

Increasing the Effectiveness of Psychotherapy in Routine Care through Transdiagnostic Online Modules? Randomized Controlled Trial Investigating Blended Care

Carmen Schaeuffele^{1*}, Augustin Mutak^{1,3}, Solveig Behr¹, Friederike Fenski², Leona Hammelrath¹, Marie Puetz², Steffi Pohl¹, Christine Knaevelsrud^{1**} & Johanna Boettcher^{2**}

¹Freie Universität Berlin, Department of Education and Psychology

²Psychologische Hochschule Berlin

³University of Zagreb, Faculty of Humanities and Social Sciences

---- note: this manuscript is a preprint version and has not been peer-reviewed ----

Author Note

** equal contribution

* corresponding author:

Carmen Schaeuffele, Freie Universität Berlin, Education and Psychology, Clinical-Psychological Intervention, Schwendener Str. 27, 14195 Berlin, carmen.schaeuffele@fu-berlin.de

The data, analysis procedure, and code for all analyses is provided on OSF

https://osf.io/5cn37/?view_only=c6e736a45cb14ea7ba0b99b466e15d70 [note: this is a view-only link created for peer-review; we will make the project publicly available on OSF and provide the doi upon acceptance of this manuscript].

Abstract

Objective: Blended Care (BC), the integration of Internet-based interventions into psychotherapy (PT), is thought of as a promising approach to enhance PT's effectiveness and efficiency. This randomized controlled trial aimed to investigate the usage, implementation, and effectiveness of BC with transdiagnostic online modules compared to PT in routine care.

Method: Psychotherapists in routine care recruited 1159 patients who were randomized to BC or PT. Outcomes included self-reported anxiety, depression, satisfaction with life and level of functioning, as well as therapist-rated severity and changes. Outcomes were measured at baseline, 6 weeks, 12 weeks, and 6 months. We examined whether BC and PT groups changed differently over time using linear mixed models. We also investigated differences in sessions and terminations and report usage metrics of the BC platform.

Results: Contrary to our hypotheses, we did not find differences between BC and PT in outcomes, including anxiety, depression, satisfaction with life, level of functioning, therapist-rated severity / changes, and satisfaction with treatment at 6-months post-randomization (all $p > 0.05$). BC and PT did not differ in the number of sessions or terminations. Regarding usage of the BC platform, 534 patients (91.6%) received at least one online chapter, with $M = 7.26$ ($SD = 7.01$) online chapters assigned on average.

Conclusions: In real-world applications of BC, therapists have considerable flexibility in integrating IBIs with sessions. Our findings suggest that the benefits observed in more controlled BC settings may not fully translate to routine care, possibly due to variations in implementation and adherence to BC.

Keywords: Blended Care, Blended Therapy, Routine Care, Transdiagnostic, Personalization

Public health significance statement

This study suggests that blending digital components and psychotherapy may not result in increased effectiveness of psychotherapy or reduce the number of sessions in routine care.

Increasing the Effectiveness of Psychotherapy in Routine Care through Transdiagnostic Online Modules? Randomized Controlled Trial Investigating Blended Care

Mental disorders are highly prevalent, with comorbidity being the rule rather than the exception (Jacobi et al., 2014). Mental disorders significantly impact the quality of life, social functioning, and even physical health (Arias et al., 2022). At the same time, less than 25% of patients with a mental disorder seek professional help (Heinig et al., 2021). For those who do, the average waiting time to start PT in German outpatient care is approximately five months, due to the current demand for PT exceeding available treatment capacities (Bundespsychotherapeutenkammer, 2011). Integrating Internet-based interventions (IBI) into routine psychotherapeutic treatment (“BC”) has the potential to alleviate the strain on the mental healthcare system. By IBIs assuming some treatment responsibilities, such as providing psychoeducation, interactive therapeutic exercises, and symptom monitoring, IBIs may free up psychotherapeutic resources and expand the reach of mental health support. While the evidence for the efficacy of standalone IBI is relatively sound, BC as a research field is only emerging. A first systematic review found that BC may boost the effects of PT and increase cost-effectiveness and adherence rates (Erbe et al., 2017). Given the heterogeneous possibilities of blending IBIs and PT and the dynamic research activity on BC in recent years, it is worthwhile to nuance these findings with more recent empirical evidence. The vast majority of studies have investigated BC with a disorder-specific focus in the context of depression. Studies investigating the additional benefit of using an IBI as an add-on to traditional PT without integrating the IBI into the sessions show small to medium between-group effects in favor of BC (Berger et al., 2018; Schuster et al., 2020). More integrated BC setups with a higher degree of blend between face-to-face therapy and IBIs show comparable effects between BC and PT (Kemmeren et al., 2023; Kooistra et al.,

2019; Mathiasen et al., 2022). Most frequently, the BC setup is manualized and involves alternating sessions between patients working on the IBI by themselves and a therapy session with a psychotherapist. The comparable effectiveness suggests that integrating IBIs into therapy can optimize mental health care delivery, offering individuals the same level of therapeutic benefits, often with fewer sessions. A recent pilot study in depressed patients investigated an integrated, modular BC setup to PT with an equal number of PT sessions between both groups and found indications that BC may be superior in reducing symptoms of depression and anxiety with small effects, albeit the study was underpowered (Kalde et al., 2023). Another recent study in a transdiagnostic sample of outpatients indicated that an emotion regulation-focused IBI added to PT shows comparable effects to PT alone (Bielinski et al., 2023).

Outside of study contexts, the uptake of BC in routine care is still limited, which may be overcome by adapting the intervention to routine care: For one, by involving patients and therapists in the development and implementation process, and secondly, by enabling psychotherapists to flexibly use the online materials and tailor the blended treatment to each patient's specific needs (Titzler et al., 2018). For Germany specifically, it may also require the BC intervention to be applicable for different therapeutic approaches beyond CBT, as a large proportion of psychotherapists in routine care are not CBT-based. To address these implementation barriers, we developed a transdiagnostic, transtheoretical, and modular BC intervention called TONI that addresses shared mechanisms and symptoms across the whole spectrum of mental health and applies a common language of PT (Behr et al., 2024).

This study aimed to investigate the usage, implementation, and effectiveness of BC with TONI compared to regular face-to-face PT in routine care in Germany. We expected BC with TONI to be superior in alleviating anxiety and depression to routine PT with a small effect.

Beyond that, we analyzed the differences between BC and PT in secondary outcomes, including level of functioning, satisfaction with life, number of sessions, and therapist-rated symptom distress and changes.

Method

Trial Design

In a parallel-group, two-arm randomized controlled superiority trial conducted in routine care in Germany, this study aims to assess the comparative effectiveness of BC to PT. We reported the study procedure in detail in the study protocol (Schaeuffele et al., 2022) The study was preregistered (DRKS00028536) and approved by the ethics committee of Psychologische Hochschule Berlin (EK202121).

Procedure and Participants

All psychotherapists, irrespective of psychotherapeutic approach, participating in routine PT within the public health insurance system in Germany with internet access in their practice were eligible to participate. Psychotherapists enrolled their patients into the trial according to these inclusion criteria: Patients were over 18 years old, had internet access, had sufficient German language skills, and were able to read and write. There were no exclusion criteria based on diagnoses or symptom severity due to the naturalistic nature of the trial. Patients needed to be enrolled during the initial consultation sessions with a therapist that are used to establish diagnosis and rapport before applying for coverage through their public health insurance. Patients and therapists provided electronic informed consent before participating.

Randomization

We implemented randomization at the patient level and thus chose block randomization to ensure that psychotherapists treat patients of both groups. After patients provided their consent

and baseline data, the TONI platform automatically randomly allocated patients to either BC or routine PT in a 1:1 ratio applying a pre-specified, concealed randomization algorithm.

Interventions

Routine PT

All patients received routine PT. Psychotherapists were either psychologists or physicians who underwent post-graduate training in PT and had a license to practice as psychotherapists. Psychotherapists could be licensed in CBT, psychodynamic (including psychoanalysis), systemic therapy, or a combination. Routine PT comprises heterogeneous treatment lengths and procedures. Due to the naturalistic nature of this trial, additional (mental) healthcare utilization was not limited.

BC (Routine PT and TONI)

In the BC group, psychotherapists could use TONI in addition to routine PT. TONI is a comprehensive digital toolkit that includes twelve transdiagnostic and transtheoretical modules (see Table 1), as well as a range of self-monitoring tools. Psychotherapists could individualize BC by selecting and ordering specific modules or chapters and by implementing TONI flexibly into the therapy process. Due to the naturalistic structure of the trial, we did not instruct psychotherapists to apply a specific implementation of TONI, i.e., they could choose the frequency, duration, and format of TONI. In addition, patients could request access to modules. To allow for a high blend of sessions and the IBI, psychotherapists were able to view patients' entries in the modules' record forms. Patients were free to share additional TONI components, like symptom monitoring. Psychotherapists could enable messaging to allow for written communication with their patients through the platform.

Table 1. Transdiagnostic online modules of TONI

Module	Description of major concepts and components
Development	Chapter 1: biopsychosocial model of distress; Chapter 2: biographical history of symptoms; Chapter 3: genogram and intergenerational mental health
Values & Goals	Chapter 1: motivation and barriers to change; Chapter 2: identifying personal values; Chapter 3: therapy goals
Mindfulness	Chapter 1: introduction to mindfulness; Chapter 2: mindfulness-based exercises (internal and external)
Getting active	Chapter 1: relationship between activity and mood; Chapter 2: identifying fulfilling and enjoyable activities; Chapter 3: implementing activities in daily life
Thoughts	Chapter 1: relationship between thoughts and emotions; Chapter 2: life story shaping thoughts and beliefs; Chapter 3: mentalization and different perspectives
Understanding feelings	Chapter 1: understanding feelings and emotional stance; Chapter 2: messages behind feelings; Chapter 3: life story shaping emotional experience.
Dealing with feelings	Chapter 1: tolerating emotions; Chapter 2: avoidance/defense mechanisms and new reactions
Self-worth & strengths	Chapter 1: concept of self-worth and personal standards; Chapter 2: kindness and self-compassion; Chapter 3: identifying strengths and resources
Communication	Chapter 1: meanings of messages; Chapter 2: communication patterns and reflexion of relationships; Chapter 3: nonviolent communication of needs and wishes
Body & well-being	Chapter 1: relationship between body and mind; Chapter 2: eating and body image; Chapter 3: sleep; Chapter 4: movement; Chapter 5: pain and bodily complaints; Chapter 6: sexuality; Chapter 7: stress
Addictive substances and behaviors	Chapter 1: core concepts of substance and non-substance use; Chapter 2: pros and cons of substance use/addictive behavior; Chapter 3: dealing with cravings; Chapter 4: relapse prevention and dealing with relapses
Collaboration	This module is targeted at a patient's significant other Chapter 1: significant other's own mental well-being; Chapter 2: basics of therapy and emotional well-being; Chapter 3: joint exercises for patients and significant others

Outcomes

All outcomes were assessed at baseline, 6 weeks, 12 weeks, and 6 months post-randomization, unless otherwise specified.

Primary outcome. The primary outcome was self-reported anxiety and depression, assessed with the GAD-7 and PHQ-8. The Patient Health Questionnaire-8 (PHQ-8) is an eight-item tool for depressive symptoms (Gräfe et al., 2004; Kroenke et al., 2009; Kroenke & Spitzer, 2002) that omits the suicidal ideation item from the PHQ-9. Participants rate their emotional well-being over the past two weeks on a 4-point scale (0 = "not at all" to 3 = "nearly every day"). Scores range from 0 to 24, with higher scores indicating more severe depression. The PHQ-8 is internally reliable ($\alpha = .88$) and valid in clinical and non-clinical settings (Wu et al., 2020). The Generalized Anxiety Disorder Screener (GAD-7) is a seven-item measure of general anxiety (Löwe et al., 2008; Spitzer et al., 2006). Participants rate anxiety symptoms over the past two weeks on a 4-point scale (0 = "not at all" to 3 = "nearly every day"). The GAD-7 is unidimensional and internally reliable ($\alpha = .89$).

Secondary outcome. Secondary self-reported outcomes included satisfaction with life (SWLS) and level of functioning (AQoL-8D). The Satisfaction with Life Scale (SWLS) measures overall life satisfaction (Diener et al., 1985; Glaesmer et al., 2011). Participants rate five items on a 7-point Likert scale (1 = "strongly disagree" to 7 = "strongly agree"). Scores range from 5 to 35, with higher scores indicating greater satisfaction. The SWLS is internally reliable ($\alpha = .92$; (Glaesmer et al., 2011)). The Assessment of Quality of Life (AQoL-8D) is a 35-item instrument assessing overall quality of life (Richardson et al., 2014). It covers eight domains: independent living, pain, senses, mental health, happiness, coping, relationships, and self-worth. The AQoL-8D is internally reliable ($\alpha = .80$; (Richardson et al., 2014)). We also

measured satisfaction with treatment at 6-months post-treatment with an adaptation of the CSQ (Attkisson & Zwick, 1982; Schmidt et al., 1989). Comprising eight items, it measures various aspects of patient contentment, such as overall satisfaction, perceived benefits, and willingness to recommend the treatment to others. The CSQ has demonstrated good reliability ($\alpha = .92$).

Therapist ratings. Therapists rated symptom distress with the Clinical Global Impression Scale (CGI). The CGI evaluates illness severity and improvement over time (Berk et al., 2008). It includes two components: CGI-S rates current severity on a seven-point scale from "normal" to "among the most extremely ill patients." CGI-I rates improvement or deterioration from "very much improved" to "very much worse." and was accordingly not measured at baseline. Both scales are reliable ($\alpha = .71$; (Berk et al., 2008)). Therapists also provided information on the number of sessions at 6-months. Since TONI was intended for use during active PT, they were also asked to indicate on the platform or during the 6-month assessment whether therapy had concluded. If so, they were prompted to indicate whether the PT ended prematurely and who initiated the termination.

Sample size

We hypothesized that BC would be superior to routine PT, with an incremental effect of $d = 0.2$. 400 patients per group are needed to detect this effect at 6-months post-randomization in an RCT with a t-test and alpha set to 0.05 (two-sided). The nesting of patients in therapists was taken into account by considering therapists' effects: An assumed average caseload of 5 new patients per year and an intraclass correlation of 0.05 (Schiefele et al., 2017) result in a design effect of 1.2 [design effect = $1 + (nC - 1) \times ICC = 1 + (5 - 1) \times 0.05 = 1.20$]. In addition, a dropout rate of 20% was assumed. This results in a sample size of at least ($400 \times \text{design effect} \times$

missing values = $400 \times 1.2 \times 1.2$) 1152 patients (~ 231 psychotherapists) to detect the assumed effect of $d = 0.2$.

Statistical Analyses

All analyses were performed using R 4.4.1 (R Core Team, 2013). For the primary analyses, we included data from all randomized patients (intention-to-treat (ITT) analysis) and conducted the analyses blindly with respect to the intervention group label. Analyses on the usage of TONI were conducted separately once the main analyses were completed. The data, analysis procedure, and code for all analyses is provided on OSF

https://osf.io/5cn37/?view_only=c6e736a45cb14ea7ba0b99b466e15d70 [note: this is a view-only link created for peer-review; we will make the project publicly available on OSF and provide the doi upon acceptance of this manuscript].

Before analyzing the data, we multiply imputed the missing values of the primary and secondary outcomes using predictive mean matching in the mice package for R (van Buuren & Groothuis-Oudshoorn, 2011). In the performed imputation, the multilevel structure of the data was taken into account with the patient ID as the cluster variable. The variables used for the imputation can be divided into two main groups. The first group consists of questionnaires measuring clinically relevant variables. The full overview of the battery is available at the OSF. We used patients' questionnaire data for imputation, with the exception of the WAI-SR questionnaire, where therapists' data was used. Regarding the item-level data imputation, we used item-level data to impute other items from the same scale or subscale, but only scale- or subscale-level data to impute items from other scales or subscales to avoid collinearity issues. Regarding the scale- or subscale-level data imputation, scale and subscale scores were passively imputed. The second group of variables used for imputation consists of certain demographic and

medical characteristics of the patients and of study-related variables: patient age, gender, location, socioeconomic status, employment status, relationship status, education, distance, preferred and applied PT method, existence of previous diseases, family history of diseases, medication, therapy termination, whether the patient experiences discrimination and whether they attended therapy. The variables were chosen on a theoretical basis, with the exception of diseases, for which we found significant correlation with the outcomes. We imputed 30 values for each missing value. We used predictive mean matching as the imputation method, both for the time point- and patient-level variables. The imputation procedure converged normally with $\hat{R} < 1.1$. Imputation was performed using Freie Universität Berlin's High Performance Cluster (Bennett et al., 2020).

Although we randomized patients to intervention groups, we tested for group differences in all baseline variables. Variables with group differences of at least a small effect size were controlled for in the subsequent analyses. For evaluating the effectiveness of blended versus routine PT, we first conducted linear multilevel regression using the lme4 package (Bates et al., 2014). We conducted a separate analysis for every primary and secondary outcome. The predictors included in the multilevel models were timepoint, study group and their interaction. We also included education level as a covariate, as we found group differences in this variable between the two study groups. We fitted a random intercept model with a random intercept for every patient. We assessed assumptions of the multilevel models by checking diagnostic plots of residuals versus fitted values, distribution of the residuals, and the distribution of the random intercepts. No significant deviations from the model assumptions was detected. The plots pertaining to the assumption tests can be found in the supplementary materials. The results were pooled across imputations. Secondly, we calculated the proportion of patients who achieved

reliable change based on the data. Reliable change was determined based on PHQ-8 (change of six points or more) and GAD-7 scores (change of four points or more) (Porter et al., 2024). For the significance tests, we applied the multivariate t-distribution correction (Kotz & Nadarajah, 2004).

Transparency and Openness

We report how we determined our sample size. De-identified data and analysis code are available at https://osf.io/5cn37/?view_only=c6e736a45cb14ea7ba0b99b466e15d70 [link created for masked peer review. We will make this project publicly available with a stable identifier upon acceptance of this manuscript]. All analyses were conducted using R 4.4.1. This RCT was preregistered prospectively at [removed for masked review] and the study procedures were detailed in a study protocol [removed for masked review].

Results

The study's flow is depicted in Figure 1. Table 2 shows baseline characteristics for patients randomized to BC vs. PT. Participants were mostly female, in their mid-thirties, in a relationship, educated, and employed. A total of 39.54% lived in a rural area, and 32.01% were on medication. Additionally, 42.02% reported at least one health issue, and 47.89% had undergone PT in the past. The most common diagnoses were affective, anxiety, and adjustment disorders, with half of the sample having at least one comorbid diagnosis. There were significant differences with at least small effect size between BC and routine PT regarding education, which we subsequently controlled for in statistical analyses (see Table 5). There were no considerable or significant differences in baseline values of the outcomes between the two groups (Table 4 and 5). Table 3 depicts therapist characteristics. Most therapists were licensed in CBT (82.68%),

followed by psychodynamic therapy (19.55%). On average, therapists included $M = 6.23$ patients ($SD = 4.60$; range: 1-32).

Table 2 Sociodemographic Characteristics of Patients at Baseline

	BC (<i>n</i> = 583)		TAU (<i>n</i> = 576)		Total (<i>n</i> = 1159)		Statistical test of group difference		
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>df</i>	X^2_a	<i>p</i>
Gender									0.20 ^b
Female	400	68.61	417	72.40	817	70.49			
Male	172	29.50	155	26.91	327	28.21			
Non-binary	4	0.68	2	0.35	6	0.52			
Gender-fluid	4	0.68	0	0.00	4	0.35			
Other	1	0.17	0	0.00	1	0.09			
No answer	2	0.34	2	0.35	4	0.35			
Age									
Mean (SD)	35.82 (12.57)		37.43 (12.96)		36.62 (12.79)		1154	-2.15	0.03* ^c
Range	18-89		18-83		18-89				
Relationship							1	2.38	0.12
Single	205	35.16	177	30.73	382	32.69			
In a relationship	378	64.84	399	69.27	777	67.04			
Highest education									<0.01** ^b
No degree	2	0.34	2	0.35	4	0.35			1.00 ^b
Nine years of education	40	6.86	25	4.34	65	5.61			0.07 ^b
Secondary school certificate	79	13.55	107	18.58	186	16.05			0.02* ^b
College entrance qualification	322	55.23	268	46.53	590	50.91			<0.01** ^b
Vocational education	136	23.33	162	28.13	298	25.71			0.07 ^b
Other	4	0.67	12	2.08	16	1.38			0.05* ^b
Employment status							5	7.61	0.18
Employed	333	57.12	350	60.76	683	58.93			
Self-employed	22	3.77	36	6.25	58	5.00			
Unemployed	51	8.75	38	6.60	89	7.68			
Pensioned	24	4.12	21	3.65	45	3.88			
In education	108	18.52	91	15.80	199	17.17			

[illegible]

	BC (<i>n</i> = 583)		TAU (<i>n</i> = 576)		Total (<i>n</i> = 1159)		Statistical test of group difference			
Mean number of minutes (SD)	23.93 (44.64)		22.28 (28.30)		23.11 (37.42)		1157	0.75	0.45 ^c	
Mean treatment preference ^g										
For BC (SD)	7.47 (2.40)		7.49 (2.52)		7.48 (2.46)		1157	-0.10	0.92 ^c	
For TAU (SD)	6.20 (2.86)		6.10 (2.94)		6.15 (2.90)		1157	0.61	0.54 ^c	
Primary diagnosis ^e										
Mental and behavioral disorders due to substance use	8	1.37	5	0.87	13	1.12	1	0.29	0.59	
Schizophrenia, schizotypal and delusional disorders	3	0.51	3	0.52	6	0.52				1.00 ^b
Affective disorders	311	53.34	305	52.95	616	53.15	1	0.01	0.94	
Anxiety disorders	134	22.98	138	23.96	272	23.47	1	0.10	0.75	
Obsessive- compulsive disorder	15	2.57	14	2.43	29	2.50	1	<0.01	1.00	
Acute stress reaction and post-traumatic stress disorder	55	9.43	33	5.73	88	7.59	1	5.15	0.02 [*]	
Adjustment disorder	93	15.95	98	17.01	191	16.48	1	0.17	0.68	
Somatoform disorders	30	5.15	31	5.38	61	5.26	1	0.02	0.96	
Eating disorders	24	4.12	17	2.95	41	3.54	1	0.84	0.36	
Personality disorders	26	4.46	30	5.21	56	4.83	1	0.21	0.64	
Hyperkinetic and attention-deficit disorders	15	2.57	12	2.08	27	2.33	1	0.13	0.72	
Other disorders ^h	21	3.64	25	4.39	46	4.01	1	0.24	0.62	
Missing diagnosis	8	1.37	6	1.04	14	1.21	1	0.06	0.81	

	BC (<i>n</i> = 583)		TAU (<i>n</i> = 576)		Total (<i>n</i> = 1159)		Statistical test of group difference		
At least one comorbid diagnosis	333	57.12	307	53.30	640	55.22	1	1.56	0.21
Family history of similar mental health issues	301	51.63	290	50.35	591	50.99	1	0.14	0.71

^a = For samples with *n* < 40 Yates continuity correction was applied

^b = Fisher's Exact Test

^c = t-test for independent samples

^d = Based on the German version of the MacArthur Scale (SSS-D)

^e = Multiple answers possible

^f = Unplausible values were removed

^g = On a scale from 0 (not at all) to 10 (very much), patients were asked how much both treatment options (BC vs. TAU) appealed to them.

^h = Dissociative disorders, other and unspecified reactions to severe stress, other neurotic disorders, sexual dysfunction, enduring personality change after catastrophic experience, Habit and impulse disorders, pervasive developmental disorders, emotional disorders with onset specific to childhood, nonorganic sleep disorders

* = *p* < 0.05

** = *p* < 0.01

Table 3 Sociodemographic Characteristics of Therapists at Baseline

Demographic variable	<i>n</i> = 178 (%)
Gender	
Female	138 (77.53)
Male	38 (21.35)
No answer	2 (1.12)
Age ^a	
Mean (SD)	45.54 (9.42)
Range	27-74
Years of professional experience	
Mean (SD)	12.03 (8.16)
Range	1-44
Licensed in ^b	
Cognitive-behavioral therapy	148 (82.68)
Psychodynamic therapy	35 (19.55)
Psychoanalytic therapy	6 (3.35)
Systemic therapy	7 (3.91)
More than one approach	11 (6.14)
Application of ^b	
Cognitive-behavioral therapy	147 (82.12)
Psychodynamic therapy	37 (20.67)
Psychoanalytic therapy	5 (2.79)
Systemic therapy	22 (12.29)
More than one approach	30 (16.76)
Location of practice	
Rural area	60 (33.70)

Demographic variable	<i>n</i> = 178 (%)
Medium-sized city	61 (34.27)
Large city	57 (32.02)

Note. Demographic characteristics are available for *n* = 178 therapists. *N* = 179 therapists recruited patients and are included in later analysis.

^a = Unplausible values were removed

^b = Multiple answers possible

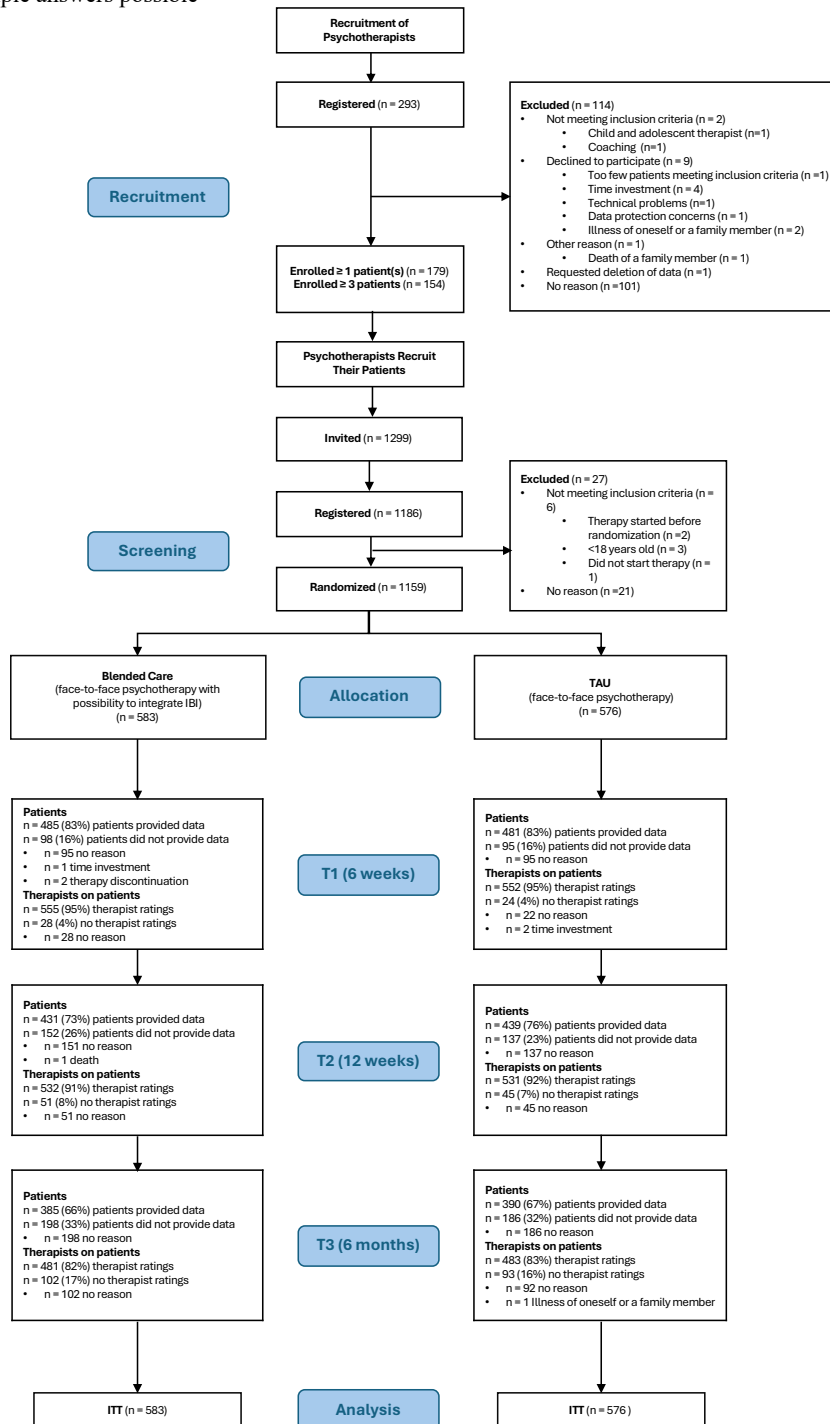


Figure 1. CONSORT flow chart.

Intervention Effects on Primary and Secondary Outcomes

Mean scores and standard errors for all primary and secondary outcome measures are depicted in Table 4. Findings from the linear mixed models and Cohen's d effect sizes are presented in Table 5. Time was a significant predictor of all outcomes. There was no significant difference between BC or PT nor any indication that the two groups changed differently over time in primary or secondary outcomes (see Table 5). Between-group effect sizes at 6-months post-randomization were negligibly small (Cohen's d ranging from -0.06 to 0.06) and non-significant.

Table 4: Estimated Means and Standard Errors for Blended Care and Psychotherapy by Measurement Points based on Multiply Imputed Data

Outcome	Group	Baseline		6 weeks		12 weeks		6 months	
		M	SE	M	SE	M	SE	M	SE
Mental Distress (composite)	BC	23.55	0.7	19.66	0.71	18.43	0.71	17.2	0.72
Mental Distress (composite)	TAU	23.61	0.7	19.53	0.7	18.15	0.7	16.78	0.7
Anxiety (GAD-7)	BC	10.82	0.37	9.02	0.37	8.48	0.38	7.87	0.38
Anxiety (GAD-7)	TAU	10.88	0.36	9.08	0.36	8.42	0.37	7.73	0.37
Depression (PHQ-8)	BC	12.73	0.4	10.64	0.4	9.94	0.41	9.33	0.41
Depression (PHQ-8)	TAU	12.73	0.4	10.45	0.4	9.73	0.4	9.05	0.4
Satisfaction with Life (SWLS)	BC	18.36	0.51	18.43	0.53	19.57	0.53	19.57	0.53
Satisfaction with Life (SWLS)	TAU	18.65	0.5	18.55	0.51	19.9	0.51	20.06	0.52
Level of Functioning (AQoL-8D)	BC	92.32	1.48	89.55	1.5	87.73	1.49	86.03	1.5

		Baseline		6 weeks		12 weeks		6 months	
Level of Functioning (AQoL-8D)	TAU	92.26	1.47	89.38	1.47	86.51	1.47	84.78	1.48
Coping Subscale (AQoL-8D)	BC	9.08	0.18	8.82	0.18	8.53	0.18	8.27	0.19
Coping Subscale (AQoL-8D)	TAU	8.98	0.18	8.65	0.18	8.28	0.18	8.16	0.18
Happiness Subscale (AQoL-8D)	BC	11.71	0.23	11.5	0.23	11.25	0.23	11.02	0.23
Happiness Subscale (AQoL-8D)	TAU	11.61	0.22	11.45	0.22	11.12	0.23	10.88	0.23
Independent Living Subscale (AQoL-8D)	BC	8.67	0.24	8.36	0.24	8.22	0.25	8.09	0.24
Independent Living Subscale (AQoL-8D)	TAU	8.55	0.24	8.28	0.24	8.05	0.24	7.85	0.25
Mental Health Subscale (AQoL-8D)	BC	24.68	0.41	23.53	0.41	22.95	0.41	22.45	0.42
Mental Health Subscale (AQoL-8D)	TAU	24.81	0.4	23.69	0.41	22.72	0.41	22.15	0.41
Therapist-rated Severity (CGI-S)	BC	4.66	0.1	4.45	0.1	4.25	0.1	4.03	0.1
Therapist-rated Severity (CGI-S)	TAU	4.66	0.1	4.45	0.1	4.31	0.1	4.07	0.1
Therapist-rated Change (CGI-C)	BC			3.14	0.07	2.95	0.07	2.62	0.07
Therapist-rated Change (CGI-C)	TAU			3.15	0.07	2.89	0.07	2.66	0.07
Therapist-rated Change due to therapy (CGI-C)	BC			4.17	0.19	4.6	0.19	5.34	0.2
Therapist-rated Change due to therapy (CGI-C)	TAU			4.3	0.19	4.64	0.19	5.18	0.19

Note. Means and SEs as predicted by the linear mixed models, based on the imputed data.

Results are averaged over the levels of education. Therapists did not rate changes at baseline.

AQoL-8D: Assessment of Quality of Life-8D; CGI-C: Clinical Global Impression-Change; CGI-S: Clinical Global Impression-Severity; GAD-7: Generalized Anxiety Disorder-7; PHQ-8: Patient Health Questionnaire-8; SWLS: Satisfaction with Life Scale.

Table 5: Results of the Linear Mixed Models and Between-Group Effect Sizes (Cohen's d) based on Multiply Imputed Data

Outcome	Time	Group	Education	Time x Group		Baseline		6 weeks		12 weeks		6 months	
						d	95%CI	d	95%CI	d	95%CI	d	95%CI
Mental Distress (composite)	F(3,3468) = 276.35, p <0.001	F(1,1151) = 0.24, p = 1.00	F(5,1151) = 4.65, p = 0.033	F(3,3468) = 0.34, p = 1.00	BC vs. TAU	0	-0.08 - 0.07	0.01	-0.07 - 0.09	0.02	-0.06 - 0.1	0.03	-0.05 - 0.11
Anxiety (GAD-7)	F(3,3468) = 179.97, p <0.001	F(1,1151) = 0.01, p = 1.00	F(5,1151) = 2.79, p = 1.00	F(3,3468) = 0.22, p = 1.00	BC vs. TAU	-0.01	-0.08 - 0.07	-0.01	-0.08 - 0.07	0.01	-0.07 - 0.08	0.02	-0.06 - 0.09
Depression (PHQ-8)	F(3,3468) = 235.55, p <0.001	F(1,1151) = 0.61, p = 1.00	F(5,1151) = 5.26, p = 0.009	F(3,3468) = 0.38, p = 1.00	BC vs. TAU	0	-0.08 - 0.08	0.03	-0.05 - 0.11	0.03	-0.05 - 0.11	0.04	-0.04 - 0.12
Satisfaction with Life (SWLS)	F(3,3468) = 35.1, p <0.001	F(1,1151) = 1.17, p = 1.00	F(5,1151) = 4.65, p = 0.033	F(3,3468) = 0.43, p = 1.00	BC vs. TAU	-0.04	-0.12 - 0.05	-0.02	-0.1 - 0.07	-0.04	-0.12 - 0.04	-0.06	-0.14 - 0.02
Level of Functioning (AQoL-8D)	F(3,3468) = 109.63, p <0.001	F(1,1151) = 0.65, p = 1.00	F(5,1151) = 8.21, p <0.001	F(3,3468) = 1.29, p = 1.00	BC vs. TAU	0	-0.09 - 0.09	0.01	-0.08 - 0.1	0.06	-0.03 - 0.15	0.06	-0.03 - 0.15
Coping Subscale (AQoL-8D)	F(3,3468) = 63.87, p <0.001	F(1,1151) = 2.45, p = 1.00	F(5,1151) = 2.74, p = 1.00	F(3,3468) = 0.7, p = 1.00	BC vs. TAU	0.03	-0.05 - 0.11	0.06	-0.02 - 0.14	0.08	0 - 0.16	0.03	-0.05 - 0.11
Happiness Subscale (AQoL-8D)	F(3,3468) = 35.9, p <0.001	F(1,1151) = 0.63, p = 1.00	F(5,1151) = 3.05, p = 0.972	F(3,3468) = 0.14, p = 1.00	BC vs. TAU	0.03	-0.05 - 0.11	0.01	-0.07 - 0.09	0.03	-0.05 - 0.11	0.03	-0.05 - 0.12
Independent Living Subscale (AQoL-8D)	F(3,3468) = 29.26, p <0.001	F(1,1151) = 1.28, p = 1.00	F(5,1151) = 9.07, p <0.001	F(3,3468) = 0.43, p = 1.00	BC vs. TAU	0.03	-0.05 - 0.12	0.02	-0.06 - 0.1	0.04	-0.04 - 0.13	0.06	-0.02 - 0.15
Mental Health Subscale (AQoL-8D)	F(3,3468) = 116.84, p <0.001	F(1,1151) = 0.06, p = 1.00	F(5,1151) = 3.5, p = 0.382	F(3,3468) = 1.43, p = 1.00	BC vs. TAU	-0.02	-0.1 - 0.06	-0.02	-0.1 - 0.06	0.03	-0.05 - 0.11	0.04	-0.04 - 0.12
Therapist-rated Severity (CGI-S)	F(3,3468) = 115.28, p <0.001	F(1,1151) = 0.22, p = 1.00	F(5,1151) = 4.16, p = 0.095	F(3,3468) = 0.4, p = 1.00	BC vs. TAU	0	-0.09 - 0.08	0	-0.08 - 0.08	-0.04	-0.12 - 0.05	-0.02	-0.1 - 0.06
Therapist-rated Change (CGI-C)	F(3,3468) = 94.04, p <0.001	F(1,1151) = 0.01, p = 1.00	F(5,1151) = 2.6, p = 1.00	F(3,3468) = 0.68, p = 1.00	BC vs. TAU	-	-	-0.01	-0.07 - 0.06	0.03	-0.03 - 0.1	-0.02	-0.08 - 0.04
Therapist-rated Change due to therapy (CGI-C)	F(3,3468) = 40.5, p <0.001	F(1,1151) = 0.01, p = 1.00	F(5,1151) = 1.75, p = 1.00	F(3,3468) = 0.77, p = 1.00	BC vs. TAU	-	-	-0.03	-0.09 - 0.04	-0.01	-0.07 - 0.05	0.03	-0.03 - 0.1

Note: Therapists did not rate changes at baseline. AQoL-8D: Assessment of Quality of Life-8D; CGI-C: Clinical Global Impression-Change; CGI-S: Clinical Global Impression-Severity; GAD-7: Generalized Anxiety Disorder-7; PHQ-8: Patient Health Questionnaire-8; SWLS: Satisfaction with Life Scale.

Therapy Duration and Satisfaction with Treatment

Patients receiving BC had on average, 16.6 ($SD = 8.75$) face-to-face sessions and patients receiving conventional PT had 16.4 ($SD = 7.03$) sessions in six months. There was no significant group difference in the number of sessions ($t(914.29) = 0.57, p = 0.566$). For 25.26% ($n = 147$) patients in BC and 26.22% ($n = 151$) patients in PT, therapists indicated that treatment was terminated during the six-month study period. In BC, for 67.35 % ($n = 99$) patients it ended earlier than planned and for 32.65% ($n = 48$) it was terminated as planned. In conventional PT, it ended earlier than planned for 59.60 % ($n = 90$) of patients and for 40.40 % ($n = 61$) it ended as planned. There was no significant difference in termination rates between study groups ($\chi^2(1) = 1.61, p = 0.205$). Treatment satisfaction was high in both groups, with a mean CSQ score (maximum of 32 points) of 27.2 ($SD = 3.98$) in BC and 26.9 ($SD = 4.38$) in routine PT. There was no significant difference in overall satisfaction with treatment between groups ($t(755.73) = 0.88, p = 0.379$).

Reliable Change

The percentage of patients achieving GAD-7 reliable change were comparable for both groups (BC: 41.9% - 45.2%; PT: 44.1% - 47.6% across imputations) as well as for the PHQ-8 (BC: 31.6% - 34.2%; PT: 30.9% - 33.9%).

Platform Usage

The platform asks therapists to assign modules or chapters within modules to their patients. In addition, patients can request their therapist to assign a module. Therapists assigned patients, on average, $M = 7.26$ ($SD = 7.01$) out of 39 online chapters in TONI. 534 patients (91.6%) received at least one chapter, while 273 (46.83%) of patients requested at least one

module. They requested $M = 1.44$ ($SD = 2.3$) modules on average. Figure S6 in the supplementary materials shows frequencies of assigned online modules (operationalized as receiving at least one chapter). The most frequently assigned modules were “Development” and “Values & Goals” with 227 patients (38.94%) assigned respectively. Therapists assigned the Collaboration module least frequently (33 patients (5.66%)). $N = 174$ patients (29.85%) pinned at least one page, with $M = 1.36$ ($SD = 4.5$) of pinned pages. $N = 152$ patients (26.07%) had at least one entry in any tracker. Usage of trackers varied widely ($M = 11.23$, $SD = 85.11$, with a range from 0 to 1907). $N = 73$ patients (12.52%) completed at least one diary entry. Regarding overall platform usage, patients logged in $M = 19.73$ ($SD = 24.66$) times and spent $M = 367.14$ ($SD = 338.27$) minutes on the platform. 117 therapists enabled messaging. Patients who had messaging enabled sent $M = 0.42$ ($SD = 1.2$) messages to their therapist and received $M = 0.94$ ($SD = 1.56$) messages from their therapist.

Discussion

This study investigated the effects of BC in a naturalistic setting and compared its effects to routine PT. Contrary to our hypotheses, we did not find differences between BC and PT in outcomes, including anxiety, depression, satisfaction with life, level of functioning, therapist-rated severity / changes and satisfaction with treatment. We also did not find a reduced number of therapeutic sessions or differences in therapy termination.

We hypothesized that integrating TONI, a transdiagnostic and transtheoretical modular digital intervention, would enhance the effects of PT in routine outpatient care. This assumption was grounded on (1) initial evidence suggesting that BC can improve PT effectiveness, as indicated by the first systematic review on the topic (Erbe et al., 2017), (2) the ability for

therapists to outsource certain materials, such as psychoeducation, to TONI while adapting face-to-face sessions flexibly to patient needs, (3) the potentially higher therapy dosage and (4) the empowerment patients may receive by working on additional digital material between sessions.

Our findings are in line with more recent studies on BC for depression and anxiety, suggesting comparable clinical effects between BC and PT (Bielinski et al., 2023; Kemmeren et al., 2023; Kooistra et al., 2019; Mathiasen et al., 2022; Nakao et al., 2018; Romijn et al., 2021). Studies indicating small to moderate between-group effect sizes in favor of BC, offered the online component as an adjunctive tool to usual care rather than an integrated blended treatment protocol (Berger et al., 2018; Schuster et al., 2020). While integrated BC setups usually involve alternating sessions with online components and thus reduce the number of sessions while maintaining effectiveness (Kemmeren et al., 2023), we did not find any indication that BC reduces the number of sessions. In routine care in Germany, therapists have little incentive to reduce the number of sessions, as their compensation is session-based. They thus might not have implemented BC to the full extent, resulting in a very comparable dosage of treatment across both study groups.

With TONI, we encouraged a personalized and flexible use of online modules. Usage data showed that the first three modules (Values & Goals, Development, Mindfulness) were assigned most often. This could reflect a particularly good fit for this content and the digital format as well as a broad applicability across patients and therapeutic schools (Behr et al., 2024). It could, however, also reflect an inability or reluctance of psychotherapists to make individual choices for their patients, with the first three modules being a safe and convenient choice for all of their patients. In BC, as an emerging field, nearly all therapists are novices with no to very limited experience in how to integrate digital tools in their PT. Psychotherapists repeatedly

identified this lack of knowledge and confidence as a barrier to implementing BC in qualitative studies (Dech et al., 2022). Their high caseload and the limited scope of the study (on average, therapists enrolled six patients in the study which were then randomized to the two groups) may also discourage transformative changes to their practice. The limited research on BC seems to show more robust results for additive designs, where the online material works independently of the face-to-face therapists, compared to integrative designs. While qualification and confidence in using digital tools will likely increase in the future, it still seems plausible that not all therapists want to actively engage with online materials. Future studies should investigate additive vs. integrative BC setups. Future studies may also focus on investigating the rationale and motivation of therapists to include online material in their therapy.

Strengths and Limitations

We developed and tested a flexible, modular, and transtheoretical BC tool that allows psychotherapists of all orientations to integrate digital components across the whole spectrum of mental health. Main strengths of the trial are the naturalistic setting as well as the transdiagnostic application and evaluation. As such it offers evidence for implementation in practice for a broad clientele. Other strengths include the trial's sufficient power to detect small differences, statistical analyses based on masked data, the use of self- as well as therapist-rated measures, and the open science principles adopted, including sharing data and code. The integration of all therapeutic orientations as well as all diagnoses mirrors routine practice and increases external validity. However, there are still limitations to the generalizability of our findings: Despite extensive efforts to recruit therapists from diverse orientations, most therapists were CBT therapists. The largest proportion of patients was diagnosed by their therapist with an affective disorder, followed by anxiety and adjustment disorders. The study team did not conduct

additional clinical interviews to corroborate diagnoses, and the diagnostic status at 6-month randomization is unknown. Psychotherapists were responsible for recruiting their patients, which might have constrained the scope of patient recruitment. While qualitative interviews with psychotherapists in our study indicated that they did not restrict patient inclusion into the trial (Jordan et al., 2024) this approach may still limit generalizability. Randomization to treatment allowed for a causal interpretation of the results. We found some small imbalances between groups in baseline demographic variables, which we accounted for by subsequently controlling for these variables in the analyses. To deal with systematic dropout, we collected a large number of possible variables that may explain the missing data mechanism and controlled for it via multiple imputation.

Conclusion

BC holds the potential to increase therapy effectiveness and reduce therapist burden. Our findings suggest that the benefits observed in more controlled BC settings may not fully translate to routine care, potentially due to variations in implementation and adherence.

Conflict of Interest Statement: The authors have no conflict of interest to declare.

Funding Sources: This study was funded by the Innovation Committee (Innovationsausschuss) of the Joint Federal Committee (Gemeinsamer Bundesausschuss, GBA, no: 01VSF19041). The funders had no role in the study design, collection, analysis, and interpretation of data, writing of the manuscript, or the decision to submit the manuscript for publication.

Author Contributions: Authors contributions according to the CRediT taxonomy:

Conceptualization: CS, SB, FF, LH, MP, SP, CK, JB; Data Curation: SB, FF, LH, MP; Formal Analysis: CS, AM; Funding Acquisition: SP, CK, JB; Investigation: CS, SB, FF, LH, MP, CK, JB; Methodology: CS, SB, FF, LH, MP, SP, CK, JB; Project Administration: CS, CK, JB; Resources: n/a; Software: LH; Supervision: CK, JB, SP; Validation: CS, AM; Visualization: CS, AM; Writing – Original Draft: CS, SB, FF, LH, MP, JB; Writing – Review & Editing: CS, AM, SB, FF, LH, MP, SP, CK, JB

Data transparency: This manuscript represents the main outcome paper of the RCT. Secondary analyses of this study population but relating to other research questions and measures are in preparation but have not been submitted or published yet. We also conducted several interview studies with subsamples of patients and psychotherapists. One qualitative interview study for which we have interviewed psychotherapists about their BC inclusion process to gain insights into potential inclusion or exclusion criteria for BC has been submitted elsewhere.

References

- Arias, D., Saxena, S., & Verguet, S. (2022). Quantifying the global burden of mental disorders and their economic value. *eClinicalMedicine*, 54, 101675.
<https://doi.org/10.1016/j.eclinm.2022.101675>
- Attkisson, C. C., & Zwick, R. (1982). The client satisfaction questionnaire. *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)
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv:1406.5823 [Stat]*. <http://arxiv.org/abs/1406.5823>
- Behr, S., Fenski, F., Boettcher, J., Knaevelsrud, C., Hammelrath, L., Kovacs, G., Schirmer, W., Petrick, H., Becker, P., & Schaeuffele, C. (2024). TONI - One for all? Participatory development of a transtheoretic and transdiagnostic online intervention for blended care. *Internet Interventions*, 35, 100723. <https://doi.org/10.1016/j.invent.2024.100723>
- Bennett, L., Melchers, B., & Proppe, B. (2020). *Curta: A General-purpose High-Performance Computer at ZEDAT, Freie Universität Berlin* (p. 5 S.). Freie Universität Berlin.
<https://doi.org/10.17169/REFUBIUM-26754>
- Berger, T., Krieger, T., Sude, K., Meyer, B., & Maercker, A. (2018). Evaluating an e-mental health program (“deprexis”) as adjunctive treatment tool in psychotherapy for depression: Results of a pragmatic randomized controlled trial. *Journal of Affective Disorders*, 227, 455–462. <https://doi.org/10.1016/j.jad.2017.11.021>
- Berk, M., Ng, F., Dodd, S., Callaly, T., Campbell, S., Bernardo, M., & Trauer, T. (2008). The validity of the CGI severity and improvement scales as measures of clinical effectiveness

- suitable for routine clinical use. *Journal of Evaluation in Clinical Practice*, 14(6), 979–983. <https://doi.org/10.1111/j.1365-2753.2007.00921.x>
- Bielinski, L. L., Krieger, T., Kley, M. A., Moggi, F., & Berger, T. (2023). Blending an internet-based emotion regulation intervention with face-to-face psychotherapy: Findings from a pilot randomized controlled trial. *Internet Interventions*, 33, 100650. <https://doi.org/10.1016/j.invent.2023.100650>
- Bundespsychotherapeutenkammer. (2011). *BPtK-Studie zu Wartezeiten in der ambulanten psychotherapeutischen Versorgung*. Abrufbar unter . https://www.bptk.de/uploads/media/110622_BPtK-Studie_Langfassung_Wartezeiten-in-der-Psychotherapie_01.pdf
- Dech, M., Klawohn, J., & Romanczuk-Seiferth, N. (2022). Das Beste aus zwei Welten: Eine systematische Übersicht zu Faktoren der Implementierung einer “Blended Therapy” (Gemischte Therapie) in der Psychotherapeutischen Routineversorgung. *Verhaltenstherapie*, 1–12. <https://doi.org/10.1159/000524332>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13
- Erbe, D., Eichert, H.-C., Riper, H., & Ebert, D. D. (2017). Blending Face-to-Face and Internet-Based Interventions for the Treatment of Mental Disorders in Adults: Systematic Review. *Journal of Medical Internet Research*, 19(9), e306. <https://doi.org/10.2196/jmir.6588>
- Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the satisfaction with life scale (SWLS). *European Journal of Psychological Assessment*, 27(2), 127–132. <https://doi.org/10.1027/1015-5759/a000058>

- Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem “Gesundheitsfragebogen für Patienten (PHQ-D)“. *Diagnostica*, 50(4), 171–181.
<https://doi.org/10.1026/0012-1924.50.4.171>
- Heinig, I., Wittchen, H.-U., & Knappe, S. (2021). Help-Seeking Behavior and Treatment Barriers in Anxiety Disorders: Results from a Representative German Community Survey. *Community Mental Health Journal*, 57(8), 1505–1517.
<https://doi.org/10.1007/s10597-020-00767-5>
- Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., Busch, M. A., Hapke, U., Maske, U., Seiffert, I., Gaebel, W., Maier, W., Wagner, M., Zielasek, J., & Wittchen, H.-U. (2014). Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: The Mental Health Module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, 23(3), 304–319. <https://doi.org/10.1002/mpr.1439>
- Jordan, S., Becker, P., Behr, S., Fenski, F., Knaevelsrud, C., Boettcher, J., & Schaeuffele, C. (2024). “Who blends in and why (not)?” A qualitative study on psychotherapists’ patient inclusion in blended care. *PsyArXiv*. <https://doi.org/10.31234/osf.io/nvjhd>
- Kalde, J., Atik, E., Stricker, J., Schückes, M., Neudeck, P., Pittig, A., & Pietrowsky, R. (2023). Enhancing the effectiveness of CBT for patients with unipolar depression by integrating digital interventions into treatment: A pilot randomized controlled trial. *Psychotherapy Research*, 0(0), 1–16. <https://doi.org/10.1080/10503307.2023.2277866>
- Kemmeren, L. L., van Schaik, A., Draisma, S., Kleiboer, A., Riper, H., & Smit, J. H. (2023). Effectiveness of Blended Cognitive Behavioral Therapy Versus Treatment as Usual for Depression in Routine Specialized Mental Healthcare: E-COMPARED Trial in the

- Netherlands. *Cognitive Therapy and Research*, 47(3), 386–398.
<https://doi.org/10.1007/s10608-023-10363-y>
- Kooistra, L. C., Wiersma, J. E., Ruwaard, J., Neijenhuijs, K., Lokkerbol, J., van Oppen, P., Smit, F., & Riper, H. (2019). Cost and Effectiveness of Blended Versus Standard Cognitive Behavioral Therapy for Outpatients With Depression in Routine Specialized Mental Health Care: Pilot Randomized Controlled Trial. *Journal of Medical Internet Research*, 21(10), e14261. <https://doi.org/10.2196/14261>
- Kotz, S., & Nadarajah, S. (2004). *Multivariate t-distributions and their applications*. Cambridge university press. https://books.google.com/books?hl=en&lr=&id=dmxtU-TxTi4C&oi=fnd&pg=IA1&dq=Kotz+%26+Nadarajah,+2004+t+distribution&ots=kzwNH4iuZE&sig=2ATLve9NZAPxPMDYNO0Y6_fnz6U
- Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, 32(9), 509–515. <https://doi.org/10.3928/0048-5713-20020901-06>
- Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114(1), 163–173. <https://doi.org/10.1016/j.jad.2008.06.026>
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274.
<https://doi.org/10.1097/MLR.0b013e318160d093>
- Mathiasen, K., Andersen, T. E., Lichtenstein, M. B., Ehlers, L. H., Riper, H., Kleiboer, A., & Roessler, K. K. (2022). The Clinical Effectiveness of Blended Cognitive Behavioral

- Therapy Compared With Face-to-Face Cognitive Behavioral Therapy for Adult Depression: Randomized Controlled Noninferiority Trial. *Journal of Medical Internet Research*, 24(9), e36577. <https://doi.org/10.2196/36577>
- Nakao, S., Nakagawa, A., Oguchi, Y., Mitsuda, D., Kato, N., Nakagawa, Y., Tamura, N., Kudo, Y., Abe, T., Hiyama, M., Iwashita, S., Ono, Y., & Mimura, M. (2018). Web-Based Cognitive Behavioral Therapy Blended With Face-to-Face Sessions for Major Depression: Randomized Controlled Trial. *Journal of Medical Internet Research*, 20(9), e10743. <https://doi.org/10.2196/10743>
- Porter, A., Franklin, M., De Vocht, F., d'Apice, K., Curtin, E., Albers, P., & Kidger, J. (2024). Estimating the effectiveness of an enhanced 'Improving Access to Psychological Therapies' (IAPT) service addressing the wider determinants of mental health: A real-world evaluation. *BMJ Open*, 14(1), e077220. <https://doi.org/10.1136/bmjopen-2023-077220>
- R Core Team. (2013). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Richardson, J., Iezzi, A., Khan, M. A., & Maxwell, A. (2014). Validity and Reliability of the Assessment of Quality of Life (AQoL)-8D Multi-Attribute Utility Instrument. *The Patient - Patient-Centered Outcomes Research*, 7(1), 85–96. <https://doi.org/10.1007/s40271-013-0036-x>
- Romijn, G., Batelaan, N., Koning, J., Van Balkom, A., De Leeuw, A., Benning, F., Hakkaart van Roijen, L., & Riper, H. (2021). Acceptability, effectiveness and cost-effectiveness of blended cognitive-behavioural therapy (bCBT) versus face-to-face CBT (ftfCBT) for

- anxiety disorders in specialised mental health care: A 15-week randomised controlled trial with 1-year follow-up. *PLoS One*, 16(11), e0259493.
- Schaeuffele, C., Heinrich, M., Behr, S., Fenski, F., Hammelrath, L., Zagorscak, P., Jansen, A., Pohl, S., Boettcher, J., & Knaevelsrud, C. (2022). Increasing the effectiveness of psychotherapy in routine care through blended therapy with transdiagnostic online modules (PsyTOM): Study protocol for a randomized controlled trial. *Trials*, 23(1), 830. <https://doi.org/10.1186/s13063-022-06757-0>
- Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., Kopta, M., Schulte, D., Saxon, D., Nielsen, S. L., & Lambert, M. J. (2017). Reliability of Therapist Effects in Practice-Based Psychotherapy Research: A Guide for the Planning of Future Studies. *Administration and Policy in Mental Health and Mental Health Services Research*, 44(5), 598–613. <https://doi.org/10.1007/s10488-016-0736-3>
- Schmidt, J., Lamprecht, F., & Wittmann, W. W. (1989). Zufriedenheit mit der stationären Versorgung. Entwicklung eines Fragebogens und erste Validitätsuntersuchungen. *Psychotherapie Psychosomatik Medizinische Psychologie*, 39(7), 248–255.
- Schuster, R., Laireiter, A.-R., Berger, T., Moritz, S., Meyer, B., Hohagen, F., & Klein, J. P. (2020). Immediate and long-term effectiveness of adding an Internet intervention for depression to routine outpatient psychotherapy: Subgroup analysis of the EVIDENT trial. *Journal of Affective Disorders*, 274, 643–651. <https://doi.org/10.1016/j.jad.2020.05.122>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092. <https://doi.org/10.1001/archinte.166.10.1092>

- Titzler, I., Saruhanjan, K., Berking, M., Riper, H., & Ebert, D. D. (2018). Barriers and facilitators for the implementation of blended psychotherapy for depression: A qualitative pilot study of therapists' perspective. *Internet Interventions*, *12*, 150–164.
<https://doi.org/10.1016/j.invent.2018.01.002>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67.
- Wu, Y., Levis, B., Riehm, K. E., Saadat, N., Levis, A. W., Azar, M., Rice, D. B., Boruff, J., Cuijpers, P., Gilbody, S., Ioannidis, J. P. A., Kloda, L. A., McMillan, D., Patten, S. B., Shrier, I., Ziegelstein, R. C., Akena, D. H., Arroll, B., Ayalon, L., ... Thombs, B. D. (2020). Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: A systematic review and individual participant data meta-analysis. *Psychological Medicine*, *50*(8), 1368–1380. <https://doi.org/10.1017/S0033291719001314>