensitive to change (Abramowitz et al., 2005) and has good psychometric properties for both the total scale and the subscales (Abramowitz & Deacon, 2006; Gönner et al., 2008; Huppert et al., 2007). In the present study, internal consistency was also good, with a Cronbach's α of 0.789.

We used quality of life as a secondary outcome, measured by the global item of the World Health Organization Quality of Life Questionnaire (WHOQOL-BREF, German version: Angermeyer et al., 2000; English version: Skevington et al., 2004). Quality of life is measured on a five-point Likert scale ranging from 5 = *very good* to 1 = *very bad*.

The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001; German version: Löwe et al., 2004) was used to assess comorbid depression as a secondary outcome. The PHQ-9 is a well-known self-report scale for the evaluation of depression symptoms. Its nine items are measured on a four-point Likert scale, with each item being rated on a scale from 0 to 3 and total scores ranging from 0 to 27. It has been shown to possess excellent internal consistency (Cronbach's $\alpha = 0.86$ –0.89) and good criteria validity (Kroenke et al., 2001). In the present study, internal consistency was also good, with a Cronbach's α of 0.84.

To measure self-esteem as a secondary outcome, we used the Rosenberg Self-Esteem Scale (RSE, Rosenberg, 1965), a valid and reliable scale consisting of ten items assessed on a scale from 1 to 4 with total scores ranging from 10 to 40 (with higher scores indicating higher self-esteem). In the present study, internal consistency was excellent, with a Cronbach's α of 0.908.

We employed the Brief Experiential Avoidance Questionnaire (BEAQ, Gámez et al., 2014; German version: Schaeuffele et al., 2022) to assess experiential avoidance. The 15-item questionnaire is a short form of the Multidimensional Experiential Avoidance Questionnaire (MEAQ), which has been shown to have good internal consistency and validity (Schaeuffele et al., 2022). It is rated on a six-point scale ranging from 1 = *strongly disagree* to 6 = *strongly agree*. In the present study, internal consistency was also good, with a Cronbach's α of 0.85.

The General Self-Efficacy Scale (GSES, Schwarzer & Jerusalem, 2003) was used to measure self-efficacy. The GSES measures subjective

Table 1
Description of the cERP and the historical inpatient control sample.

	cERP (n = 33)	Historical inpatient controls (HIS)(n = 33)	Statistics
Gender			
female	20 (60.6%)	16 (48.5%)	$\chi^2(1) = 0.978, p = .323$
male	13 (39.4%)	17 (51.5%)	
Age in years	35.46 (12.94)	32.79 (12.06)	$t(64) = 0.866, p = .390$
Education			
≥ 12 years of school	30 (90.9%)	23 (71.9%)	$\chi^2(1) = 3.910, p = .048$
< 12 years of school	3 (9.1%)	9 (28.1%)	
Illness duration	17.94 (14.73)	15.78 (12.86)	$t(59) = 0.608, p = .546$
Medication status			
Medicated	17 (51.5%)	27 (81.8%)	$\chi^2(1) = 6.818, p = .009$
Non-medicated	16 (48.5%)	6 (18.2%)	
Y-BOCS total score	25.18 (3.96)	24.39 (6.79)	$t(51.50) = 0.576, p = .567$
OCI-R	24.67 (9.20)	27.39 (12.07)	$t(59.79) = 1.033, p = .306$
OCD content*			
contamination	14 (50.0)	14 (50.0)	$\chi^2(1) = 1.000, p = 1.00$
checking	14 (53.8)	12 (46.2)	$\chi^2(1) = 0.614, p = 0.801$
symmetry	7 (58.3)	5 (41.7)	$\chi^2(1) = 0.407, p = 0.751$
taboo thoughts	24 (57.1)	18 (42.9)	$\chi^2(1) = 2.357, p = 0.125$

Notes. OCD = obsessive compulsive disorder. cERP = concentrated exposure and response prevention.

Y-BOCS = Yale-Brown Obsessive Compulsive Scale. OCI-R = Obsessive-Compulsive Inventory-Revised. *According to the Y-BOCS checklist.

controllability and general competence expectations. We applied the 10-item short version for overall self-efficacy, which uses a four-point rating scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, 4 = *strongly agree*). The GSES has shown good internal consistency (Cronbach's $\alpha = 0.82$ –0.86). In the present study, internal consistency was good, with a Cronbach's α of 0.846.

The Global Assessment of Functioning (GAF, APA, 1994) scale was used to measure patients' overall psychosocial functioning, including psychological, social, and occupational functioning. The GAF is rated on a scale from 1 = *persistent danger of severely hurting oneself or others* to 100 = *superior functioning in a wide range of activities*.

2.4.3. Side effects (SE)

We assessed SE with the SIAN (Dietrichkeit et al., 2021), a questionnaire which was developed based on the structure of the Inventory for the Assessment of Negative Effects of Psychotherapy (INEP; Ladwig et al., 2014) at T1_{cERP} and T2_{cERP}. We modified the items for the current study; that is, we specifically mentioned the name of the intervention in each item (i.e., "... due to the B4DT...") and adjusted the items related to psychopathology to OCD symptomatology (e.g., "my intrusive thought" instead of "my depression"). The final questionnaire had a total of 21 items, with 13 items rated on a seven-point bipolar scale (agreement/disagreement from -3 to +3) and eight items rated on a four-point scale (1 = *true*, 2 = *somewhat true*, 3 = *somewhat not true*, 4 = *not true*). To count the number of SE, we coded values for items 1 to 13 from -3 to -1 (=agreement) as well as for items 14 to 21 for values 1 and 2 (1 = *true*, 2 = *rather true*) as "yes" (=1), indicating the occurrence of an SE. All other values were coded 0 and were not coded as SE. All items are displayed in e-supplements A and B. In the present study, internal consistency was good, with a Cronbach's α of 0.931.

The symptoms assessed with the SIAN include both OCD-related and general psychopathological symptoms (e.g., helplessness, guilt, rumination) as well as items regarding social context (i.e., coming across to others and relationship with others). Some items that might not apply to all participants could be answered with "does not apply." The internal consistency ($n = 32$), with Cronbach's $\alpha_{12} = 0.908$, was excellent in our sample. The quantity of the SE overall, as well as for each item, was calculated for assessment at post-assessment (T1_{cERP}) and at three-month follow-up (T2_{cERP}).

2.4.4. Feasibility

The Client Satisfaction Questionnaire-8 (CSQ-8; Attkisson, 1982; German version [ZUF-8]: Schmidt & Wittman, 2008) was employed to measure the patients' subjective appraisal of the brief cERP at

post-assessment (T1_{cERP}). The eight-item scale is rated on a four-point Likert scale with scores ranging from 8 to 32 (higher scores indicate higher satisfaction). It has good internal consistency (Cronbach's $\alpha = .83$ –0.93) and moderate convergent validity (Attkisson, 1982; Kriz et al., 2008). In the present study, internal consistency was good, with a Cronbach's α of 0.857.

2.5. Sample size

For sample size, we calculated that the sample size for the dependent sample *t*-test should be at least 59, aiming for a test power of 0.85 and an alpha level of 0.05 (two-tailed) with a large effect size ($f = 0.40$). To account for potential drop-outs, we aimed at a sample size of 33 patients in each group.

2.6. Data analysis

To test change in primary and secondary outcomes, we planned repeated measure analyses of variance (ANOVA) and Student's *t*-test for dependent samples. For comparison between the intervention group (cERP) and the historical inpatient control group (HIS), we planned analyses of covariance (ANCOVA) with the group as the between-subject factor (cERP vs. HIS). Assessment points differed between samples, with the cERP group being assessed at baseline (T0_{cERP}), post-treatment (T1_{cERP}; 14 days after T0_{cERP}), and three months later (T2_{cERP}) and the HIS group being assessed at baseline (T0_{HIS}) and eight weeks later (post; T1_{HIS}). To compare treatment effects of cERP and inpatient treatment (historical controls), we compared the change in HIS group (T0_{HIS} to T1_{HIS}) to the change in the brief cERP group for the time intervals (1) T0_{cERP} to T1_{cERP} and (2) T0_{cERP} to T2_{cERP}. All analyses used baseline scores as covariates (to control for regression to the mean (Barnett et al., 2005; Rausch et al., 2003)).

Also, we calculated response ($\geq 35\%$ decline on the Y-BOCS) and remission rates (≤ 12 points on the Y-BOCS total score) according to the international consensus (Mataix-Cols et al., 2016). Symptom deterioration was calculated using the RCI as described by Jacobson and Truax (1991).

We report effect sizes using partial eta squared for ANOVAs, referring to $\eta_p^2 \approx .01$ as small, $\eta_p^2 \approx 0.06$ as medium, and $\eta_p^2 \approx 0.14$ as large effect sizes, based on Kinnear and Gray (2009). For the *t*-test, we report Cohen's *d*, following the parameters of Cohen (1988) for small ($d \approx 0.2$), medium ($d \approx 0.5$), and large ($d \approx 0.8$) effect sizes.

Table 2a
Means (standard deviations) in primary and secondary outcome parameters.

	cERP			Historical controls (HIS)	
	<i>M (SD)</i>			<i>M (SD)</i>	
	Baseline (T0 _{cERP}) <i>n</i> = 33	Post (T1 _{cERP}) <i>n</i> = 33	Follow-up (T2 _{cERP}) <i>n</i> = 32	Baseline (T0 _{HIS}) <i>n</i> = 33	Post (T1 _{HIS}) <i>n</i> = 33
Y-BOCS - total score*	25.18 (3.96)	15.24 (5.63)	13.84 (7.08)	24.39 (6.79)	17.94 (7.79)
Y-BOCS - omissions	12.48 (2.31)	8.06 (2.66)	7.03 (3.69)	12.85 (3.64)	9.49 (3.81)
Y-BOCS - compulsions	12.70 (1.94)	7.18 (3.39)	6.81 (3.80)	11.55 (4.73)	8.45 (4.74)
OCI-R	24.67 (9.19)	11.79 (5.64)	13.25 (6.92)	27.40 (12.07)	20.91 (12.35) ^a
GAF	64.64 (13.51)	75.82 (12.87)	78.87 (13.19)	-	-
PHQ-9	9.91 (5.36)	6.39 (5.30)	6.22 (4.19)	-	-
	Baseline (T0 _{cERP}) <i>n</i> = 32	Post (T1 _{cERP}) <i>n</i> = 32	Follow up (T2 _{cERP}) <i>n</i> = 31	Baseline (T0 _{HIS}) <i>n</i> = 33	Post (T1 _{HIS}) <i>n</i> = 33
RSE	25.56 (3.09)	24.93 (2.40)	25.29 (2.57)	27.14 (7.23) ^b	26.87 (7.72)
BEAQ	49.97 (12.36)	44.00 (8.46)	43.45 (9.81)	-	-
GSES	23.97 (4.53)	28.31 (4.40)	28.16 (6.16)	-	-
WHOQOL-BREF	2.88 (0.87)	3.66 (0.60)	3.53 (0.84)	2.33 (1.11)	3.41 (0.84)

Notes. * primary outcome. cERP = concentrated exposure and response prevention. Y-BOCS = Yale-Brown Obsessive Compulsive Scale. OCI-R = Obsessive Compulsive Disorder-Revised. GAF = Global Assessment of Functioning. PHQ-9 = Patient-Health-Questionnaire-9. RSE = Rosenberg Self-Esteem Scale. BEAQ = Experiential Avoidance. GSES = General Self-Efficacy Scale. WHOQOL-BREF = global item of the World Health Organization Quality of Life Questionnaire.

3. Results

3.1. Sample characteristics

By design, patients treated with brief cERP did not differ from the HIS in age, gender, illness duration, OCD symptomatology at baseline (as measured by the Y-BOCS and the OCI-R), or OCD content (see Table 1). Samples differed in years of school education (a higher proportion of patients with more years of school education in the brief cERP than in the historical control group), current or lifetime diagnosis of depression (all historical inpatient controls (HC)[HC] were diagnosed with a current or lifetime depressive disorder, but only 79% of the cERP sample, $\chi^2(1) = 7.831, p = .005$), and medication status (more historical controls than brief cERP patients were being medicated).

3.2. Effectiveness

3.2.1. Within-group effects

As can be seen in Tables 2a-c and Fig. 1, the brief cERP group improved over time, with large effect sizes in the primary (Y-BOCS) and secondary outcome measures (OCI-R, GAF, PHQ-9, GSES, quality of life). Effect sizes for improvement in self-esteem (RSE) and experiential avoidance (BEAQ) were small to moderate.

3.2.2. Between-group effects

Improvement over treatment was superior in the brief cERP group compared to the HIS group with medium size for the primary outcome (Y-BOCS total score; see Fig. 1). For secondary outcomes, a large effect size was found for improvement in self-assessed OCD symptomatology

Table 2b
Analyses for primary and secondary outcomes within the intervention group.

	t-test	t-test	ANOVA
	T0 _{cERP} VS. T1 _{cERP}	T0 _{cERP} VS. T2 _{cERP}	
Y-BOCS - total score (primary outcome)	$t(32) = 11.717, p < .001, d = 2.040, CI_{95\%} [1.431, 2.637]$	$t(31) = 9.536, p < .001, d = 1.686, CI_{95\%} [1.137, 2.223]$	$F(1.68, 52.09) = 77.8, p < .001, \eta_p^2 = 0.715$
Y-BOCS - obsessions	$t(32) = 10.897, p < .001, d = 1.897, CI_{95\%} [1.317, 2.466]$	$t(31) = 8.315, p < .001, d = 1.470, CI_{95\%} [0.961, 1.967]$	$F(1.523, 47.209) = 59.891, p < .001, \eta_p^2 = 0.659$
Y-BOCS - compulsions	$t(32) = 10.208, p < .001, d = 1.777, CI_{95\%} [1.220, 2.323]$	$t(31) = 8.970, p < .001, d = 1.586, CI_{95\%} [1.056, 2.104]$	$F(2.62) = 63.721, p < .001, \eta_p^2 = 0.673$
OCI-R	$t(32) = 7.579, p < .001, d = 1.319, CI_{95\%} [0.845, 1.783]$	$t(31) = 6.753, p < .001, d = 1.194, CI_{95\%} [0.732, 1.644]$	$F(1.629, 50.496) = 42.679, p < .001, \eta_p^2 = 0.579$
GAF	$t(32) = 8.083, p < .001, d = -1.407, CI_{95\%} [-1.885, -0.917]$	$t(31) = 7.975, p < .001, d = -1.410, CI_{95\%} [-1.896, -0.912]$	$F(2.62) = 44.252, p < .001, \eta_p^2 = 0.588$
PHQ-9	$t(32) = 4.775, p < .001, d = 0.831, CI_{95\%} [0.429, 1.223]$	$t(31) = 5.398, p < .001, d = 0.954, CI_{95\%} [0.529, 1.368]$	$F(2.62) = 16.860, p < .001, \eta_p^2 = 0.352$
RSE	$t(31) = 1.360, p = .092, d = 0.240, CI_{95\%} [-0.113, 0.590]$	$t(30) = 0.312, p = .378, d = 0.056, CI_{95\%} [-0.297, 0.408]$	$F(1.520, 45.595) = 0.874, p = .397, \eta_p^2 = 0.028$
BEAQ	$t(31) = 3.289, p = .003, d = 0.581, CI_{95\%} [0.202, 0.953]$	$t(30) = 3.393, p = .002, d = 0.609, CI_{95\%} [0.221, 0.989]$	$F(1.652, 49.566) = 9.240, p < .001, \eta_p^2 = 0.235$
GSES	$t(31) = 6.619, p < .001, d = -1.170, CI_{95\%} [-1.616, -0.713]$	$t(30) = 4.541, p < .001, d = -0.816, CI_{95\%} [-1.218, -0.403]$	$F(2.60) = 18.564, p < .001, \eta_p^2 = 0.382$
WHOQOL-BREF	$t(31) = 6.256, p < .001, d = 1.106, CI_{95\%} [-1.542, -0.659]$	$t(30) = 4.768, p < .001, d = 0.856, CI_{95\%} [-1.264, -0.438]$	$F(2.60) = 20.492, p < .001, \eta_p^2 = 0.406$

Notes. cERP = concentrated exposure and response prevention. Y-BOCS = Yale-Brown Obsessive Compulsive Scale. OCI-R = Obsessive Compulsive Disorder-Revised. GAF = Global Assessment of Functioning. PHQ-9 = Patient-Health-Questionnaire-9. RSE = Rosenberg Self-Esteem Scale. BEAQ = Experiential Avoidance. GSES = General Self-Efficacy Scale. WHOQOL-BREF = global item of the World Health Organization Quality of Life Questionnaire.

Table 2c

Analyses for primary and secondary outcomes between the intervention group and the historical control group.

	ANCOVA	
	T0 _{cERP} to T1 _{cERP} vs. T0 _{HIS} to T1 _{HIS}	T0 _{cERP} to T2 _{cERP} vs. T0 _{HIS} to T1 _{HIS}
Y-BOCS - Total Score (primary outcome)	$F(1, 62) = 6.709, p = .012, \eta_p^2 = 0.098$	$F(1, 63) = 4.109, p = .047, \eta_p^2 = 0.061$
Y-BOCS - Obsessions	$F(1, 62) = 6.764, p = .012, \eta_p^2 = 0.098$	$F(1, 63) = 2.842, p = .097, \eta_p^2 = 0.043$
Y-BOCS - Compulsions	$F(1, 62) = 4.984, p = .030, \eta_p^2 = 0.074$	$F(1, 63) = 4.070, p = .048, \eta_p^2 = 0.061$
OCI-R	$F(1, 61) = 8.768, p = .004, \eta_p^2 = 0.126$	$F(1, 62) = 14.187, p < .001, \eta_p^2 = 0.186$
GAF	-	-
PHQ-9	-	-
RSE	$F(1, 56) = 0.649, p = .424, \eta_p^2 = 0.011$	$F(1, 57) = 1.475, p = .230, \eta_p^2 = 0.025$
BEAQ	-	-
GSES	-	-
WHOQOL-BREF	$F(1, 61) = 0.282, p = .597, \eta_p^2 = 0.005$	$F(1, 62) = 1.287, p = .261, \eta_p^2 = 0.020$

Notes. cERP = concentrated exposure and response prevention. Y-BOCS = Yale-Brown Obsessive Compulsive Scale. OCI-R = Obsessive Compulsive Disorder-Revised. GAF = Global Assessment of Functioning. PHQ-9 = Patient-Health-Questionnaire-9. RSE = Rosenberg Self-Esteem Scale. BEAQ = Experiential Avoidance. GSES = General Self-Efficacy Scale. WHOQOL-BREF = global item of the World Health Organization Quality of Life Questionnaire. ANCOVA: All analyses used baseline scores as covariates.

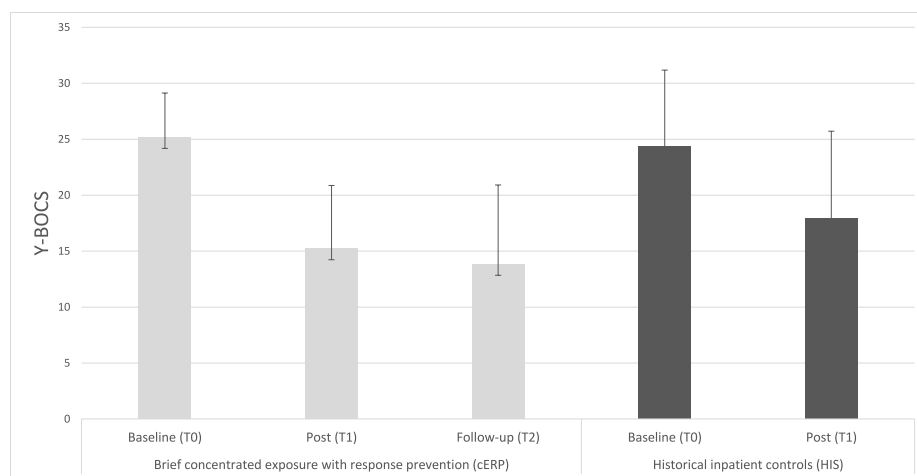


Fig. 1. Means and standard deviations of the brief concentrated exposure with response prevention (cERP) at baseline (T0), post-assessment 2 weeks later (T1) and follow-up assessment three months later (T2) and the historical inpatient control sample at baseline (T0) and post-assessment 8 weeks later (T1).

(OCI-R). The effect size for the changes in self-esteem and quality of life was small. Results are displayed in Tables 2a, 2b, and 2c.

3.2.3. Response and remission rates

In Table 3, response and remission rates are displayed for the patients in the brief cERP group for different time points (52–63% responder, 33–47% in remission) and the HIS group (36% responder, 24% in remission). Response and remission rates tended to be larger in the cERP than in the HIS group (the latter were receiving inpatient treatment) when rates at the three-month follow-up assessment for the brief cERP group (T2_{cERP}) were compared to rates at the eight weeks after baseline assessment for the HIS group (T1_{HIS}). Rates were similar between groups for the brief cERP group. Rates right after treatment

(T1_{cERP}, 14 days after baseline) were chosen and compared to rates at the eight weeks after baseline assessment for the HIS group (T1_{HIS}).

3.3. Safety

3.3.1. Side effects (SE)

At T1_{cERP} and T2_{cERP}, 53–56% of the participants reported at least one SE, agreeing fully or partially to at least one item, and 25–28% agreed to a single item. On average the participants reported 1–2 SE (T1_{cERP}: $M = 1.88, SD = 3.55$; T2_{cERP}: $M = 1.34, SD = 1.99$). Being exhausted after exposure sessions (38.2%) and being disappointed about not feeling better yet (17.6%) were expressed most frequently at T1_{cERP} (for details, see e-supplements A and B). The most frequently reported

Table 3Numbers (percent) of patients classified as responders ($\geq 35\%$ reduction in the Y-BOCS) or as in remission (≤ 12 points in the Y-BOCS).

	cERP		Historical Controls	Statistics	
	T1 _{cERP} $n = 33$	T2 _{cERP} $n = 32$	T1 _{HIS} $n = 33$	T1 _{cERP} vs. T1 _{HIS}	T2 _{cERP} vs. T1 _{HIS}
Responder	17 (51.5%)	20 (62.5%)	12 (36.4%)	$\chi^2(1) = 1.538, p = .215$	$\chi^2(1) = 4.440, p = .035$
In remission	11 (33.3%)	15 (46.9%)	8 (24.2%)	$\chi^2(1) = 0.665, p = .415$	$\chi^2(1) = 3.640, p = .056$

Notes. cERP = concentrated exposure and response prevention. Y-BOCS = Yale-Brown Obsessive Compulsive Scale. HIS = historical inpatient control group.

SE at T2_{cERP} were worries about how they came across to others (38.2%) and unpleasant feelings such as grief, disgust, or guilt (26.5%), as well as negative relationships with others (20.6%).

3.3.2. Symptom deterioration (RCI)

According to the RCI, no significant deterioration, according to the RCI, appeared from T0_{cERP} to T1_{cERP} and from T0_{cERP} to T2_{cERP} in the Y-BOCS score.

3.3.3. Feasibility: treatment satisfaction (CSQ-8)

Treatment satisfaction with the brief cERP as measured with the CSQ-8 was generally high (total score: $M = 29.30$, $SD = 3.06$). For example, all patients evaluated the quality of the brief cERP as either “excellent” or “good” and confirmed that they would recommend brief cERP to a friend with similar problems (ratings either “yes, definitely” or “yes, I think so”).

4. Discussion

The aim of the current study was to test the safety, feasibility, and effectiveness of brief cERP by conducting an uncontrolled pilot study and comparing cERP treatment results to treatment results of a historical inpatient control group.

For safety, we assessed side effects (SE) of brief cERP. SE have generally been neglected in psychotherapy research (Moritz et al., 2015) and, until this study, had never been assessed in cERP for OCD. As expected, SE were generally low in brief cERP for OCD, with an average of 1–2 SE at post-assessment and three months later and with 53–56% of the patients reporting at least one SE at either T1 or T2. The most frequently reported SE, were being exhausted after exposure sessions and worries about how they came across to others. This largely corresponds to 42% of the patients with anxiety disorders experiencing SE after intensified exposure, as measured by the INEP in a recent study by Heinig et al. (2023). According to Linden et al. (2018), SE are unwanted events that are temporally contingent on and preferably causally related to the treatment process, and they also occur during correctly executed treatment. By assessing the SE shortly after the treatment, which supervised clinicians performed to ensure that the treatment was carried out correctly, at post-assessment the SE measured in the current study are most likely related to brief cERP and no other factors (Linden et al., 2018). Still, after more time had passed since the end of treatment, patients reported more SE (such as negative feelings, negative impact on relationships, and worries about how they came across to others) at the three-month follow-up assessment. Due to the short duration of the treatment (four days), some SE inevitably needed time to occur, especially if they involved interactions with other people. Still, the total percentage of people experiencing SE remained similar between post-treatment and follow-up. As a further measure of safety, we measured symptom deterioration (assessed by the RCI) and found no clinically significant symptom deterioration, which underlines the safety of brief cERP.

Regarding feasibility, we measured patient satisfaction using the CSQ, showing high rates of satisfaction with an average total score of 29.30 ($SD = 3.06$). These findings are largely similar to those reported in Norwegian samples; for example, a mean satisfaction score of 29.5 ($SD = 2.7$) was reported by Launes et al. (2019). Thus, our results further support the acceptance of the treatment.

Primary and secondary outcome measures also improved over brief cERP treatment. Effect sizes were large for improvement in OCS (primary outcome: Y-BOCS, secondary outcome: OCI-R), global functioning (GAF), depression (PHQ-9), self-efficacy (GSES), and quality of life (global item of the WHOQOL-BREF), with a Cohen's d of up to $d = 2.040$, $CI_{95\%} [1.431, 2.637]$ for the total score of the Y-BOCS. However, effect sizes are below the effect sizes reported by the developers of the intervention in this format. Similarly, our response and remission rates at three-month follow-up are 62.5% and 46.9%, lower than the approx.

70% reported in Norwegian and Icelandic samples (e.g. (Daviðsdóttir et al., 2019; Hansen et al., 2018; Kvale et al., 2018)). However, they are similar to results in a difficult-to-treat sample (Kvale et al., 2020), in which remission rates were between 56.5% (post-treatment) and 47.9% (12-month follow-up), and they are above the remission rates of 28% reported in recent a Finnish study using the same brief cERP treatment (Silver et al., 2023).

First, allegiance effects need to be discussed, which always occur when a treatment is applied by its developers. In the current study the developers of the B4DT conducted the clinical training of the therapists and provided supervision, but they did not conduct the therapy themselves, making allegiance effects less likely. Second, the screening and treatment preparation phase may have differed from previous studies, with motivation and readiness for change being assessed and prepared differently before the start of cERP (compare Silver et al. (2023)). In line with this argument, the clinical training in cERP (B4DT) focused on the delivery of the four-day intervention, not the selection of patients and treatment preparation process, potentially leaving room for improvement in this pre-treatment phase. Generally, research is necessary to identify ‘best-fit’ indicators for different treatment formats.

As in many pilot studies aiming at testing the feasibility and safety of an unestablished intervention, inclusion of patients is maximized in terms of clinical experience with the intervention and may attract patients seeking a last-resort intervention who have a potentially higher risk of disappointment (Silver et al., 2023). Descriptively, at 57% the percentage of patients experiencing obsessions in the form of taboo thoughts was rather large. While we were able to match OCD contents between the cERP and the HIS groups, the challenges of including these patients in health care (Besiroglu et al., 2004; García-Soriano et al., 2014) and conducting ERP in this patient group have been outlined previously (e.g., Bruce et al., 2018). Moreover, it has been described that the observed duration of untreated illness (DUI) is particularly high among patients with taboo thoughts (Dell’Osso et al., 2014) and that a longer DUI is associated with a worse long-term outcome (Perris et al., 2023). Potentially, the high percentage patients with taboo thoughts may have negatively affected treatment responses in our sample.

To compensate for the uncontrolled design, we compared the cERP group's data to a historical sample treated in the same clinic, which showed superiority of the cERP with moderate to large effect sizes for improvement in OCS in comparison to inpatient treatment. In a previous publication by Hansen et al. (2018), results of an uncontrolled study were mathematically compared to effects of CBT reported in a meta-analysis (Öst et al., 2015), which summarize many different formats of CBT and settings. Up until now, brief cERP has only been compared in an RCT with a self-help intervention and a waiting condition, with rather weak control conditions and suboptimal randomization (see introduction). Using a matched historical inpatient sample that was treated at the same site is a rather strong and effective active control condition, despite certain limitations that are outlined in the following section.

The several limitations of the current study mainly refer to the use of a historical control group. First, the comparability of the two samples should be discussed. Our control samples represent a sample of patients receiving inpatient (or residential) treatment. This includes inpatients receiving ERP and other interventions during treatment (e.g., occupational therapy). Still, ERP was the main focus of treatment, and the ERP dose was numerically similar between the two groups, with approximately 20 h ERP during B4DT and an average of 21 h of ERP in the inpatient sample. Even so, we cannot attribute all treatment effects of the inpatient stay to ERP as the patients' treatment was multifaceted.

Moreover, the two samples may be from different populations. Although carefully matched, the two samples differed in years of school education (a higher proportion of patients had more years of school education in the cERP than in the historical control group), medication status (more historical controls than cERP patients were being medicated), and diagnosis of depression (more patients with a current or

lifetime diagnosis of depression in the inpatient group). This may point to rather complex and severe cases in the control group. Generally, inpatient treatment may include more complex and severe cases than outpatient treatment. However, this is not necessarily the case. Inpatient or residential treatment may also be indicated in Germany if no specialized OCD treatment is locally available and/or would require a long waiting time. Unfortunately, this is often the case for OCD patients because ERP treatment is unavailable and underused in Germany (Jelinek et al., 2024; Moritz et al., 2019; Voderholzer et al., 2015), forcing patients to seek help in a ward specializing in exposure treatment.

Additionally, a substantial proportion of the B4DT patients was recruited via our inpatient ward, potentially leading to more complex cases in this group, too (see discussion above on a difficult-to-treat sample (Kvale et al., 2020)). In conclusion, while our use of a historical control group is of value, it does have limitations and does not meet the high standards of an RCT; the current results need to be interpreted against the background of a potential selection bias. Second, side effects were not measured in the historical control group. Thus, we were not able to compare the groups on all measures. Third, assessment times differed between the intervention group and the historical control group.

Fourth, we did not compare the treatments regarding costs. We can assume that B4DT offers significant savings as long-term, residential/inpatient therapies require sustained resource allocation over extended periods, driving up direct and indirect costs for healthcare systems. If B4DT, by contrast, achieves comparable outcomes in a fraction of the time, as suggested by the current evidence, fewer days of illness, for example, would allow a faster return to the workplace and reduce indirect health costs. Moreover, B4DT's concentrated treatment nature allows it to transcend geographical barriers. Patients outside the typical catchment area can access treatment, making specialized care more accessible to those who might otherwise face logistical and financial challenges (e.g., including longer travel times to receive treatment).

In addition to the limitations regarding the control group, we only included a rather short-term follow-up interval. Also, our small sample size is limited, which is important in such a heterogeneous disorder as OCD and we used a previous version of the OCI-R, in which the neutralizing items were found to have problematic psychometric properties; this version also includes hoarding items, which the DSM-5 does not identify as diagnostic of OCD. Future studies should consider using the OCI-12 version (Abramovitch et al., 2021).

5. Conclusions

Our results confirmed that cERP is safe, feasible, and effective in the treatment of individuals with OCD and has marginal negative side effects (e.g., exhaustion). Future research should thus include a randomized controlled design and should also investigate factors of the mechanism contributing to the efficacy and effectiveness of the treatment (Tjelle et al., 2021), including the role of the selection process and pre-treatment phase.

CRedit authorship contribution statement

Lena Jelinek: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Anna Serve:** Writing – review & editing, Project administration, Formal analysis. **Saskia Pampuch:** Writing – review & editing, Project administration, Methodology, Investigation, Conceptualization. **Josephine Schultz:** Writing – review & editing, Project administration. **Franziska Miegel:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization. **Bjarne Hansen:** Writing – review & editing, Supervision,

Funding acquisition. **Kristen Hagen:** Writing – review & editing, Supervision. **Frances Bohnsack:** Writing – review & editing, Resources, Project administration, Investigation, Funding acquisition. **Jürgen Gallinat:** Writing – review & editing, Resources, Conceptualization. **Amir H. Yassari:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Statement of ethics

This study protocol was reviewed and approved by the local ethics committee of the center for psycho social medicine, approval number #LPEK-0512. Consent to participate statement: Written informed consent was obtained from all participants to participate in the study.

Funding sources

The study was partially funded by the Peter Möhrle Stiftung, the Hamburger Spendenparlament, and the Werner Otto Foundation (WOI). Training of staff in brief concentrated exposure treatment was funded by the Kvali Trust fund.

Declaration of competing interest

Lena Jelinek, Anna Serve, Saskia Pampuch, Jakob Scheunemann, Frances Bohnsack, Josephine Schultz, Franziska Miegel, Kristen Hagen, Jürgen Gallinat, and Amir H. Yassari report no conflict of interest. Bjarne Hansen developed the B4DT.

Acknowledgments

The authors thank Maren Schäfer, Josephine Gehlenborg, Luisa Tegtmeier, Holger Zapf, Anastasia Basmanov, Karin Liebert, Maren Duwe, and Violetta Burmeister for their help.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jocrd.2024.100913>.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References

- Abramovitch, A., Abramowitz, J. S., & McKay, D. (2021). The OCI-12: A syndromally valid modification of the obsessive-compulsive inventory-revised. *Psychiatry Research*, 298, Article 113808. <https://doi.org/10.1016/j.psychres.2021.113808>
- Abramowitz, J. S., & Deacon, B. J. (2006). Psychometric properties and construct validity of the Obsessive-Compulsive Inventory-Revised: Replication and extension with a clinical sample. *Journal of Anxiety Disorders*, 20(8), 1016–1035. <https://doi.org/10.1016/j.janxdis.2006.03.001>
- Abramowitz, J. S., Foa, E. B., & Franklin, M. E. (2003). Exposure and ritual prevention for obsessive-compulsive disorder: Effects of intensive versus twice-weekly sessions. *Journal of Consulting and Clinical Psychology*, 71(2), 394–398. <https://doi.org/10.1037/0022-006X.71.2.394>. American Psychological Association.
- Abramowitz, J. S., Tolin, D. F., & Diefenbach, G. J. (2005). Measuring change in OCD: Sensitivity of the obsessive-compulsive inventory-revised. *Journal of Psychopathology and Behavioral Assessment*, 27(4), 317–324. <https://doi.org/10.1007/s10862-005-2411-y>
- Angermeyer, M. C., Kilian, R., & Matschinger, H. (2000). *WHOQOL-100 und WHOQOL-Bref. Handbuch für die deutschsprachige Version der WHO Instrumente zur Erfassung von Lebensqualität [WHOQOL-100 und WHOQOL-Bref. Handbook for the German version of the WHO instruments for the assessment of quality of life]*. Hogrefe.
- APA. (1994). In A. P. Association (Ed.), *Diagnostic and statistical manual of mental disorders* (4th ed.).
- Attkisson, C. C. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning*, 5(3), 233–237. [https://doi.org/10.1016/0149-7189\(82\)90074-X](https://doi.org/10.1016/0149-7189(82)90074-X)

- Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34(1), 215–220. <https://doi.org/10.1093/ije/dyh299>
- Belloch, A., del Valle, G., Morillo, C., Carrió, C., & Cabedo, E. (2009). To seek advice or not to seek advice about the problem: The help-seeking dilemma for obsessive-compulsive disorder. *Social Psychiatry and Psychiatric Epidemiology*, 44(4), 257–264. <https://doi.org/10.1007/s00127-008-0423-0>
- Besiroglu, L., Cilli, A. S., & Askin, R. (2004). The predictors of health care seeking behavior in obsessive-compulsive disorder. *Comprehensive Psychiatry*, 45(2), 99–108. <https://doi.org/10.1016/j.comppsy.2003.12.010>
- Bruce, S. L., Ching, T. H. W., & Williams, M. T. (2018). Pedophilia-themed obsessive-compulsive disorder: Assessment, differential diagnosis, and treatment with exposure and response prevention. *Archives of Sexual Behavior*, 47(2), 389–402. <https://doi.org/10.1007/s10508-017-1031-4>
- Cicek, E., Cicek, I. E., Kayhan, F., Uguz, F., & Kaya, N. (2013). Quality of life, family burden and associated factors in relatives with obsessive-compulsive disorder. *General Hospital Psychiatry*, 35(3), 253–258. <https://doi.org/10.1016/j.genhosppsych.2013.01.004>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. 2nd edition. Academic Press.
- Coluccia, A., Fagiolini, A., Ferretti, F., Pozza, A., Costoloni, G., Bognesi, S., & Goracci, A. (2016). Adult obsessive-compulsive disorder and quality of life outcomes: A systematic review and meta-analysis. *Asian Journal of Psychiatry*, 22, 41–52. <https://doi.org/10.1016/j.ajp.2016.02.001>
- Davíðsdóttir, S. D., Sigurjónsdóttir, Ö., Ludvígsdóttir, S. J., Hansen, B., Laukvik, I. L., Hagen, K., ... Kvale, G. (2019). Implementation of the Bergen 4-day treatment for obsessive compulsive disorder in Iceland. *Clinical Neuropsychiatry*, 16(1), 33–38.
- Dell'Osso, B., Benatti, B., Oldani, L., Spagnolin, G., & Altamura, A. C. (2014). Differences in duration of untreated illness, duration, and severity of illness among clinical phenotypes of obsessive-compulsive disorder. *CNS Spectrums*, 20(5), 474–478. <https://doi.org/10.1017/S1092852914000339>
- Dietrichkeit, M., Hagemann-Goebel, M., Nestoriuc, Y., Moritz, S., & Jelinek, L. (2021). Side effects of the metacognitive training for depression compared to a cognitive remediation training in patients with depression. *Scientific Reports*, 11(1), 7861. <https://doi.org/10.1038/s41598-021-87198-8>
- Dragioti, E., Dimoliatis, I., Fountoulakis, K. N., & Evangelou, E. (2015). A systematic appraisal of allegiance effect in randomized controlled trials of psychotherapy. *Annals of General Psychiatry*, 14(1), 25. <https://doi.org/10.1186/s12991-015-0063-1>
- Fawcett, E. J., Power, H., & Fawcett, J. M. (2020). Women are at greater risk of OCD than men. *The Journal of Clinical Psychiatry*, 81(4). <https://doi.org/10.4088/JCP.19r13085>
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment*, 14(4), 485–496. <https://doi.org/10.1037/1040-3590.14.4.485>
- Foa, E. B., & Kozak, M. J. (1997). *Mastery of obsessive-compulsive disorder: A cognitive-behavioral approach therapist guide*. Oxford University Press.
- Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., Suzuki, N., & Watson, D. (2014). The brief experiential avoidance questionnaire: Development and initial validation. *Psychological Assessment*, 26(1), 35–45. <https://doi.org/10.1037/a0034473>
- García-Soriano, G., Rufer, M., Delsignore, A., & Weidt, S. (2014). Factors associated with non-treatment or delayed treatment seeking in OCD sufferers: A review of the literature. *Psychiatry Research*, 220(1–2), 1–10. <https://doi.org/10.1016/j.psychres.2014.07.009>
- Glazier, K., & Wetterneck, C. (2015). Stigma and shame as barriers to treatment for obsessive-compulsive and related disorders. *Journal of Depression & Anxiety*, 4(3). <https://doi.org/10.4172/2167-1044.1000191>
- Göner, S., Leonhart, R., & Ecker, W. (2007). Das Zwangsinventar OCI-R: die deutsche version des obsessive-compulsive inventory-revised. *PPmP-Psychotherapie-Psychosomatik: Medizinische Psychologie*, 57(09/10), 395–404.
- Göner, S., Leonhart, R., & Ecker, W. (2008). The Obsessive-Compulsive Inventory-Revised (OCI-R): Validation of the German version in a sample of patients with OCD, anxiety disorders, and depressive disorders. *Journal of Anxiety Disorders*, 22(4), 734–749. <https://doi.org/10.1016/j.janxdis.2007.07.007>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown obsessive compulsive scale. I. Development, use, and reliability. *Archives of General Psychiatry*, 46(11), 1006–1011. <https://doi.org/10.1001/archpsyc.1989.01810110048007>
- Goodman, W. K., Rasmussen, S. A., Price, L. H., Mazure, C., Heninger, G. R., & Charney, D. S. (1986). *Yale-Brown obsessive compulsive scale (Y-BOCS)*.
- Hand, I., & Büttner-Westphal, H. (1991). Die Yale-Brown Obsessive Compulsive Scale (Y-BOCS): Ein halbstrukturiertes Interview zur Beurteilung des Schweregrades von Denk- und Handlungszwängen [The Yale-Brown Obsessive Compulsive Scale (Y-BOCS): A semi-structured interview for the assessment of ob. *Verhaltenstherapie*, 1(3), 223–225. <https://doi.org/10.1159/000257972>
- Hansen, B., Hagen, K., Öst, L.-G., Solem, S., & Kvale, G. (2018). The bergen 4-day OCD treatment delivered in a group setting: 12-Month follow-up. *Frontiers in Psychology*, 9.
- Hansen, B., Kvale, G., Hagen, K., Havnen, A., & Öst, L.-G. (2019). The bergen 4-day treatment for OCD: Four years follow-up of concentrated ERP in a clinical mental health setting. *Cognitive Behaviour Therapy*, 48(2), 89–105. <https://doi.org/10.1080/16506073.2018.1478447>
- Hauschildt, M., Jelinek, L., Randjbar, S., Hottenrott, B., & Moritz, S. (2010). Generic and illness-specific quality of life in obsessive-compulsive disorder [Journal Article]. *Behavioural and Cognitive Psychotherapy*, 38(4), 417–436. <https://doi.org/10.1017/S1352465810000275>
- Havnen, A., Hansen, B., Haug, E. T., Prescott, P., & Kvale, G. (2013). Intensive group treatment of obsessive-compulsive disorder: A pilot study. *Clinical Neuropsychiatry*, 10(3 SUPPL.1), 48–55.
- Havnen, A., Hansen, B., Öst, L. G., & Kvale, G. (2014). Concentrated ERP delivered in a group setting: An effectiveness study. *Journal of Obsessive-Compulsive and Related Disorders*, 3(4), 319–324. <https://doi.org/10.1016/j.jocrd.2014.08.002>
- Heinig, I., Knappe, S., Hoyer, J., Wittchen, H. U., Richter, J., Arolt, V., Deckert, J., Domschke, K., Hamm, A., Kircher, T., Lucken, U., Margraf, J., Neudeck, P., Rief, W., Straube, B., Ströhle, A., Pauli, P., & Pittig, A. (2023). Effective and tolerable: Acceptance and side effects of intensified exposure for anxiety disorders. *Behavior Therapy*, 54(3), 427–443. <https://doi.org/10.1016/j.beth.2022.11.001>
- Hollander, E., Doernberg, E., Shavitt, R., Waterman, R. J., Soreni, N., Veltman, D. J., Sahakian, B. J., & Fineberg, N. A. (2016). The cost and impact of compulsivity: A research perspective. *European Neuropsychopharmacology*, 26(5), 800–809. <https://doi.org/10.1016/j.euroneuro.2016.02.006>
- Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., Roiser, J. P., Bocking, C. L. H., O'Connor, R. C., Shafraan, R., Moulds, M. L., & Craske, M. G. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry*, 5(3), 237–286. [https://doi.org/10.1016/S2215-0366\(17\)30513-8](https://doi.org/10.1016/S2215-0366(17)30513-8)
- Huppert, J. D., Walther, M. R., Hajcak, G., Yadin, E., Foa, E. B., Simpson, H. B., & Liebowitz, M. R. (2007). The OCI-R: Validation of the subscales in a clinical sample. *Journal of Anxiety Disorders*, 21(3), 394–406. <https://doi.org/10.1016/j.janxdis.2006.05.006>
- Jacobsen, D., Kloss, M., Fricke, S., Hand, I., & Moritz, S. (2003). Reliabilität der deutschen Version der Yale-Brown Obsessive Compulsive Scale. *Verhaltenstherapie*, 13(2), 111–113. <https://doi.org/10.1159/000072184>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19. <https://doi.org/10.1037/0022-006X.59.1.12>
- Jelinek, L., Balzar, A., Moritz, S., Lieberher, S., & Yassari, A. H. (2024). Exposure therapy in patients with harm-related obsessive-compulsive disorder: The theory-practice gap and its relation to experiential avoidance and negative beliefs about exposure. *Journal of Obsessive-Compulsive and Related Disorders*, 40(June 2023), Article 100860. <https://doi.org/10.1016/j.jocrd.2024.100860>
- Jelinek, L., Zielke, L., Hottenrott, B., Miegel, F., Cludius, B., Sure, A., & Demiralay, C. (2018). Patients' perspectives on treatment with metacognitive training for OCD. *Zeitschrift für Neuropsychologie*, 29(1), 20–28. <https://doi.org/10.1024/1016-264X/a000211>
- Jónsson, H., Kristensen, M., & Arendt, M. (2015). Intensive cognitive behavioural therapy for obsessive-compulsive disorder: A systematic review and meta-analysis. *Journal of Obsessive-Compulsive and Related Disorders*, 6, 83–96. <https://doi.org/10.1016/j.jocrd.2015.04.004>
- Kessler, R., Petukhova, M., Sampson, N. A., Zaslavsky, A., & Wittchen, H. U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, 21(3), 169–184. <https://doi.org/10.1002/mpr.1359>
- Kinney, P. R., & Gray, C. D. (2009). *SPSS 16 made simple*. Psychology Press.
- Kochar, N., Ip, S., Vardanaga, V., Sireau, N. T., & Fineberg, N. A. (2023). A cost-of-illness analysis of the economic burden of obsessive-compulsive disorder in the United Kingdom. *Comprehensive Psychiatry*, 127(September), Article 152422. <https://doi.org/10.1016/j.comppsy.2023.152422>
- Kriz, D., Nübling, R., Steffanowski, A., Rieger, J., & Schmidt, J. (2008). Patientenzufriedenheit: Psychometrische Reanalyse des ZUF-8. [Client Satisfaction: psychometric reanalysis of the ZUF-8]. *DRV-Schriften Band*, 77, 84–85. Deutsche Rentenversicherung Bund.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kvale, G., Hansen, B., Björgvinsson, T., Børtveit, T., Hagen, K., Haseth, S., Kristensen, U. B., Launes, G., Ressler, K. J., Solem, S., Strand, A., van den Heuvel, O. A., & Öst, L.-G. (2018). Successfully treating 90 patients with obsessive compulsive disorder in eight days: The Bergen 4-day treatment. *BMC Psychiatry*, 18(1), 323. <https://doi.org/10.1186/s12888-018-1887-4>
- Kvale, G., Hansen, B., Hagen, K., Abramowitz, J. S., Børtveit, T., Craske, M. G., Franklin, M. E., Haseth, S., Himle, J. A., Hystad, S., Kristensen, U. B., Launes, G., Lund, A., Solem, S., & Öst, L.-G. (2020). Effect of D-Cycloserine on the effect of concentrated exposure and response prevention in difficult-to-treat obsessive-compulsive disorder. *JAMA Network Open*, 3(8), Article e2013249. <https://doi.org/10.1001/jamanetworkopen.2020.13249>
- Ladwig, I., Rief, W., & Nestoriuc, Y. (2014). Welche Risiken und Nebenwirkungen hat Psychotherapie? - Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP). *Verhaltenstherapie*, 24(4), 252–263. <https://doi.org/10.1159/000367928>
- Launes, G., Hagen, K., Öst, L.-G., Solem, S., Hansen, B., & Kvale, G. (2020). The Bergen 4-day treatment (B4DT) for obsessive-compulsive disorder: Outcomes for patients treated after initial waiting list or self-help intervention. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00982>
- Launes, G., Laukvik, I. L., Sunde, T., Klovning, I., Hagen, K., Solem, S., Öst, L.-G., Hansen, B., & Kvale, G. (2019). The Bergen 4-day treatment for obsessive-compulsive disorder: Does it work in a new clinical setting? *Frontiers in Psychology*, 10.
- Leeuwierik, T., Cavanagh, K., & Strauss, C. (2019). Patient adherence to cognitive behavioural therapy for obsessive-compulsive disorder: A systematic review and meta-analysis. *Journal of Anxiety Disorders*, 68, Article 102135. <https://doi.org/10.1016/j.janxdis.2019.102135>

- Linden, M., Strauß, B., Scholten, S., Nestoriuc, Y., Brakemeier, E.-L., & Wasilewski, J. (2018). Definition und Entscheidungsschritte in der Bestimmung und Erfassung von Nebenwirkungen von Psychotherapie. *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie*, 68(09/10), 377–382. <https://doi.org/10.1055/a-0619-5949>
- Löwe, B., Kroenke, K., Herzog, W., & Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders*, 81, 61–66. [https://doi.org/10.1016/S0165-0327\(03\)00198-8](https://doi.org/10.1016/S0165-0327(03)00198-8)
- Macy, A. S., Theo, J. N., Kaufmann, S., & Ghazzaoui, R. B. (2013). Quality of life in obsessive compulsive disorder. *CNS Spectrums*, 18(1), 21–33. <https://doi.org/10.1017/S1092852912000697>
- Mataix-Cols, D., Fernandez de La Cruz, L., Nordsletten, A. E., Lenhard, F., Isomura, K., & Simpson, H. B. (2016). Towards an international expert consensus for defining treatment response, remission, recovery and relapse in obsessive-compulsive disorder. *World Psychiatry*, 15(1), 80–81. <https://doi.org/10.1002/wps.20299>
- Moritz, S., Fieker, M., Hottenrott, B., Seeralan, T., Cludius, B., Kolbeck, K., Gallinat, J., & Nestoriuc, Y. (2015). No pain, no gain? Adverse effects of psychotherapy in obsessive-compulsive disorder and its relationship to treatment gains. *Journal of Obsessive-Compulsive and Related Disorders*, 5, 61–66. <https://doi.org/10.1016/j.jocrd.2015.02.002>
- Moritz, S., Külz, A., Voderholzer, U., Hillebrand, T., McKay, D., & Jelinek, L. (2019). “Phobie à deux” and other reasons why clinicians do not apply exposure with response prevention in patients with obsessive-compulsive disorder. *Cognitive Behaviour Therapy*, 48(2), 162–176. <https://doi.org/10.1080/16506073.2018.1494750>
- National Institute for Health and Care Excellence (NICE). (2005). Obsessive-compulsive disorder: Core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder. NICE clinical guideline 31. In *NICE clinical guidelines*. British psychological society.
- Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: A review of meta-analytic findings. *Psychiatric Clinics of North America*, 33(3), 557–577. <https://doi.org/10.1016/j.psc.2010.04.002>
- Olatunji, B. O., Davis, M. L., Powers, M. B., & Smits, J. A. J. (2013). Cognitive-behavioral therapy for obsessive-compulsive disorder: A meta-analysis of treatment outcome and moderators. *Journal of Psychiatric Research*, 47(1), 33–41. <https://doi.org/10.1016/j.jpsychires.2012.08.020>
- Olatunji, B. O., Deacon, B. J., & Abramowitz, J. S. (2009). The cruelest cure? Ethical issues in the implementation of exposure-based treatments. *Cognitive and Behavioral Practice*, 16(2), 172–180. <https://doi.org/10.1016/j.cbpra.2008.07.003>
- Öst, L.-G., Havnén, A., Hansen, B., & Kvale, G. (2015). Cognitive behavioral treatments of obsessive-compulsive disorder: A systematic review and meta-analysis of studies published 1993–2014. *Clinical Psychology Review*, 40, 156–169. <https://doi.org/10.1016/j.cpr.2015.06.003>
- Perris, F., Cipolla, S., Catapano, P., Sampogna, G., Luciano, M., Giallonardo, V., Del Vecchio, V., Fabrizio, M., Fiorillo, A., & Catapano, F. (2023). Duration of untreated illness in patients with obsessive-compulsive disorder and its impact on long-term outcome: A systematic review. *Journal of Personalized Medicine*, 13(10), 1453. <https://doi.org/10.3390/jpm13101453>
- Rausch, J. R., Maxwell, S. E., & Kelley, K. (2003). Analytic methods for questions pertaining to a randomized pretest, posttest, follow-up design. *Journal of Clinical Child and Adolescents*, 32(3), 467–486. https://doi.org/10.1207/S15374424JCCP3203_15
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.
- Schaeuffele, C., Knaevelsrud, C., Renneberg, B., & Boettcher, J. (2022). Psychometric properties of the German brief experiential avoidance questionnaire (BEAQ). *Assessment*, 29(7), 1406–1421. <https://doi.org/10.1177/10731911211010955>
- Schmidt, J., & Wittman, W. W. (2008). Fragebogen zur Messung der Patientenzufriedenheit - ZUF-8. In *Diagnostische Verfahren in der Rehabilitation* (pp. 248–255).
- Schwarzer, R., & Jerusalem, M. (2003). SWE. Skala zur Allgemeinen Selbstwirksamkeitserwartung [GSES. General Self-Efficacy Scale]. In *Open test archive*. Leibniz-Institut für Psychologie (ZPID). <https://doi.org/10.23668/psycharchives>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. (1998). The MINI International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview. *Journal of Clinical Psychiatry*, 59(suppl. 20), 22–23.
- Silver, E., Isometsä, E., Baryshnikov, I., Heino, H., Stenberg, J.-H., & Saarni, S. E. (2023). Bergen 4-Day Treatment (B4DT) for obsessive-compulsive disorder – an observational pilot study of a treatment protocol in Finland. *Psychiatria Fennica*, 198–217. <https://www.psychiatriantutkimussaatio.fi/wp-content/uploads/2023/10/Psychiatria-Fennica-2023-Silver.pdf>
- Skevington, S. M., Lotfy, M., & O’Connell, K. A. (2004). The World Health Organization’s WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 13(2), 299–310. <http://www.ncbi.nlm.nih.gov/pubmed/15085902>
- Tjelle, K., Opstad, H. B., Solem, S., Launes, G., Hansen, B., Kvale, G., & Hagen, K. (2021). Treatment adherence as predictor of outcome in concentrated exposure treatment for obsessive-compulsive disorder. *Frontiers in Psychiatry*, 12. <https://doi.org/10.3389/fpsyt.2021.667167>
- Voderholzer, U., Favreau, M., Rubart, A., Staniloiu, A., Wahl-Kordon, A., Zurowski, B., & Kathmann, N. (2022). Therapie der Zwangsstörungen: Empfehlungen der revidierten S3-Leitlinie Zwangsstörungen. *Nervenarzt, Der*, 93(7), 678–687. <https://doi.org/10.1007/s00115-022-01336-9>
- Voderholzer, U., Schlegel, S., Diedrich, A., Külz, A. K., Thiel, N., Hertenstein, E., Schwartz, C., Rufer, M., Herbst, N., Nissen, C., Hillebrand, T., Osen, B., Stengler, K., Jelinek, L., & Moritz, S. (2015). Supply with cognitive behaviour therapy as first-line therapy in patients with obsessive-compulsive disorder. *Verhaltenstherapie*, 25.
- Wasmann, K. A., Wijsman, P., Van Dieren, S., Bemelman, W., & Buskens, C. (2019). Partially randomised patient preference trials as an alternative design to randomised controlled trials: Systematic review and meta-analyses. *BMJ Open*, 9(10), 1–12. <https://doi.org/10.1136/bmjopen-2019-031151>
- World Medical Association. (2013). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Journal of the American Medical Association*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>