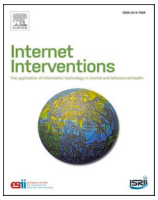
Internet Interventions 35 (2024) 100716

Contents lists available at ScienceDirect

Internet Interventions

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



Evaluation of an online-based self-help program for patients with generalized anxiety disorder - A randomized controlled trial

J. Rubel [a,1], T. Väth [a,1], S. Hanraths [a,1], L. Pruessner [b,2], C. Timm [b,2], S. Hartmann [b,2], S. Barnow [b,2], C. Lalk [a,*,1]

[a] Department of Humanities, Institute of Psychology, Osnabrück University, Germany
[b] Department of Psychology, Heidelberg University, Germany

ARTICLE INFO

Keywords:
Randomized controlled trial
Internet-based interventions
Generalized anxiety disorder

ABSTRACT

Objectives: This study aimed to evaluate the effects of an online self-help intervention for generalized anxiety disorder (GAD). Our primary outcomes were generalized anxiety symptoms, measured using the Generalized Anxiety Disorder – 7 (GAD-7; Spitzer et al., 2006), and wellbeing based on the World Health Organization Wellbeing Index - 5 (WHO-5; Topp et al., 2015).
Methods: A total of 156 German-speaking patients aged 18 to 65 with a diagnosis of GAD and internet access were included in this randomized controlled trial. The intervention group ($N = 78$) received access to a 12-week online self-help program, while the waitlist control group ($N = 78$) received access after the 12-week waiting period.
Results: The intervention group showed a significant improvement in generalized anxiety symptoms compared to the control group ($t(df = 123.73) = 4.52$, $p < .001$) with a large effect size ($d = 0.88$, 95 %-CI: 0.50; 1.26). Additionally, the intervention group demonstrated a significant increase in wellbeing compared to the control group ($t(df = 87{,}86) = 3.48$, $p < .001$), with a moderate effect size ($d = 0.62$, 95 % CI: 0.27; 0.98). However, no significant effects were observed for secondary outcomes of functional impairments, work productivity, mental health literacy, and healthcare demands. For exploratory outcomes, improvement was found for anxiety and worry symptoms.
Conclusions: These findings suggest that an online-based self-help intervention effectively reduces GAD symptoms and improves overall wellbeing. Future research should explore the long-term effects of this intervention and investigate potential mechanisms underlying its efficacy.
Public health implications: Online-based self-help programs provide a promising treatment option for individuals with GAD who face barriers to traditional face-to-face therapy.

1. Introduction

Generalized anxiety disorder (GAD) is a common mental health condition marked by excessive and uncontrollable worry about past, present, or future (DSM-5; American Psychiatric Association, 2013). GAD causes distress and impairment in daily live (Judd et al., 1998), often co-occurring with depressive disorders (Tyrer and Baldwin, 2006). GAD's economic costs are a challenge to healthcare systems (Bereza et al., 2009; Grant et al., 2004), with an estimated lifetime prevalence of 1.6 %–10.5 % (Preti et al., 2021). Hoffman et al. (2008) reviewed three studies assessing economic costs of GAD based on healthcare use and absenteeism. The largest cost factors were outpatient treatment (42 % of costs), sick leave (33 %), and hospitalization (21 %). GAD patients with a comorbidity, most often depression (26 %), substance use disorders (24 %), and gastroenterological symptoms (14.6 %), had almost twice of the costs than patients without a comorbidity. Altogether, the costs of GAD were estimated to be significant cost drivers due to referrals to family doctors and specialists on the grounds of unclear somatic

* Corresponding author at: Lise-Meitner-Str. 3, 49076 Osnabrück, Germany.
E-mail addresses: julian.rubel@uni-osnabrueck.de (J. Rubel), luise.pruessner@psychologie.uni-heidelberg.de (L. Pruessner), christina.timm@psychologie.uni-heidelberg.de (C. Timm), steffen.hartmann@psychologie.uni-heidelberg.de (S. Hartmann), sven.barnow@psychologie.uni-heidelberg.de (S. Barnow), Christopher.lalk@uni-osnabrueck.de (C. Lalk).
[1] Address: Lise-Meitner-Straße 3, 9076 Osnabrück, Germany.
[2] Address: Hauptstraße 49, 69117 Heidelberg, Germany.

https://doi.org/10.1016/j.invent.2024.100716
Received 15 September 2023; Received in revised form 7 December 2023; Accepted 17 January 2024
Available online 23 January 2024

illnesses. This may be a result of misdiagnosing and mistreating GAD as a somatic illness (Hoffman et al., 2008).

GAD is treated with pharmacotherapy or psychotherapy (Carl et al., 2020; Portman et al., 2012) with lower response rates than other anxiety disorders, likely due to delayed treatment (Becker and Hoyer, 2005), personality disorder comorbidity (Friborg et al., 2013), and misdiagnosis and mistreatment (Mowbray et al., 2006). Recently, technology-based solutions provide new treatment approaches (Aguilera, 2015). Cognitive behavioral therapy (CBT) is suitable for adaptation into an online intervention due to its structured nature and its emphasis on psychoeducation and homework assignments (Berger et al., 2011). Internet-based CBT (iCBT) effectively reduces GAD symptoms ($g = 0.70–0.79$) compared to control (Andrews et al., 2018; Eilert et al., 2021). Andrews et al. (2018) conducted at meta-analysis with 9 trials on GAD and over 1100 subjects comparing iCBT with support to a waitlist-control condition with a large effect size for iCBT. More recently, Eilert et al. (2021) conducted another meta-analysis of iCBT with a waitlist-control condition. Again, with the exception of two trials, iCBT was conducted with additional support via email, telephone, or other means. They found large effect sizes for GAD severity and worry, and moderate to large effects for depression, functional impairments, and quality of life.

Mewton et al. (2014) conducted a review of 37 RCTs to assess the effects of iCBT with minimal guidance (e.g., email support) versus face-to-face CBT and found no differential effect. However, therapists spent 13 times as much time for each patient in the face-to-face CBT condition. Also, they found no indication that level of therapist guidance (e.g., face-to-face meetings, phone or email contact, or no support) was predictive of iCBT effects. This is particularly relevant considering the rising healthcare costs associated with GAD treatment (Hoffman et al., 2008) and the prolonged waiting times for outpatient therapy in Germany, which have been further exacerbated by the COVID-19 pandemic (Bundespsychotherapeutenkammer, 2018, 2021).

Notably, no differences in symptom reduction, completion rates, and user satisfaction emerged in a large RCT ($n = 338$) between guided and unguided iCBT for GAD (Dear et al., 2015). Furthermore, specific aspects of guidance, such as the frequency of supportive monitoring or the experience level of support personnel, may not significantly impact GAD symptom reduction (Hadjistavropoulos et al., 2017; Robinson et al., 2010). Taken together, minimally guided online interventions offer low-threshold access to mental healthcare. However, most studies on iCBT for GAD still involve high levels of guidance (e.g., support by therapist or researcher via email, telephone, or text messaging) (Eilert et al., 2021). Therefore, there is a need to strengthen the evidence-base for interventions without therapist support.

In addition to their effects on symptom reduction, iCBT programs have demonstrated positive outcomes in terms of quality of life and functioning (Eilert et al., 2021). However, effects on wellbeing, mental health literacy and healthcare usage are not yet thoroughly investigated: In contrast to quality of life, wellbeing is a more general and more subjective assessment of life satisfaction (Skevington and Böhnke, 2018). It encompasses cognitive and emotional reactions to one's life, while quality of life is measured as life satisfaction across several areas of life. Yet, in a combined model of quality of life and wellbeing, positive emotions are central (Skevington and Böhnke, 2018), underscoring the importance of wellbeing. It has utility across different study domains (e.g., suicidology and pain; Topp et al., 2015) and is positively associated with longevity (Martín-María et al., 2017). Further, wellbeing allows individuals to explore and develop new resources (Fredrickson, 2004), leading to decreased long-term mental health risk (Burns et al., 2022). Mental health literacy refers to knowledge about mental health and treatment, as well as attitudes that promote recognition and help-seeking (O'Connor and Casey, 2015). Levels are higher among people with mental illness and people with prior treatment (O'Connor and Casey, 2015). Also, it has been shown that lower levels can impede help-seeking behavior (Furnham and Swami, 2018), which is particularly relevant for GAD because of notoriously late treatment-seeking (Becker and Hoyer, 2005). Healthcare usage refers to the scope of specific healthcare services utilized by the patient, allowing for an analysis of the associated costs (Chisholm et al., 2000). Due to restricted financial resources in health care systems around the globe, healthcare usage provides a complementary perspective on treatment effectiveness by focusing on a health economic perspective. It is worth noting that the impact of iCBT on mental health literacy and healthcare burdens has yet to be thoroughly investigated (Eilert et al., 2021).

1.1. Objectives

We aim to assess effects of an iCBT intervention for GAD (*Selfapy*) in conditions comparable to routine care. The outcomes symptom reduction, well-being, functioning, mental health literacy, and healthcare burdens will be analyzed. For this purpose, we conducted a randomized controlled trial comparing the self-help intervention group (IG) with a waitlist control group (CG).

1.2. Hypotheses

The primary hypotheses focused on disorder-specific symptoms and well-being. The secondary hypotheses focused on functioning in daily life and work ability, mental health literacy, and healthcare burden. Mental health literacy has been selected since lack thereof could impede treatment-seeking behavior.

1.2.1. Main hypotheses

The primary hypotheses were compared against the control group (CG) after 12 weeks:

a. A greater reduction in Generalized Anxiety Symptomatology than in the CG.
b. A greater improvement in perceived wellbeing than in the CG.

1.2.2. Secondary hypotheses

a. A greater reduction in difficulties in daily life than in the CG.
b. A greater improvement in working ability than in the CG.
c. A greater improvement in health literacy than in the CG.
d. A greater reduction in the extent of healthcare use than in the CG.

1.2.3. Exploratory hypotheses

a. A greater reduction in anxiety symptoms than in the CG.
b. A greater reduction in depressive symptoms than in the CG
c. A greater reduction in worry symptoms than in the CG.
d. Using the intervention does not lead to more negative effects in the IG compared to the CG.

2. Methods

2.1. Study design

A two-arm randomized controlled trial was conducted as preregistered (https://drks.de/search/en/trial/DRKS00023799) to test the efficacy of the online self-help intervention for GAD (Rubel et al., 2023). Before randomization and to check the inclusion and exclusion criteria, a structured diagnostic interview with the Diagnostic Interview for Mental Disorders-Open Access (DIPS-OA, Margraf et al., 2017a; Margraf et al., 2017b) was conducted via video call with every subject. The DIPS-OA allows reliable diagnoses according to the DSM-5 and ICD 10. Regarding DSM-IV-TR criteria, the DIPS-OA was found to have acceptable interrater (0.78) and retest (0.76) reliability for anxiety disorders (Suppiger et al., 2008). Patients who met the criteria were then randomized into either the CG or the IG in a nonstratified 1:1 fashion.

Patients in the IG had immediate access to the intervention, while the CG were only permitted access after a waiting period of 12 weeks. Interim and final evaluations occurred 6 (T2) and 12 (T3) weeks after randomization. This study and its report adhere to the CONSORT statement (Schulz et al., 2010) by reporting details (e.g., randomization procedure) about the study to ensure methodological rigor and replicability.

2.2. Recruitment

Study recruitment in Germany utilized social media ads, flyers to medical professionals, therapists, and self-help groups, as well as a university newsletter. Interested individuals took a screening questionnaire to assess the inclusion criteria age, internet access, and language skills, and to screen for the diagnostic inclusion and exclusion criteria. For this purpose, we used items from the DIPS-OA and the whole Generalized Anxiety Disorder-7 and Patient-Health-Questionnaire-9 (see outcomes section). Subjects scoring above threshold in the screening questionnaire underwent diagnostic video interviews with the DIPS-OA afterwards. In total, 4361 started screening, with 156 (3.58 %) enrolled in the trial after meeting the criteria.

2.3. Inclusion and exclusion criteria

To clarify the inclusion and exclusion criteria, video call structured interviews were conducted to assess study eligibility via the DIPS-OA. All interviews were conducted by trained interviewers with a master's degree in psychology. Furthermore, they were closely supervised by a certified psychotherapist (CBT).

Eligible subjects 1) were between 18 and 65 years of age, 2) had sufficient knowledge of the German language, 3) had uninterrupted Internet access, 4) provided electronic informed consent to participate in the study, and 5) currently met criteria for a diagnosis of GAD (DSM-5300.02; American Psychiatric Association, 2013).

Subjects were excluded if they met any of the following criteria: 1) past or current diagnosis of bipolar disorder, 2) past or current diagnosis of psychotic disorder, 3) current diagnosis of substance dependence, 4) current diagnosis of severe major depressive episode, 5) acute suicidality. A primary diagnosis of one or more disorders other than GAD, in general, was not an exclusion criterion, as we wanted to represent routine care. Substance dependence, bipolar disorder, or psychotic disorders, however, were selected as exclusion criteria as they are deemed to interfere with the successful implementation of the intervention according to German regulations. Subjects who did not meet our inclusion criteria but presented with a diagnosis were encouraged to seek therapeutic help. For all included patients, an individual emergency plan was developed in case of symptom deterioration or suicidality. Adequate language skills were determined during the initial interview.

2.4. Intervention

Selfapy is an online self-help program for the treatment of GAD. The program is based on evidence-based methods, CBT exercises, and elements from Mindfulness-Based Therapy (e.g., Hoyer et al., 2016; Volz and Stieglitz, 2010). The online course consists of a core course, which includes mandatory and optional exercise content and a subsequent set of in-depth modular areas that are individually selectable. It can be accessed via the web and on mobile devices. The course is divided into 12 modules that can be worked on for a week each. Each module covers a specific topic, such as exposure, mindfulness, or problem-solving training. The modules contain informative texts, videos, audio, and interactive exercises. The first seven modules are considered the main course, while subsequent modules include additional content. An overview of all modules of the intervention can be found in the Online supplementary material (OSM) 51.

The online course was completed independently by the patient. However, as part of the patient safety concept, a psychologist monitored the participant's progress to respond to adverse events such as suicidality. Via an integrated messaging function, the patients could contact a psychologist in the event of safety concerns. Active communication by the psychologist occurred only in case of safety concerns and to maintain patient safety.

Subjects in the intervention group had immediate access to the 12-week internet-based self-help treatment. They were advised to spend at least 15 to 20 min daily on the program. Patients in the CG did not receive any treatment from the researchers during the first 12 weeks after the initial survey. However, they could seek other assistance, including pharmacological and psychological treatments. All concurrent treatments were measured repeatedly using self-reports.

2.5. Control group

A waitlist design was chosen for the CG to control for changes due to spontaneous improvement, e.g., regression to the mean. Besides, 12 weeks is a typical waiting time for psychotherapy in Germany (Bundespsychotherapeutenkammer, 2018), meaning that this comparison will allow us to assess the additional benefit of the intervention in the current situation.

2.6. Patient characteristics

The sociodemographic characteristics of all patients are displayed in Table 1. Eighty-three (53.5 %) patients fulfilled the criteria for a diagnosis of social phobia, and 29 (18.7 %) fulfilled the criteria for panic disorder in addition to a diagnosis of GAD. Also, 25 (16.0 %) patients were diagnosed with major depressive disorder, and 20 (12.9 %) patients had dysthymia. Depression levels were comparably low because patients with severe depression were excluded from participation. An overview of the frequency of all past or present diagnoses is shown in OSM 50.

2.7. Safety monitoring and ethical standards

All patients gave informed consent. The study was approved by the ethics committee at the study center at Heidelberg University (Ethics Committee-No. AZ Prüß 2021 1/1). Patients were asked about suicidality at T1, T2, and T3. If they reported 1 (“on individual days”) or higher on a 4-point Likert scale (0: “not at all” - 3: “almost every day”) to the extent that they “had thoughts [that] you would rather be dead or would like to harm yourself” in the last two weeks, the patients were contacted by phone, and an emergency plan was drawn up with them. If contact by phone was not possible, contact was made via email. If patients were contacted due to suicidality, it was ethically necessary to implement a protocol that prevented further participation in completing questionnaires for these individuals to prioritize their immediate needs and ensure appropriate support. However, all data collected up to this point was still used. Exclusion due to suicidality had to be conducted once during the trial in the CG.

Table 1
Sociodemographic characteristics of the study cohort at baseline.

Characteristic	Treatment (*N* = 78)		Control (*N* = 78)		Total sample (*N* = 156)	
	N	%	*N*	%	*N*	%
Sex						
Female	69	88.46	59	75.64	128	82.05
Male	8	10.26	18	23.08	26	16.67
Non-binary	1	1.28	1	1.28	2	1.28
Age in years (*M*, *SD*)	33.2	10.5	37.2	12.4	35.2	11.6
Health care use						
Psychotherapy	19	24.36	23	29.49	42	26.92
Psychopharmacotherapy	26	33.33	27	34.62	53	33.97

2.8. Outcomes

Primary and secondary measures were conducted at the beginning of the intervention (T1, baseline), after 6 weeks (T2, mid-treatment), and after 12 weeks (T3, post-treatment). The patients had to complete questionnaires regarding the primary, secondary, and exploratory hypotheses at all measurement points, except for the negative effects questionnaire, which was only given at T2 and T3. Additionally, at T1, a demographic questionnaire had to be filled out.

2.8.1. Primary outcome measures

The change in generalized anxiety disorder symptoms was evaluated using the *Generalized Anxiety Disorder Scale-7* (*GAD-7*; Spitzer et al., 2006). As a reliability measure for our data, we calculated McDonald's omega (Hayes and Coutts, 2020) with $\omega = 0.75$ for the *GAD-7* at T1. Wellbeing was assessed by the five-item *World Health Organization Well-being Index* (*WHO-5*; Topp et al., 2015). The reliability of the *WHO-5* was high, with $\omega = 0.81$ at the T1.

2.8.2. Secondary outcome measures

Four secondary outcomes were collected. Self-reported difficulties in daily functioning were measured by the Work and Social Adjustment Scale (*WSAS*; Mundt et al., 2002). The *WSAS* is a five-item scale with a test-retest correlation of $r = 0.73$ (Mundt et al., 2002) and had good reliability with $\omega = 0.73$ at T1.

Working ability was measured with the *iMTA Productivity Cost Questionnaire* (*iPCQ*; Bouwmans et al., 2013) to assess the amount of lost working hours due to the mental disorder in the last four weeks. The *iPCQ* consists of 18 questions aimed at measuring these hours. The *iPCQ* has been shown to be feasible and easily understandable with a retest-reliability of $icc = 0.83$ (Bouwmans et al., 2013, 2015).

Mental health literacy was measured with the *Mental Health Literacy Scale* (*MHLS*; O'Connor and Casey, 2015). The *MHLS* contains 35 items about knowledge, attitudes, and competencies regarding mental health with high reliability ($\omega = 0.85$ at T1) and good retest-reliability of $r = 0.79$ (O'Connor and Casey, 2015).

The extent of therapy-related healthcare usage was collected using three subscales of the *Client Sociodemographic and Service Receipt Inventory* (*CSSRI*; Chisholm et al., 2000): *CSSRI partly inpatient* to assess the amount of time of partly inpatient treatment (e.g., psychiatric day care), *CSSRI complementary* to assess attendance of complementary services (e.g., self-help groups), and *CSSRI outpatient* to assess attendance of outpatient services (e.g., psychotherapy treatment, medical treatment). Sousa et al. (2013) report excellent interrater-reliability of the *CSSRI* with coefficients between 0.8 and 1.0.

2.8.3. Exploratory outcome measures

The following outcomes were deemed as exploratory, since they seem less central and there is already some data on these outcomes (e.g., Eilert et al., 2021). Adverse treatment effects were collected with the *Negative Effects Questionnaire* (NEQ; Rozental et al., 2016) as an exploratory measure. The NEQ contains 32 items and shows satisfactory reliability with $\omega = 0.71$ at T3. Also, the Beck Anxiety Inventory (*BAI*; Beck et al., 1988) assessed general symptoms of anxiety. The *BAI* is a 21-item scale with a high internal consistency with $\omega = 0.88$ at baseline measurement and retest-reliability of $r = 0.78$ (Geissner and Huetteroth, 2018). Depressive symptoms were collected with the Patient Health Questionnaire-9 (*PHQ-9*; Kroenke et al., 2001). The *PHQ-9* is a short scale of nine items with good reliability ($\omega = 0.75$ at T1) and good retest-reliability with $r = 0.84$ (Kroenke et al., 2001). Levels of trait worry were collected with the *Penn State Worry Questionnaire* (PSWQ; Meyer et al., 1990). The *PSWQ* is a 16-item scale for assessing the worry facet as an accompanying syndrome of Generalized Anxiety Disorder. The *PSWQ* showed high reliability with $\omega = 0.82$ at T1.

2.9. Sample size

Presupposing recent meta-analytic data for unaccompanied online psychological interventions for anxiety disorders ($d = 0.45$; e.g., McCall et al., 2021), a planned mixed model with two measurement time points with a general correlation structure (Lu et al., 2008), a directed hypothesis, 1:1 group allocation, 80 % power, and an alpha level of 0.025 after Bonferroni-Holm correction, a total of 156 patients (78 per group) were needed (*longpower*; Donohue, 2021). For the secondary outcomes, we estimated a minimal effect size of $d = 0.48$ to achieve 80 % power with an alpha level of 0.0125 (Bonferroni-Holm adjustment) based on a post-hoc power analysis of the *WSAS* (*simr*; Green and MacLeod, 2016).

2.10. Randomization and blinding

Eligible subjects were randomly allocated to either immediate online self-help for GAD (IG) or intervention after 12 weeks (CG). Randomization happened after the clinical interview and was conducted by a non-project member using computer assistance. Therefore, allocation remained concealed until after the video interview so that interviewer blindness was ensured. Subjects were assigned in a 1:1 ratio. Patients were informed of their group via email, and were told that the waiting time was randomly varied. After data collection, statistical analysis occurred blindly, managed by an independent team member. The analysis R script was prepared before data collection.

2.11. Statistical analyses

The statistical analyses were conducted following the study protocol (Rubel et al., 2023), which was prepared and agreed upon before the analysis began. The analyses were performed with *R*, version 4.2.0 (R Core Team, 2022).

For the analyses of the psychometric outcome measures, different analyses were calculated. As documented in the study protocol, all hypotheses were evaluated based on an ITT analysis with missing values replacement by multiple imputations ("Multivariate Imputation by Chained Equations"; MICE; with $n = 5$ imputations; Azur et al., 2011) based on the control arm, using the variables "age" and "gender" as predictors in addition to the measurement-repeated variable. Four additional sensitivity analyses were added: A completer analysis using only patient data with completed T1 and T3 measures, "last-observation-carried-forward" (LOCF), "baseline-observation-carried-forward" (BOCF), and a "reference-based-multiple imputation" (J2R; Carpenter et al., 2013).

The primary outcome parameters were examined regarding the fixed interaction effect between group allocation and time for T3 (last survey). In order to adjust for multiple testing, a Bonferroni-Holm adjustment was performed for the primary and secondary outcomes. The results reported in the text refer to the MICE imputation analyses unless specified otherwise, as these were the decisive criteria for significance.

The confirmatory analysis of the primary endpoints consisted of calculating a mixed model with two measurement time points with a general correlation structure (Lu et al., 2008). A random effect for the subject was calculated (random intercept), and three fixed effects (group assignment, time, and the interaction of the two effects). The two measurement time points were nested within subjects. Secondary confirmatory outcomes were calculated only after success in the primary analysis similarly.

Independent *t*-tests and $\chi 2$-tests were used to estimate differences between groups in pretreatment sample characteristics. Also, *t*-tests were used to identify differences in negative effects in the NEQ at T2 and T3. In addition to the ITT sample, a "per-protocol" sample sensitivity analysis was defined for exploratory analyses, including all IG patients who completed at least 4 of the 12 modules.

To assess the magnitude of the treatment effects, the fixed interaction effect of time and group assignment was divided by the root of the

summed variances of the random effects (Westfall et al., 2014). Effect sizes can be roughly interpreted according to Cohen's *d*: Effect sizes of 0.20 are considered small, 0.50 moderate, and 0.80 large (Cohen, 1988). Differences in response rates and rates of use of additional care system services were examined with *t*-tests and χ2-tests. Additionally, reliable change was assessed using the Reliable Change Index (RCI; Jacobson and Truax, 1992) to calculate reliable improvement or deterioration. The RCI[3] was calculated for the MICE imputation by calculating change scores and taking into account the variability of these scores to assess if change is systematic or just random. Patients were either classified as reliably deteriorated, unchanged, or reliably improved.

The *CSSRI* questionnaire was split into three subscales (*CSSRI partially inpatient*, *CSSRI outpatient*, and *CSSRI complementary*) subject to Bonferroni-Holm adjustment by dividing by 3, 2, and 1, respectively.

Due to highly skewed data, the *iPCQ* and the *CSSRI partially inpatient* scales were log10-transformed. The *CSSRI outpatient* and *CSSRI complementary* scales wer0e dichotomized because of rare extreme outliers, in which case a transformation is not useful (Lei et al., 2017). The analysis was kept as close as possible to the preregistration for these dichotomized measures by calculating a mixed logistic regression model with random intercept and three fixed effects (group, time, and group*time). Odds ratios were calculated as the effect size.

The exploratory outcome measures *BAI*, *PHQ-9*, and *PSWQ* were analyzed using the same model as the primary and secondary outcomes but without alpha adjustment and as a two-sided test because of the exploratory nature of these outcomes.

2.12. Data and code

All data and the respective analysis code have been made publicly available at the OSF repository and can be accessed at https://osf.io/tj4re/. Materials about the content of the intervention are reported in the study protocol (Rubel et al., 2023).

3. Results

3.1. Participant flow

Recruitment spanned from 17th February 2021 to 1st April 2022 (see Fig. 1). In total, 4361 participants filled out the screening questionnaire, and 764 met participation criteria, while 160 withdrew. Of the remaining 604, 292 (48.3 %) lacked appropriate diagnosis, and 156 (26.0 %) were excluded due to panic disorder eligibility and were treated in another trial targeting panic disorder (Lalk et al., 2023). Due to the parallel enrolment for iCBT for GAD and for panic disorder, the number of exclusions is larger than it would be expected for a single trial of this size. Unfortunately, it is impossible to assign excluded subjects to a specific trial, leading to these high numbers. Chi-square testing did reveal significant differences between the groups (see Table 2) regarding their relationship status ($\chi^2 = 9.53$, $p = .013$) with fewer married and more singles in the IG. Further, the IG was younger on average ($t = -2.13$, $p = .035$). Regarding diagnoses, significant differences were found for obsessive-compulsive disorder (5.2 % in IG versus 18.0 % in CG; $p = .013$).

No differences were found for current psychopharmacology ($\chi^2 = 0.029$, $p = .866$), psychotherapy ($\chi^2 = 0.521$, $p = .470$), or baseline scores of any primary or secondary outcome. Altogether, 156 patients were randomized, with 78 patients in the IG and 78 in the CG. All randomized patients were included in the analysis.

3.2. Adverse events

One patient from the CG reported suicidality and had to be excluded from the study. However, their previous data were still included and missing data was analyzed adhering to ITT principles.

3.3. Missing data

Non-completion rates for GAD-7 were 16.0 % at post-treatment and 19.9 % for WHO-5 resulting in 76.9 % ($n = 60$) completion of GAD-7 and 73.1 % ($n = 57$) completion of WHO-5 in the IG, and 91.0 % ($n = 71$) completion of GAD-7 and 87.2 % ($n = 68$) completion of WHO-5 in the CG. Logistic regression indicated no associations between non-completion and baseline variables like group, sex, age, fitness to work, medication, psychotherapy, or baseline values of outcomes. Dropouts occurred in 18 (23.1 %) patients from the IG and 6 (7.7 %) from the CG. Group differences were non-significant at T2 and T3 but significant overall ($\chi^2(1) = 7.09$, $p = .008$). No significant gender differences in dropout emerged.

3.4. Adherence

Due to a technical problem, usage data was missing for three persons in the IG. Based on data from 75 patient, the IG completed an average of 7.1 (SD = 4.0) modules out of the total 12 modules. 21 (28.0 %) patients finished the whole course, 40 (53.3 %) patients finished seven modules (main course), and 53 (70.7 %) underwent the first four modules, which was chosen as a sensitivity analysis to assess a basic amount of engagement.

3.5. Primary outcomes

Descriptive statistics for the primary outcome measures (*GAD-7* and *WHO-5*) for each assessment point are shown in Table 2 and OSM 2–3 for ITT raw scores and MICE imputation.

Regarding the main interaction effect at T3 (see Table 3), the *GAD-7* measure was highly significant ($t = -4.52$, $p < .001$) across all imputations (see OSM 4). Further, the *WHO-5* measure ($t = 3.48$, $p < .001$) was also highly significant with Bonferroni adjustment and across all imputations (see OSM 5). The imputations and effects can be seen in OSM 2–11. For *GAD-7*, the effect size was large (−0.88, 95 %-CI: −1.26; −0.50) and for *WHO-5*, it was moderate (0.62 95 %-CI: 0.27; 0.98).

Within-group effect sizes were large in the intervention group (1.02, 95 %-CI: 0.74; 1.29) and marginal to small in the control group for *GAD-7* (0.14, 95 %-CI: −0.11; 0.39). The within-group effects for the *WHO-5* were moderate to large in the intervention group (0.65, 95 %-CI: 0.36; 0.94) and marginal in the control group (0.03, 95 %-CI: −0.15; 0.21).

3.5.1. Minimal clinical important difference

Reliable improvement was achieved in the *GAD-7* by 43.8 % of the patients in the IG and 13.8 % in the CG. In the IG, 4.6 % deteriorated, while 9.7 % deteriorated in the CG. Therefore, a significant group difference in improvement and deterioration was identified ($p < .001$). Reliable improvement was achieved in the *WHO-5* by 32.1 % of IG patients and 7.2 % of CG patients, while 5.6 % in the IG and 8.2 % in the CG deteriorated, constituting a significant difference ($p < .001$).

3.5.2. Per-protocol analyses

Additionally, per-protocol sensitivity analyses were calculated for patients that underwent the first four modules. Because no hypotheses were specified, the effects are presented as two-tailed tests without alpha adjustment. A significant group*time interaction ($t = 6.11$, $p < .001$) was found for the *GAD-7* with a large effect size ($d = 1.24$, 95 %-CI: 0.84; 1.63) and the *WHO-5* ($t = 4.56$, $p < .001$; $d = 0.93$, 95 %-CI: 0.53; 1.33).

[3] $RCI = \frac{y_1 - y_0}{s_{diff}}$ with $s_{diff} = \sqrt{2\left(s_0\sqrt{1 - r_{yy}}\right)^2}$, s_0 is the standard deviation of y_0 and r_{yy} the retest-reliability of y.

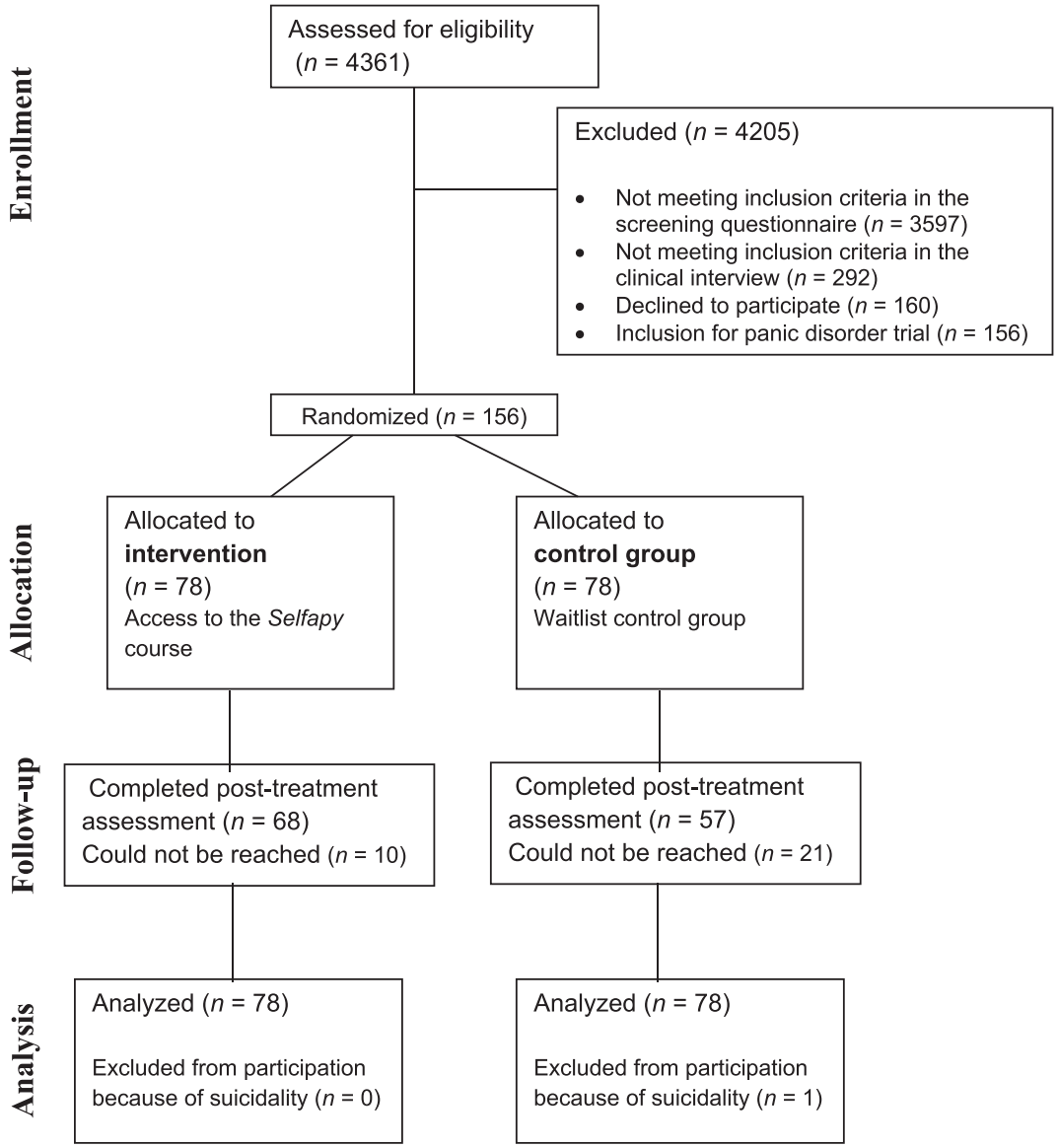


Fig. 1. CONSORT flow diagram.
Note. Deviations from the sample size did occur for some of the secondary and exploratory outcomes.

Table 2
Descriptive statistics for the primary outcomes.

Imputation/ group	T1		T2		T3	
	N	*M* (*SD*)	*N*	*M* (*SD*)	*N*	*M* (*SD*)
GAD-7						
Treatment	78	19.36 (3.23)	64	16.72 (3.99)	60	14.83 (3.54)
Control	78	19.14 (3.49)	70	18.39 (3.79)	71	18.59 (3.53)
WHO-5						
Treatment	78	2.33 (0.74)	61	2.77 (0.97)	57	2.99 (1.03)
Control	78	2.39 (0.69)	66	2.46 (0.80)	68	2.46 (0.77)

Table 3
Bonferroni-Holm adjustment for the primary outcomes.

Primary outcome	*t* (*df*)	*p* (one-sided)	Adjustment factor	Adjusted *p*	Effect size *d* (95 % CI)
GAD-7	4.52 (123.73)	<.001***	2	<.001***	0.88 (0.50; 1.26)
WHO-5	3.48 (87.86)	<.001***	1	<.001***	0.62 (0.27; 0.98)

Note. * indicates $p \leq .05$, ** indicates $p \leq .01$, *** indicates $p \leq .001$.

3.6. Secondary outcomes

Descriptive statistics can be found in Tables 4 and 5. Additional descriptive statistics, effect sizes and significance levels across the different imputations can be found in the OSM (OSM 12–35). None of the interaction effects was significant after the Bonferroni-Holm adjustment (see Table 6) and no effects in the expected directions were found in the imputations.

No interaction effect was found for all outcomes, even before the Bonferroni-Holm adjustment. Only in the WSAS a within-group effect was found but only for the IG (0.32 95 %-CI: 0.10; 0.53).

Table 4
Descriptive statistics for the secondary outcomes.

Group	T1		T2		T3	
	N	*M* (*SD*)	*N*	*M* (*SD*)	*N*	*M* (*SD*)
WSAS						
Treatment	78	4.69 (1.64)	61	4.29 (1.67)	57	4.03 (1.81)
Control	78	4.86 (1.57)	66	4.58 (1.62)	68	4.69 (1.74)
MHLS						
Treatment	78	4.33 (0.38)	61	4.33 (0.39)	57	4.35 (0.51)
Control	78	4.25 (0.43)	66	4.20 (0.48)	68	4.26 (0.50)
iPCQ						
Treatment	77	0.52 (0.65)	64	0.39 (0.53)	58	0.38 (0.55)
Control	77	0.39 (0.57)	68	0.38 (0.54)	71	0.38 (0.59)
CSSRI outpatient (log-transformed)						
Treatment	71	1.20 (1.11)	47	1.28 (1.10)	48	1.22 (1.25)
Control	73	1.15 (1.21)	55	1.11 (1.13)	59	0.96 (1.09)

Table 5
Descriptive statistics for the dichotomized secondary outcomes.

Imputation/group	T1		T2		T3	
	N	% cases	*N*	% cases	*N*	% cases
CSSRI partly inpatient						
Treatment	73	5.1 %	47	4.3 %	48	4.2 %
Control	69	13.0 %	50	10.0 %	59	5.1 %
CSSRI complementary						
Treatment	74	9.5 %	47	19.1 %	47	10.6 %
Control	73	15.1 %	49	18.4 %	59	11.9 %

3.7. Exploratory outcomes

Descriptive values, significance levels, and effect sizes across all imputations can be obtained from the OSM for the exploratory outcomes (OSM 36–47). The effects of the exploratory outcomes are displayed in Table 7. A significant interaction was found for the BAI ($t = 2.87$, $p = .004$) with a small to moderate effect size (0.38, 95 %-CI: 0.12; 0.64). Significance was maintained across all imputations (see OSM 43). Effect sizes for the within-group changes were moderate to large for the IG (0.64, 95 %-CI: 0.45; 0.84) and small for the CG (0.27, 95 %-CI: 0.10; 0.43).

For the *PHQ-9*, no significant interaction could be found ($t = 1.83$, $p = .067$), though all imputations were significant (OSM 39). Within-group effects were small to moderate (0.37, 95 %-CI: 0.14; 0.59) in the IG and were minimal (0.10, 95 %-CI -0.08; 0.28) in the CG.

For the PSWQ, a significant interaction was found ($t = 2.58$, $p = .009$) with a moderate effect size (0.47, 95 %-CI: 0.11; 0.82). Significance was only maintained across the completer cases and BOCF imputations (OSM 47). Only a moderate within-group effect (0.46, 95 %-CI: 0.21; 0.71) was found in the IG with no within-group effect in the CG (0.01, 95 %-CI: −0.22; 0.23).

Regarding side effects, significant differences were found in the *Negative Effects Questionnaire* (NEQ, Rozental et al., 2016) with fewer adverse effects in the IG ($t = 2.85$, $p = .005$; see Table 8 and OSM 48–49). In a sensitivity analysis without the NEQ items that specifically addressed treatment (without items 20–32; e.g., "I did not always understand my treatment"), still fewer negative effects were reported in the IG ($t = 3.98$, $p < .001$).

4. Discussion

The present study aimed to investigate the effects of an online self-help intervention for patients with GAD in conditions similar to routine care. Patients using the online self-help intervention revealed a more pronounced reduction in generalized anxiety symptoms compared to the waitlist control group. Additionally, well-being improved significantly compared to the CG. No effects were found for the secondary outcomes functioning in daily life, mental health literacy, working ability, and healthcare use. Regarding the exploratory outcomes, more substantial reductions were found for anxiety and rumination symptoms but not for depression.

The investigated iCBT intervention resulted in substantial within-group reductions in GAD symptom severity over the course of the

Table 7
Linear mixed model and effect sizes for secondary exploratory outcomes – MICE T3-T1.

Outcome	Group * time			
	t	*df*	*p* (two-sided)	Interaction effect (95 %-CI)
BAI	2.87	286.40	0.004**	0.38 (−0.12; 0.64)
PHQ-9	1.83	130.74	0.067	0.27 (−0.02; 0.56)
PSWQ	2.58	223.05	0.009**	0.47 (−0.11; 0.82)

Note. * indicates $p \leq .05$, ** indicates $p \leq .01$, *** indicates $p \leq .001$. MICE, Multiple Imputation by Chained Equations. BAI, Beck Anxiety Inventory; PHQ-9, Patient Health Questionnaire-9; PSWQ, Penn State Worry Questionnaire.

Table 8
The most common negative effects in the intervention group.

Negative effect	Frequency
Unpleasant memories resurfaced	36 (46.2 %)
I experienced more unpleasant feelings	18 (23.1 %)
I felt more worried	16 (20.5 %)
I felt like I was under more stress	14 (17.9 %)
I had more problems with my sleep	13 (16.7 %)
I did not always understand my treatment	11 (14.1 %)
I experienced more anxiety	9 (11.5 %)
I felt that my expectations for the treatment were not fulfilled	9 (11.5 %)
I felt that the treatment did not suit me	9 (11.5 %)
I felt more dejected	8 (10.3 %)
I started thinking that the issue I was seeking help for could not be made any better	8 (10.3 %)

Table 6
Bonferroni-Holm adjustment for secondary outcomes.

Secondary outcome	*t*	*p* (one-sided)	Adjustment factor	Adjusted *p*	Effect size *d* (95 %-CI)
WSAS	1.37	.085	4	.273	0.20 (0.09; 0.49)
iPCQ	1.12	.130	3	.375	0.01 (−0.24; 0.27)
MHLS	0.10	.460	2	.912	0.20 (−0.15; 0.54)
CSSRI partly inpatient	0.22	1.000[a]	3 * 1[a]	1.000[a]	0.06 (−0.27; 0.40)
CSSRI outpatient	0.38	1.000[a]	2 * 1[a]	1.000[a]	1.45[b] (0.05; 40.8)
CSSRI complementary	0.53	1.000[a]	1 * 1[a]	1.000[a]	1.82[b] (0.19; 17.2)

[a] The *CSSRI* subscales were additionally corrected with Bonferroni-Holm correction by multiplication with the additional factors 3, 2, and 1, respectively.
[b] The scales were dichotomized and odds ratios were used as effect size.

treatment ($d = 1.02$). Symptom reduction was significantly more pronounced in patients who received the iCBT compared to those who did not receive treatment ($d = 0.88$). This effect is comparable to those reported in meta-analytic reviews on iCBT for GAD ($g = 0.79$, Eilert et al., 2021) and even traditional CBT ($g = 0.84$; Cuijpers et al., 2014). Also, similar to other iCBT trials (e.g., Ritola et al., 2022), effect size seemed to be affected by adherence: Patients with at least 4 of altogether 12 completed modules had an even bigger effect size ($d = 1.24$). The robustness of the effect is supported by the effects on worry ($d = 0.47$) and anxiety ($d = 0.38$). Therefore, it seems like the intervention is targeting both cognitive processes associated with GAD (i.e. worrying), as well as physical symptoms of anxiety. Yet, in their meta-analysis on the effects of iCBT on anxiety and depression across 19 trials, Păsărelu et al. (2017) found large effects on both anxiety and depression and moderate effects on anxiety comorbidities. Therefore, the effects on worry, anxiety and depression are smaller than would be expected, which may be in part due to the exclusion of severe depression as discussed later.

Regarding wellbeing, the effects were moderate to large ($d = 0.65$) within-group and moderate in comparison to the waitlist ($d = 0.62$). These effects are stronger than meta-analytic findings on quality of life ($d = 0.33$, Eilert et al., 2021). However, this might also be due to the different constructs used since the WHO-5 assesses wellbeing (Topp et al., 2015), while Eilert et al. (2021) looked at broader quality of life outcomes. Since quality of life is measured across different areas (e.g., mental health, physical health, relationships, purpose, …) it may be less sensitive towards change in mental health intervention studies. In contrast, wellbeing in general and the WHO-5 specifically are focused on current overall life satisfaction (Skevington and Böhnke, 2018). Also, the WHO-5 is known to be sensitive to change and is also used as a depression measure (Topp et al., 2015). To our knowledge, no iCBT study on GAD investigated the effects of the intervention on psychological wellbeing.

Regarding the secondary hypotheses, no significant interaction effects were found. However, there was a small to moderate within-group effect on functioning ($d = 0.32$) for the intervention with a marginal effect ($d = 0.11$) in the CG. No within-group effects were found for the MHLS or any CSSRI subgroup. In the iPQC, a small within-group effect ($d = 0.19$) was detected in the IG and no effect ($d = 0.01$) in the CG. There are several reasons for the absence of significant findings in the secondary outcomes. Firstly, the trial was not specifically powered to detect effects in the secondary outcomes, which was further impeded by the additional Bonferroni-Holm adjustment. A post-hoc power analysis showed that only moderate effects ($d = 0.48$) could be detected with 80 % power after alpha adjustment. Secondly, both the iPCQ and CSSRI subscales exhibited a strong floor effect (Ho and Yu, 2015) due to the rare occurrence of lost working hours (46.1 % reported none) and additional treatments (33.8 % no outpatient treatments, 90.8 % no inpatient appointments, 87.7 % no complementary treatments) at the beginning of the study. The MHLS scale also presented a ceiling effect (mean of 4.3 and a range from 2.9 to 5.0 on a scale from 1 to 5). Thirdly, the CSSRI subscales measured treatment use over the past three months, including the initial phase of the intervention, which might not have captured its full impact. Similarly, the iPCQ assessed the past six weeks, potentially leading to similar issues. In summary, the lack of significant results in the secondary outcomes can be attributed to the inadequate power analysis, substantial floor and ceiling effects, and extended measurement timeframes. Regarding the WSAS, meta-analytic evidence by Eilert et al. (2021) reported a substantial effect on functioning ($d = 0.66$) across ten studies with 343 patients, contradicting our findings. Nevertheless, this study has the most power to assess functioning and should not be discounted easily.

The exploratory outcomes revealed a significant interaction effect for anxiety and significantly lower levels of trait worry in favor of the IG. The lack of effects on depressive symptoms via the PHQ-9 is a surprise since Eilert et al. (2021) found a large effect in their meta-analysis. One reason might be the exclusion of patients with severe depression from the trial, which lead to high levels of patients with only mild or no depression in both the IG (82.0 %) and the CG (76.9 %).

Regarding negative treatment effects, patients in the IG reported fewer negative treatment effects than the negative effects of waiting reported in the CG. This was confirmed in a sensitivity analysis without items that directly focused on treatment effects. The reported negative effects are comparable in quality and quantity to the negative effects of face-to-face psychotherapy (e.g., increase in unpleasant memories and unpleasant feelings; Strauss et al., 2021). As such, we can conclude that the iCBT for GAD tested in the present study can be safely used for the included study population.

4.1. Implications for future research

The effects of iCBT on functioning at home and work need to be investigated more thoroughly. Even though current meta-analytical evidence suggests an effect (Eilert et al., 2021), most studies, including this one, are too small and not adequately powered with a minimal detectable effect size of $d = 0.48$ for the secondary outcomes. Regarding health care usage and lost working hours, assessing these outcomes in a sample with higher impairment and a follow-up measurement to account for the floor effects and long measurement timeframes might be worthwhile.

4.2. Limitations

The study has several limitations. First, patient blinding wasn't feasible due to the study design. Second, choosing a waitlist CG might overestimate treatment effect by delaying remission and help-seeking in the CG (e.g., Cuijpers et al., 2016). Still, in the present trial we decided that this is the most appropriate condition given our primary goal to evaluate the added value of iCBT for GAD in comparison to the current situation where treatment delays are the rule. As such, the *true* differential effects may in fact be slightly smaller with the extent of the overestimation remaining uncertain. To counteract this, CG patients were encouraged to seek additional help, minimizing perceived passivity. Third, the study sample was predominantly female (82.1 %), highly educated (42.6 % with university degrees), and relatively young ($M = 35.2$, $SD = 11.6$). This demographic might result from higher female anxiety disorder prevalence (Bandelow and Michaelis, 2015), women's increased help-seeking (Kessler et al., 1981), and recruitment methods (e.g., university and social media). Therefore, caution is warranted in applying findings to older or male populations. Still, existing studies show no barriers for males (Carl et al., 2020) or older individuals (Silfvernagel et al., 2018; Cremers et al., 2019) in online GAD interventions. Fourth, although the online course structure is standardized, we had no insight into how much time each patient spent on the course content. Though effect sizes were more pronounced for patients who finished at least four modules, still substantial effects were identified on the primary outcomes irrespective of the time spent on the course. Fifth, due to limited resources, we did not include a structured diagnostic interview post-treatment, so that we did not assess how many patients recovered, which would be an important additional assessment.

Despite these limitations, the present trial was comparably well-powered and provides additional support for the beneficial effects of iCBT without therapist support in GAD patients by striking a balance between external (e.g., conditions comparable to routine care: patients may have comorbid diagnoses, additional treatments) and internal validity (randomization and blinding of investigators). Further, there is little research on the effects of iCBT on mental health literacy, working ability, and health care use.

Ethics approval

Approved by the ethics committee of the behavioral and empirical cultural science faculty at the study center at Heidelberg University

(Ethics Committee-No. AZ Prüß 2021 1/1).

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

External organizations

Not applicable.

Funding

The study was funded by a European Regional Development Fund awarded to Selfapy. The Professorship for Adult Clinical Psychology and Psychotherapy at Osnabrück University was asked by Selfapy to independently evaluate its product to be assessed to the best of our knowledge and following the standards of good scientific practice. Selfapy made suggestions for the selection of survey instruments in the study's design but had no influence on the data collection, analysis, and interpretation or production of the resulting publications.

CRediT authorship contribution statement

J. Rubel: Supervision, Methodology, Project administration, Writing – original draft. **T. Väth:** Writing – original draft, Formal analysis. **S. Hanraths:** Writing – original draft, Formal analysis. **L. Pruessner:** Writing – review & editing. **C. Timm:** Writing – review & editing. **S. Hartmann:** Writing – review & editing. **S. Barnow:** Writing – review & editing. **C. Lalk:** Formal analysis, Investigation, Software, Writing – original draft.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.invent.2024.100716.

References

Aguilera, A., 2015. Digital technology and mental health interventions: opportunities and challenges. Arbor 191 (771), a210. https://doi.org/10.3989/arbor.2015.771n1012.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, Fifth edition. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596.

Andrews, G., Basu, A., Cuijpers, P., Craske, M.G., McEvoy, P., English, C.L., Newby, J.M., 2018. Computer therapy for anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J. Anxiety Disord. 55, 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001.

Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J., 2011. Multiple imputation by chained equations: what is it, and how does it work? Int. J. Methods Psychiatr. Res. 20 (1), 40–49. https://doi.org/10.1002/mpr.329.

Bandelow, B., Michaelis, S., 2015. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin. Neurosci. 17 (3), 327–335. https://doi.org/10.31887/DCNS.2015.17.3/bbandelow.

Beck, A.T., Epstein, N., Brown, G., Steer, R.A., 1988. An inventory for measuring clinical anxiety: psychometric properties. J. Consult. Clin. Psychol. 56 (6), 893–897. https://doi.org/10.1037/0022-006x.56.6.893.

Becker, E.S., Hoyer, J., 2005. Generalisierte Angststörung: Vol. Band 25. Hogrefe.

Bereza, B.G., Machado, M., Einarson, T.R., 2009. Systematic review and quality assessment of economic evaluations and quality-of-life studies related to generalized anxiety disorder. Clin. Ther. 31 (6), 1279–1308. https://doi.org/10.1016/j.clinthera.2009.06.004.

Berger, T., Caspar, F., Richardson, R., Kneubühler, B., Sutter, D., Andersson, G., 2011. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. Behav. Res. Ther. 49 (3), 158–169. https://doi.org/10.1016/j.brat.2010.12.007.

Bouwmans, C., De Jong, K., Timman, R., Zijlstra-Vlasveld, M., Van der Feltz-Cornelis, C., Tan, S.S., Hakkaart-van Roijen, L., 2013. Feasibility, reliability, and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). BMC Health Serv. Res. 13 (1), 217 https://doi.org/10.1186/1472-6963-13-217.

Bouwmans, C., Krol, M., Severens, H., Koopmanschap, M., Brouwer, W., Hakkaart-van Roijen, L., 2015. The iMTA productivity cost questionnaire: a standardized instrument for measuring and valuing health-related productivity losses. Value Health 18 (6), 753–758. https://doi.org/10.1016/j.jval.2015.05.009.

Bundespsychotherapeutenkammer, 2018. Wartezeit für psychisch kranke Menschen weiterhin zu lang: BPtK zum G-BA-Gutachten zur Bedarfsplanung (Press release, Berlin, October 15).

Bundespsychotherapeutenkammer, 2021. BPtK-Auswertung: Monatelange Wartezeiten bei Psychotherapeut*innen: Corona-Pandemie verschärft das Defizit an Behandlungsplätzen (March 29, Press release, Berlin).

Burns, R.A., Windsor, T., Butterworth, P., Anstey, K.J., 2022. The protective effects of wellbeing and flourishing on long-term mental health risk. SSM - Mental Health 2, 100052. https://doi.org/10.1016/j.ssmmh.2021.100052.

Carl, E., Witcraft, S.M., Kauffman, B.Y., Gillespie, E.M., Becker, E.S., Cuijpers, P., Van Ameringen, M., Smits, J.A.J., Powers, M.B., 2020. Psychological and pharmacological treatments for generalized anxiety disorder (GAD): a meta-analysis of randomized controlled trials. Cogn. Behav. Ther. 49 (1), 1–21. https://doi.org/10.1080/16506073.2018.1560358.

Carpenter, J.R., Roger, J.H., Kenward, M.G., 2013. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions and inference via multiple imputation. J. Biopharm. Stat. 23 (6), 1352–1371. https://doi.org/10.1080/10543406.2013.834911.

Chisholm, D., Knapp, M.R.J., Knudsen, H.C., Amaddeo, F., Gaite, L., van Wijngaarden, B., 2000. Client Sociodemographic and Service Receipt Inventory – European Version: development of an instrument for international research: EPSILON Study 5. Br. J. Psychiatry 177 (S39), s28–s33. https://doi.org/10.1192/bjp.177.39.s28.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. L. Erlbaum Associates.

Cremers, G., Taylor, E., Hodge, L., Quigley, A., 2019. Effectiveness and acceptability of low-intensity psychological interventions on the well-being of older adults: a systematic review. Clin. Gerontol. 1–21. https://doi.org/10.1080/07317115.2019.1662867.

Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., Andersson, G., 2014. Psychological treatment of generalized anxiety disorder: a meta-analysis. Clin. Psychol. Rev. 34 (2), 130–140. https://doi.org/10.1016/j.cpr.2014.01.002.

Cuijpers, P., Cristea, I.A., Karyotaki, E., Reijnders, M., Huibers, M.J.H., 2016. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry : Official Journal of the World Psychiatric Association (WPA) 15 (3), 245–258. https://doi.org/10.1002/wps.20346.

Dear, B.F., Staples, L.G., Terides, M.D., Karin, E., Zou, J., Johnston, L., Gandy, M., Fogliati, V.J., Wootton, B.M., McEvoy, P.M., Titov, N., 2015. Transdiagnostic versus disorder-specific and clinician-guided versus self-guided Internet-delivered treatment for generalized anxiety disorder and comorbid disorders: a randomized controlled trial. J. Anxiety Disord. 36, 63–77. https://doi.org/10.1016/j.janxdis.2015.09.003.

Donohue, M.C., 2021. long power: Power and Sample Size Calculations for Longitudinal Data. R Package Version 1.0.23.

Eilert, N., Enrique, A., Wogan, R., Mooney, O., Timulak, L., Richards, D., 2021. The effectiveness of Internet-delivered treatment for generalized anxiety disorder: an updated systematic review and meta-analysis. Depress. Anxiety 38 (2), 196–219. https://doi.org/10.1002/da.23115.

Fredrickson, B.L., 2004. The broaden-and-build theory of positive emotions. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 359 (1449), 1367–1378. https://doi.org/10.1098/rstb.2004.1512.

Friborg, O., Martinussen, M., Kaiser, S., Overgård, K.T., Rosenvinge, J.H., 2013. Comorbidity of personality disorders in anxiety disorders: a meta-analysis of 30 years of research. J. Affect. Disord. 145 (2), 143–155. https://doi.org/10.1016/j.jad.2012.07.004.

Furnham, A., Swami, V., 2018. Mental health literacy: a review of what it is and why it matters. International Perspectives in Psychology 7 (4), 240–257. https://doi.org/10.1037/ipp0000094.

Geissner, E., Huetteroth, A., 2018. Beck Anxiety Inventory deutsch–Ein reliables, valides und praxisgeeignetes Instrument zur Messung klinischer Angst. PPmP-Psychotherapie· Psychosomatik· Medizinische Psychologie 68 (03/04), 118–125. https://doi.org/10.1055/s-0043-122941.

Grant, B.F., Stinson, F.S., Dawson, D.A., Chou, S.P., Dufour, M.C., Compton, W., Pickering, R.P., Kaplan, K., 2004. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders. Arch. Gen. Psychiatry 61 (8), 807–816. https://doi.org/10.1001/archpsyc.61.8.807.

Green, P., MacLeod, C.J., 2016. Simr: an R package for power analysis of generalised linear mixed models by simulation. Methods Ecol. Evol. 7 (4), 493–498. https://doi.org/10.1111/2041-210X.12504.

Hadjistavropoulos, H.D., Schneider, L.H., Edmonds, M., Karin, E., Nugent, M.N., Dirkse, D., Dear, B.F., Titov, N., 2017. Randomized controlled trial of internet-delivered cognitive behavior therapy comparing standard weekly versus optional weekly therapist support. J. Anxiety Disord. 52–15–24. https://doi.org/10.1016/j.janxdis.2017.09.006.

Hayes, A.F., Coutts, J.J., 2020. Use omega rather than Cronbach's alpha for estimating reliability. Commun. Methods Meas. 14 (1), 1–24. https://doi.org/10.1080/19312458.2020.1718629.
Ho, A.D., Yu, C.C., 2015. Descriptive statistics for modern test score distributions: skewness, kurtosis, discreteness, and ceiling effects. Educ. Psychol. Meas. 75 (3), 365–388. https://doi.org/10.1177/0013164414548576.
Hoffman, D.L., Dukes, E.M., Wittchen, H.-U., 2008. Human and economic burden of generalized anxiety disorder. Depress. Anxiety 25 (1), 72–90. https://doi.org/10.1002/da.20257.
Hoyer, J., Beesdo-Baum, K., Becker, E.S., 2016. Ratgeber generalisierte Angststörung: Informationen für Betroffene und Angehörige: Vol. Band 15 (2., aktualisierte Auflage), Hogrefe. https://elibrary.hogrefe.de/book/99.110005/9783840927089.
Jacobson, N.S., Truax, P., 1992. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. In: Kazdin, A.E. (Ed.), Methodological Issues & Strategies in Clinical Research. American Psychological Association, pp. 631–648. https://doi.org/10.1037/10109-042.
Judd, L.L., Kessler, R.C., Paulus, M.P., Zeller, P.V., Wittchen, H.-U., Kunovac, J.L., 1998. Comorbidity as a fundamental feature of generalized anxiety disorders: results from the National Comorbidity Study (NCS). Acta Psychiatr. Scand. 98 (s393), 6–11. https://doi.org/10.1111/j.1600-0447.1998.tb05960.x.
Kessler, R.C., Brown, R.L., Broman, C.L., 1981. Sex differences in psychiatric help-seeking: evidence from four large-scale surveys. J. Health Soc. Behav. 22 (1), 49–64. https://doi.org/10.2307/2136367.
Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2001. The PHQ-9. J. Gen. Intern. Med. 16 (9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.
Lalk, C., Väth, T., Hanraths, S., Pruessner, L., Timm, C., Hartmann, S., Barnow, S., Rubel, J., 2023. Evaluation of an Online-based Self-help Program for Patients With Panic Disorder - A Randomized Controlled Trial. Department of Humanities, Osnabrück University (Manuscript in preparation).
Lei, Y., Carlson, S., Yelland, L.N., Makrides, M., Gibson, R., Gajewski, B.J., 2017. Comparison of dichotomized and distributional approaches in rare event clinical trial design: a fixed Bayesian design. J. Appl. Stat. 44 (8), 1466–1478. https://doi.org/10.1080/02664763.2016.1214244.
Lu, K., Luo, X., Chen, P.-Y., 2008. Sample size estimation for repeated measures analysis in randomized clinical trials with missing data. The. Int. J. Biostat. 4 (1), 9 https://doi.org/10.2202/1557-4679.1098.
Margraf, J., Cwik, J.C., Pflug, V., Schneider, S., 2017a. Structured clinical interviews for mental disorders across the lifespan: psychometric quality and further developments of the DIPS Open Access interviews. [Structured clinical interviews for mental disorders across the lifespan: quality criteria and further developments of the DIPS procedures]. Journal of Clinical Psychology and Psychotherapy 46, 176–186. https://doi.org/10.1026/1616-3443/a000430.
Margraf, J., Cwik, J.C., Suppiger, A., Schneider, S., 2017b. DIPS Open Access: Diagnostic Interview for Mental Disorders [DIPS Open Access: Diagnostisches Interview bei psychischen Störungen]. Mental Health Research and Treament Center, Ruhr-Universität Bochum, Bochum. https://doi.org/10.13154/rub.100.89.
Martín-María, N., Miret, M., Caballero, F.F., Rico-Uribe, L.A., Steptoe, A., Chatterji, S., Ayuso-Mateos, J.L., 2017. The impact of subjective well-being on mortality: a meta-analysis of longitudinal studies in the general population. Psychosom. Med. 79 (5), 565–575. https://doi.org/10.1097/PSY.0000000000000444.
McCall, H.C., Hadjistavropoulos, H.D., Sundström, C.R.F., 2021. Exploring the role of persuasive design in unguided Internet-delivered cognitive behavioral therapy for depression and anxiety among adults: systematic review, meta-analysis, and meta-regression. J. Med. Internet Res. 23 (4), e26939 https://doi.org/10.2196/26939 (81).
Mewton, L., Smith, J., Rossouw, P., Andrews, G., 2014. Current perspectives on Internet-delivered cognitive behavioral therapy for adults with anxiety and related disorders. Psychol. Res. Behav. Manag. 7–37–46. https://doi.org/10.2147/PRBM.S40879.
Meyer, T.J., Miller, M.L., Metzger, R.L., Borkovec, T.D., 1990. Development and validation of the Penn State Worry Questionnaire. Behav. Res. Ther. 28 (6), 487–495.
Mowbray, C.T., Megivern, D., Mandiberg, J.M., Strauss, S., Stein, C.H., Collins, K., Kopels, S., Curlin, C., Lett, R., 2006. Campus mental health services: recommendations for change. Am. J. Orthop. 76 (2), 226–237. https://doi.org/10.1037/0002-9432.76.2.226.
Mundt, J.C., Marks, I.M., Shear, M.K., Greist, J.M., 2002. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br. J. Psychiatry 180 (5), 461–464. https://doi.org/10.1192/bjp.180.5.461.
O'Connor, M., Casey, L., 2015. The Mental Health Literacy Scale (MHLS): a new scale-based measure of mental health literacy. Psychiatry Res. 229 (1–2), 511–516. https://doi.org/10.1016/j.psychres.2015.05.064.
Păsărelu, C.R., Andersson, G., Bergman Nordgren, L., Dobrean, A., 2017. Internet-delivered transdiagnostic and tailored cognitive behavioral therapy for anxiety and depression: a systematic review and meta-analysis of randomized controlled trials. Cogn. Behav. Ther. 46 (1), 1–28. https://doi.org/10.1080/16506073.2016.1231219.
Portman, M.E., Riskind, J.H., Rector, N.A., 2012. Generalized anxiety disorder. In: Encyclopedia of Human Behavior. Elsevier, pp. 215–220. https://doi.org/10.1016/B978-0-12-375000-6.00175-0.
Preti, A., Demontis, R., Cossu, G., Kalcev, G., Cabras, F., Moro, M.F., Romano, F., Balestrieri, M., Caraci, F., Dell'Osso, L., Di Sciascio, G., Drago, F., Hardoy, M.C., Roncone, R., Faravelli, C., Gonzalez, C.I.A., Angermayer, M., Carta, M.G., 2021. The lifetime prevalence and impact of generalized anxiety disorders in an epidemiologic Italian National Survey carried out by clinicians by means of semi-structured interviews. BMC Psychiatry 21 (1), 48. https://doi.org/10.1186/s12888-021-03042-3.
R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Ritola, V., Lipsanen, J.O., Pihlaja, S., Gummerus, E.-M., Stenberg, J.-H., Saarni, S., Joffe, G., 2022. Internet-delivered cognitive behavioral therapy for generalized anxiety disorder in nationwide routine care: effectiveness study. J. Med. Internet Res. 24 (3), e29384 https://doi.org/10.2196/29384.
Robinson, E., Titov, N., Andrews, G., McIntyre, K., Schwencke, G., Solley, K., 2010. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 5 (6), e10942. https://doi.org/10.1371/journal.pone.0010942.
Rozental, A., Kottorp, A., Boettcher, J., Andersson, G., Carlbring, P., 2016. Negative effects of psychological treatments: an exploratory factor analysis of the Negative Effects Questionnaire for monitoring and reporting adverse and unwanted events. PLoS One 11 (6), e0157503. https://doi.org/10.1371/journal.pone.0157503.
Rubel, J.A., Quest, J., Pruessner, L., Timm, C., Hartmann, S., Barnow, S., Rittmeyer, L., Rosenbaum, D., Lalk, C., 2023. Evaluation of an online-based self-help intervention for patients with generalized anxiety disorder – study protocol for a randomized controlled trial. JMIR Res. Protocol. 12, e41440 https://doi.org/10.2196/41440.
Schulz, K.F., Altman, D.G., Moher, D., 2010. Consort 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ (Clinical Research Ed.) 340, c332. https://doi.org/10.1136/bmj.c332.
Silfvernagel, K., Westlinder, A., Andersson, S., Bergman, K., Diaz Hernandez, R., Fallhagen, L., Lundqvist, I., Masri, N., Viberg, L., Forsberg, M.-L., Lind, M., Berger, T., Carlbring, P., Andersson, G., 2018. Individually tailored internet-based cognitive behavior therapy for older adults with anxiety and depression: a randomized controlled trial. Cogn. Behav. Ther. 47 (4), 286–300. https://doi.org/10.1080/16506073.2017.1388276.
Skevington, S.M., Böhnke, J.R., 2018. How is subjective well-being related to quality of life? Do we need two concepts and both measures? Soc. Sci. Med. 1982 (206), 22–30. https://doi.org/10.1016/j.socscimed.2018.04.005.
Sousa, A., Cardoso, A., Oliveira, G.G., Mari, J.J., Razzouk, D., 2013. Standardized tool for measurement of health services costs and utilization in a sample with mental disorders: translation, cultural adaptation and interrater reliability of the client sociodemographic and service receipt inventory Brazilian version. Value Health 16 (7), A696. https://doi.org/10.1016/j.jval.2013.08.2093.
Spitzer, R.L., Kroenke, K., Williams, J.B.W., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.
Strauss, B., Gawlytta, R., Schleu, A., Frenzl, D., 2021. Negative effects of psychotherapy: estimating the prevalence in a random national sample. BJPsych Open 7 (6). https://doi.org/10.1192/bjo.2021.1025.
Suppiger, A., In-Albon, T., Herren, C., Bader, K., Schneider, S., Margraf, J., 2008. Reliabilität des Diagnostischen Interviews bei Psychischen Störungen (DIPS für DSM-IV-TR) unter klinischen Routinebedingungen. Verhaltenstherapie 18, 237–244. https://doi.org/10.1159/000169699.
Topp, C.W., Østergaard, S.D., Søndergaard, S., Bech, P., 2015. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother. Psychosom. 84 (3), 167–176. https://doi.org/10.1159/000376585.
Tyrer, P., Baldwin, D., 2006. Generalized anxiety disorder. Lancet 368 (9553), 2156–2166. https://doi.org/10.1016/S0140-6736(06)69865-6.
Volz, H.-P., Stieglitz, R.-D., 2010. Generalisierte Angststörung: Krankheitsbild, Komorbiditäten. Psycho- und Pharmakotherapie, Schattauer.
Westfall, J., Kenny, D.A., Judd, C.M., 2014. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. J. Exp. Psychol. Gen. 143 (5), 2020–2045. https://doi.org/10.1037/xge0000014.