Journal of Affective Disorders 286 (2021) 166–173

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

Efficacy of hypnotherapy compared to cognitive behavioral therapy for mild to moderate depression - Results of a randomized controlled rater-blind clinical trial

Kristina Fuhr [1], Christoph Meisner [2], Angela Broch [1], Barbara Cyrny [1], Juliane Hinkel [1], Joana Jaberg [1], Monika Petrasch [1], Cornelie Schweizer [1], Anette Stiegler [1], Christina Zeep [1], Anil Batra [1,*]

[1] Department of Psychiatry and Psychotherapy, University Hospital of Tuebingen, Tuebingen, Germany
[2] Institute for Clinical Epidemiology and Applied Biometry, University of Tuebingen, Tuebingen, Germany

ARTICLE INFO

Keywords:
Psychotherapy
Major Depression
Cognitive Behavioral Therapy
Hypnotherapy
Non-inferiority

ABSTRACT

Background: Methodologically well-designed RCTs concerning the efficacy of Hypnotherapy in the treatment of Major Depression are lacking. The aim of this study was to determine whether Hypnotherapy (HT) is not inferior to Cognitive Behavioral Therapy (CBT), the gold-standard psychotherapy, in the percentage reduction of depressive symptoms, assessed in mild to moderate Major Depression (MD).

Methods: This study reports the main results of a monocentric two-armed randomized-controlled rater-blind clinical trial. A total of 152 patients with MD were randomized to either CBT or HT receiving outpatient individual psychotherapy with 16 to 20 sessions for the duration of six months. The primary outcome was the mean percentage improvement in depressive symptoms assessed with the Montgomery-Asberg Depression Rating Scale (MADRS) before and after treatment.

Results: The difference in the mean percentage symptom reduction between HT and CBT was 2.8 (95% CI=-9.85 to 15.44) in the Intention-to-treat sample and 4.0 (95% CI=-9.27 to 17.27) in the Per Protocol sample (N=134). Concerning the pre-specified non-inferiority margin of -16.4, both results confirm the non-inferiority of HT to CBT. The results for the follow-ups six and twelve months after the end of the treatment support the primary results.

Limitations: For ethical reasons the trial did not include a control group without treatment; therefore we can only indirectly conclude that both treatment conditions are effective.

Conclusion: This is the first study to demonstrate that HT was not inferior to CBT in MD, while employing rigorous methodological standards.

1. Introduction

The lifetime prevalence for Major Depression (MD) is around 16% (Kessler et al., 2003). Only approximately 20% of the cases have received adequate treatment during the previous 12 months (Kessler et al., 2003). Furthermore, depressive disorders are the main source for chronic conditions in Europe as assessed by the World Health Organization (WHO, Global Health Estimates, 2016). In mild to moderate forms of MD, psychotherapy is known to be similarly effective as pharmacotherapy and even evinces a higher long-term efficacy, especially in the prevention of relapses (Hollon et al., 2005). The NICE guidelines on the treatment of MD (NICE, 2016) therefore recommend psychotherapy alternatively with a psychopharmacological treatment for mild to moderate unipolar depressive episodes. Cognitive Behavioral Therapy (CBT) and Interpersonal Therapy (IPT) are considered to be the evidence-based interventions with the highest empirical efficacy. The response rates of CBT and IPT, however, only reach about 50% (Luty et al., 2007). To enhance the efficacy of psychological interventions for MD, new approaches, e.g. Cognitive Behavioral Analysis System of Psychotherapy (CBASP, Wiersma et al., 2014) for the treatment of

* Corresponding author Dr. Anil Batra, Department of Psychiatry and Psychotherapy, University Hospital of Tuebingen, Calwer Str. 14, 72076 Tuebingen, Germany, phone +49 7071 29-82685, fax +49 7071 29-5384.
E-mail address: anil.batra@med.uni-tuebingen.de (A. Batra).

https://doi.org/10.1016/j.jad.2021.02.069
Received 23 November 2020; Received in revised form 10 January 2021; Accepted 28 February 2021
Available online 5 March 2021

chronic depression, have been evaluated in the last decades. It was found that CBASP alone and in combination with antidepressant medication was superior to antidepressant medication alone in patients with a history of childhood trauma (Nemeroff et al., 2003). Well-established interventions like psychodynamic approaches have been compared to CBT in non-inferiority trials and have achieved an outcome not inferior to that of CBT (Driessen et al., 2013). Following the results of differential response as found in Nemeroff et al. (2003), a greater variety of treatments strategies could increase the treatment outcome for subgroups of patients.

There is some evidence for the efficacy of Hypnotherapy (HT), one of the oldest treatment approaches, especially in the treatment of physical or psychophysiological disorders (Whorwell et al., 1984). However, concerning the treatment of mental disorders, evidence is lacking, although quite a number of psychotherapists in Germany attend further training in HT and claim to already work with HT strategies, even if HT is not covered by the current health care system. Few studies have investigated the efficacy of HT for the treatment of MD in a randomized trial. For example, Alladin and Alibhai (2007) compared "Cognitive Hypnotherapy" (CHT) to CBT in 84 patients with chronic depression according to the DSM-IV. CHT consisted of CBT enhanced with six additional HT strategies like self-hypnosis or positive mood induction. These six additional HT strategies showed additive effects on depressive symptoms at the end of the 16-weeks treatment and at follow-ups (Alladin and Alibhai, 2007). However, repeated measures ANOVA did not confirm the differences between CHT and CBT. Most of the studies only report about the content of depression treatment strategies with HT or further state that HT is indicated for depression as being an "evidence-based" treatment without reporting data (e.g. Alladin, 2009). Unfortunately, support for the efficacy of HT compared to a gold standard treatment in randomized controlled trials (RCTs) is missing. Furthermore, sample sizes of studies concerning the efficacy of HT for treatment of depressive symptoms are often too small, control groups are missing, or the allocation is not at random (Milling et al., 2019). Even a recent meta-analysis concerning the efficacy of HT did not include studies that required involving patients with the primary diagnosis of MD (Milling et al., 2019).

1.1. Aims of the present study

The present study therefore aims to compare the efficacy of HT with an evidence-based standard treatment (CBT) for mild to moderate MD. As numerous studies confirm the efficacy of CBT in the treatment for MD (NICE, 2016), we chose CBT as the standard treatment in our trial. We decided to omit a non-active control group for several ethical reasons. Firstly, providing only placebo or no treatment is not recommended in patients with mild to moderate MD (NICE, 2016; USFDA, 2016), and secondly, HT has been successfully applied in the treatment of MD in the past (Alladin and Alibhai, 2007). In the current study we tested the hypothesis that the treatment of MD with HT would be non-inferior to the treatment with CBT regarding the reduction of depressive symptoms (measured by the mean percentage of symptom reduction in the clinician-administered diagnostic questionnaire, the Montgomery-Åsperg Depression Rating Scale, MADRS). We also expected non-inferiority of HT compared to CBT six and twelve months after the end of the treatment.

2. Materials and Methods

2.1. Trial design

The clinical study was based on a single factor design with repeated measures with the factor treatment condition (CBT vs. HT). Outcome was assessed pre-treatment and post-treatment as well as six and twelve months after post-treatment. Written informed consent was obtained from all the patients after the procedures for participating in the trial had been fully explained. Afterwards, eligible participants were randomly assigned 1:1 to 20 individual sessions of either HT or CBT. A blocked randomization sequence with a fixed block size of 40 was created using nQuery 7.0 (Statsols, Cork, Ireland) for 160 patients (80:80). The treatment allocation was communicated via email between the statistical center of the trial (Institute for Clinical Epidemiology and Applied Biometry) and the study center shortly after inclusion for each patient. The details of the randomization sequence were unknown to the investigator and the coordinator. The method and result of the randomization were concealed to patients and therapists up to the start of the therapy. Even if block size was fixed, the central randomization process prevented anticipation of the randomization results. Patients were followed up after randomization and assessed every six months. Study length was in total 18 months for every participant. 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. All procedures involving human subjects/patients were approved by the Ethics Committee of the University Hospital Tuebingen (061/2015BO2, version 5.0, 23.11.2015). The trial was registered with ClinicalTrials.gov before recruiting participants (NCT02375308). The trial design and the full study protocol including the statistical plan for the primary and secondary analyses were published elsewhere (Fuhr et al., 2017). In this paper, we mainly focus on the results on the primary outcome, the non-inferiority of HT compared to CBT concerning the pre post percentage symptom reduction. Secondary endpoints were 1) a non-inferiority of HT compared to CBT related to a further symptom reduction over the follow-up period up to twelve months after end of treatment, 2) the response rate of patients defined as symptom reduction >/= 50%. The results concerning other secondary explorative endpoints will be reported elsewhere. Those were 3) the rate of remission (number of weeks) and the rate of relapses in the twelve months after the therapy, 4) the time to remission, and 5) an exploratory analysis concerning the (sociodemographic and psychopathological) predictors of the treatment response.

Beside the reduction of depressive symptoms measured by MADRS, (Serious) Adverse Events (SAEs/AEs) were assessed in order to monitor the safety of the patients. An external monitoring of inclusion and exclusion criteria, (S)AEs, as well as the study documentation and the primary endpoints was established to assure the quality of the trial.

2.2. Trial sample

The study was promoted through newspaper articles, a newsletter to the local psychiatry clinics and licensed psychiatrists and psychotherapists, and the study was announced per email to the employees at the university and the university hospital. Patients were recruited between May 2015 and December 2016 and screened in the study site at the University Hospital of Psychiatry and Psychotherapy Tuebingen. Main inclusion criterion was the diagnosis of a Major Depression (MD) with an actual mild to moderate episode according to the Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition (DSM-5, AMA,2013) and MADRS baseline scores < 35. Further inclusion criteria were: being at the age of 18-70 years, and - in case of existing anti-depressive medication - a stable medication for at least three months without planned changes during the duration of therapy was required. Assessment of inclusion and exclusion criteria at baseline was conducted by seven different raters using the Structured Clinical Interview for DSM-IV Axis I, SCID-I (First et al., 1997), adapted for DSM-5. We excluded patients with the lifetime diagnosis of a bipolar disorder or psychotic disorder, the diagnosis of chronic MD (duration >/= two years), the diagnosis of a current severe Major Depressive Episode according to SCID-I resp. MADRS > 34, acute suicidality (intended action, concrete plans or intermittent pronounced suicidal ideation), if patients were in remission of the actual episode since four weeks or longer, if patients had other severe primary mental disorders (for example, severe

personality disorder of borderline type with self-injury, active alcohol or drug dependence, current posttraumatic stress disorder, or anorexia nervosa), and if patients attended another outpatient psychotherapy during the last twelve months. Comorbid disorders like anxiety disorders or other personality disorder were assessed but not handled as exclusion criteria if the major depression was the primary disorder. Hypnotic suggestibility was not a necessary inclusion criterion for the study.

2.3. Assessments

2.3.1. Primary endpoint

The primary endpoint of the study was the change in depressive symptoms, as measured by the percentage improvement from pre to post treatment in the Montgomery–Åsberg Depression Rating Scale (MADRS, Montgomery and Asberg, 1979). As secondary endpoints the change from pre to six and from pre to twelve month follow-ups were assessed. We choose the percentage symptom reduction to enhance comparability with other trials and outcomes, see for example Khan et al., 2012. Raters were blind concerning the treatment condition of the patient. A total of seven raters were involved with baseline and five with the follow-up assessments. Raters had at least a bachelors' degree in psychology, participated in a course in clinical interviewing at university or elsewhere, and underwent a specific half-day training in the interviews for the current study. With the agreement of the patients, baseline and follow-up interviews were recorded on digital audio-tapes to calculate inter-rater-reliability of the MADRS between the original rating and a blind second rating. The raters documented whether they were un-blinded by the patients.

The internal consistency at baseline was low with α=.56 due to the fact that participants with MADRS scores ≥ 35 were excluded which limits the informative value of the score. The inter-rater reliability between two raters was ICC=.89 for the baseline MADRS score (with a randomly selected sample of n=11 interviews) and ICC=.89 for the post MADRS score (n=21 interviews).

2.3.2. Interventions

Both outpatient psychological interventions included 20 sessions of 50 minutes each during a period of six months. Up to four double sessions were allowed in order to adapt to periods of holiday breaks or illness. In the German health care system, up to 24 sessions after four initial sessions of CBT are considered to be short-term treatment. A total of eight therapists with sufficient qualifications (training in CBT or HT and at least 3 years of professional experience in the respective method) and intensive training were involved in the study treatment. There were four therapists in CBT and four in HT. In addition, none of the CBT therapists had completed any training in HT and vice versa. Because CBT treatment is more common, CBT therapists only attended a one-day workshop, HT therapists a two-day treatment manual workshop. Therapists were on average 40 years old (range 29-53), were all female, and had 3 to 20 years of experience. Each therapist was responsible for 12-24 patients. The therapy sessions were recorded after agreement of the patient.

The CBT manual was based on the well-established CBT manuals (Hautzinger, 2013) and included CBT techniques such as psychoeducation, behavioral activation, cognitive restructuring techniques, problem solving skills, and interpersonal skills. CBT is based on the cognitive model of Beck (2002) and the behavioral model of Lewinsohn (1974) assuming that depression is a combination of both, negative thinking and a loss/lack of positive reinforcements.

HT consisted of different modules which are currently used by a variety of trainers in the German hypnosis society of Milton Erickson (Milton Erickson Gesellschaft, MEG). Thus, the manual is a representative collection of internationally accepted HT techniques. The treatment manual was published as a book in Germany (Wilhelm-Goessling et al., 2020). HT is based on the theoretical humanistic assumption that depressive symptoms were developed in the past as a positive solution strategy to overcome a personal problem in the lifetime history. The modules of HT promote emotional activation and reinforcement of personal resources, the use of relevant emotionally positive and negative experiences from the biography, and the development of positive solution imagery. Formal trance induction, utilization techniques, and different types of metaphors were also used. Furthermore, the application of the different modules is guided and based on the patients' ego-structure (Rudolf, 2013). HT can be also adapted to the patients' hypnotic suggestibility.

Therapists' adherence to the manual was enhanced by regular monthly supervisions. Therapists further documented the use of the treatment strategies according to the manual. Adherence to the manual (treatment fidelity) was tested by four raters which were not involved at any time in the procedure of the RCT. They listened to randomly selected therapy sessions of 64 patients after the trial was closed. Each rater rated either one of the first, second, third, or last quarter of the sessions (1-5, 6-10, 11-15, 16-20) which resulted in a total of 253 rated therapy session (125 of HT and 128 of CBT). Inter-rater reliability was calculated over 10 randomly selected sessions that were rated by all four raters and was very high, ICC = .94 for CBT and ICC = .96 for HT. In HT, the therapists used more HT techniques (M = 9.13, SD = 6.61) than CBT techniques (M = 2.28, SD = 2.93) and in CBT, more CBT techniques (M = 10.19, SD = 4.56) than HT techniques (M = 1.40, SD = 3.43) were applied. Some of the "overlap" could have been due to the fact that some strategies were not represented by distinct but rather generic strategies like for example agenda setting or resource activation. The items used for the assessment of the treatment adherence are available on request by the authors.

2.4. Statistical Analysis Plan

2.4.1. Sample Size

The primary endpoint of the study was the comparison of HT and CBT regarding the mean percentage symptom reduction measured with the MADRS before and immediately after the end of the treatment. Using data from the trial comparing CBT to Interpersonal Therapy (Luty et al., 2007) we derived an improvement in the MADRS of 50% for CBT with a standard deviation of 32.9. Thus, to ensure that HT preserved at least two thirds of the effect of CBT, the non-inferiority margin was set at 16.4, an effect size equal to half the standard deviation. The margin was based on clinically and statistically important differences as well as ethical criteria, cost and feasibility. With 140 patients we would have a statistical power of 80% to reject the inferiority hypothesis (Change of MARDS for CBT – Change of MADRS for HT ≥ 16.4%) using a randomization proportion of 1:1 and an one-sided alpha-level of 5%. Assuming a maximal drop-out rate of 12.5%, a total of 160 patients with current mild to moderate MDE should be employed to reach at least a number of 140 assessable patients. The sample size was computed using nQuery version 7.0 (Statsols, Cork, Ireland).

2.4.2. Analysis of primary endpoint

According to the protocol the difference between the HT and CBT in the primary endpoint was estimated including a one-sided 95% Confidence Interval (CI). HT was regarded as non-inferior compared to CBT if the lower limit of the CI was greater than the pre-specified non-inferiority level of -16.4%.

The following hypotheses were tested at a one-site significance level of 5%:

H0: μCBT- μHT ≥ 16.4%
H1: μHT – μCBT < 16.4%

μ stands for the mean percentage change of MADRS between the baseline (t1) and the end of the treatment (t2) for each treatment condition.

According to the study protocol the primary confirmatory analysis was based on the per-protocol population (PP), which statistically

represents the conservative approach for the analysis of non-inferiority. The ITT analysis, which preserves the advantages of randomization, was done as supporting sensitivity analysis for the assessment of non-inferiority. Treatment participation was considered as per protocol (PP), if the patient attended at least 80% of the sessions (this corresponds to 16 sessions from a total of 20 planned sessions) and with complete data regarding the depressive symptoms (MADRS) after the end of the treatment. In view of the large amount of therapeutic content in both treatment arms, it seemed reasonable to us to only accept completion if at least 80% of the sessions were completed. We decided to replace missing data with the multiple imputations method (MI). Thus, after assuring that the missing data of the primary outcome measure

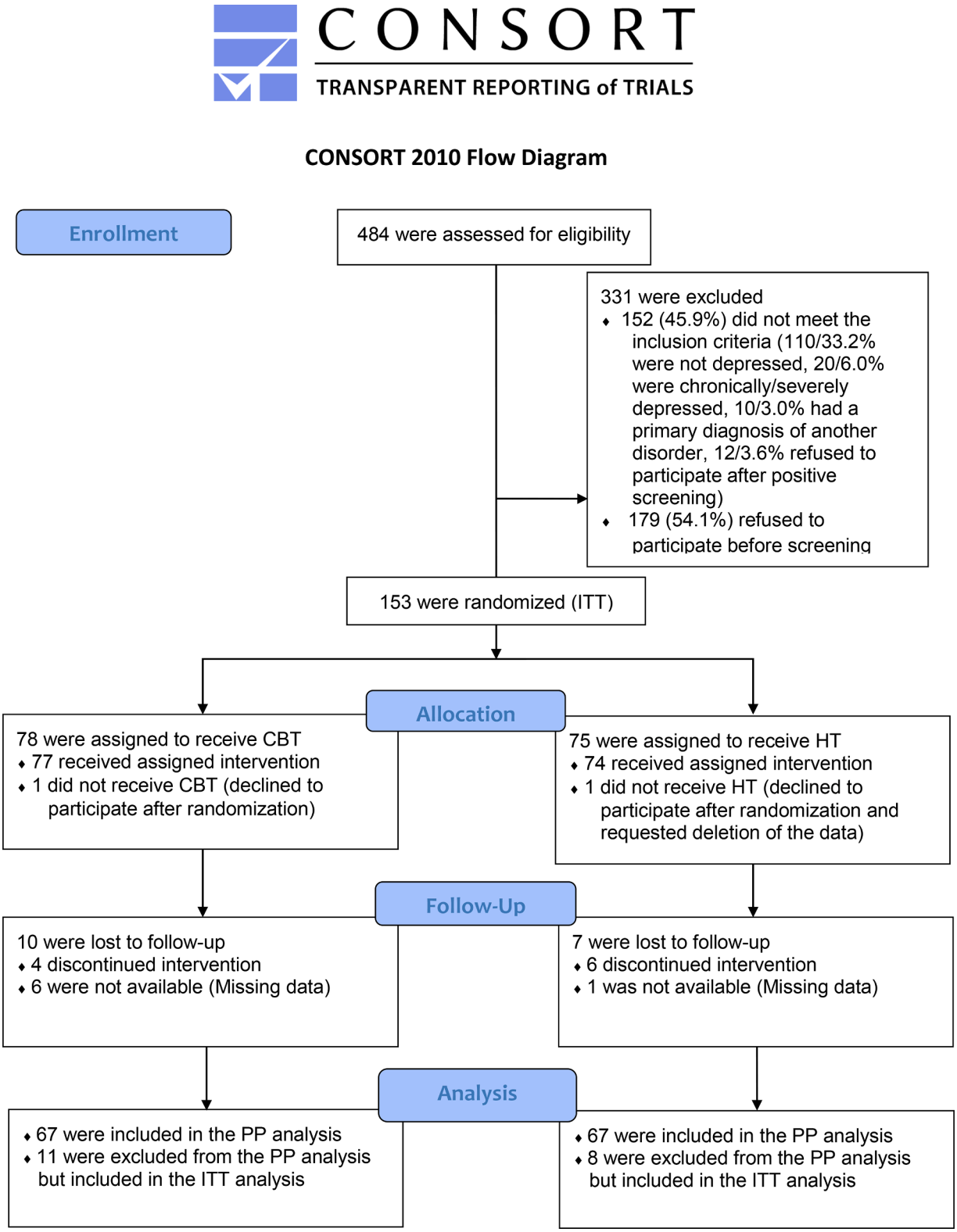

Fig. 1. Trial Enrollment, Randomization, and Follow-up

were at random, we generated five imputed data-sets based on a linear regression imputation algorithm automatically generated by SPSS. The aggregated data of the five imputations will be reported. As secondary analyses, we exploratively compared mean percentage symptom change of MADRS between baseline and six-months-follow-up (t3) as well as between baseline and twelve-months-follow-up (t4). We applied the same ITT analyses and non-inferiority test as for the primary endpoint.

As another endpoint that is derived from the primary endpoint, treatment response was defined as a symptom reduction of at least 50% in the MADRS between pre and post treatment and was compared between both treatment conditions using a Chi-Square test, including estimation for relative risk and 95% CI for relative risk according to Katz et al. (1978) (method C), which is implemented in the SPSS procedure crosstabs.

The statistical analysis was performed with SPSS (version 25), R Studio, and SAS (version 9.4). The authors and investigators of the current trial were blind concerning the primary outcome until the database was closed in November 2018.

3. Results

3.1. Sample Characteristics

In total 484 patients were interested in participation and were assessed for eligibility. After successful screening, 153 patients were included in the trial and randomized to either HT (n=75) or CBT (n=78). The dropout rate in both conditions was low (HT: n=8, 10.7%; CBT: n=11, 14.1%). Dropout was defined as declination to participate after randomization, discontinuation of the study treatment before 16 sessions, and missing data at the assessment after the end of the intervention. For details on the patient flow, see the CONSORT diagram in Fig. 1.

The characteristics of the ITT sample are displayed in Table 1, characteristics of the PP sample in the supplementary material. Since one of the patients declined to participate before the first session and requested the deletion of the data, the ITT sample consisted of n=152 patients. The PP sample was n=134 (11.8% dropout rate). Fifty-six (36.8%) of the patients were on antidepressant medication. Of those, four patients stopped antidepressant medication during treatment, six dropped out during the trial (no data at t2 available). All other patients continued medication as planned before or did never use antidepressant medication during the trial. Hypnotic suggestibility (score between 0-12 assessed with the Harvard Group Scale of Hypnotic Susceptibility) was on average in the low to high medium range (Bongartz, 1985). In total 11 patients revealed their treatment condition to the rater (five CBT and six HT patients). In all other 122 cases, raters were still unaware of the patient's treatment condition (1 missing data). Treatment preferences were assessed at the start of the study before randomization. Most of the participants had no treatment preference (40.8%). Some patients preferred CBT (31.6%) over HT (24.3%), see also Table 1.

At posttreatment, MADRS data for 134 patients were available, at t3 (six months later) for 84, at t4 (twelve months after posttreatment) for 99. Only 70 patients were available consequently at all follow-ups. Thus, the secondary analyses concerning percentage symptom change in MADRS from baseline to t3 and to t4 were only applied to the ITT sample using again MI with five imputed data-sets to replace missing data after assuring that missing data were at random.

3.2. Primary outcome

For the means and standard deviations of the MADRS in both groups, as well as the mean percentage symptom reduction, see Table 2. The means were in the range of moderate depression (20-34) at study entry and in the range for mild depression (7-19) after the end of treatment. The difference of the mean percentage reduction of depressive symptoms in the MADRS score between baseline and the end of treatment between the HT (M=39.8, SD=46.6) and CBT (M=35.8, SD=46.1) in the PP sample was 4.0 (95% CI -9.27 to 17.27). Concerning the pre-specified non-inferiority margin of -16.4 this result confirms the non-inferiority of HT compared to CBT. In the ITT sample, we found a similar result. HT (M=39.1%, SD=47.6) was also not inferior to CBT (M=36.3%, SD=46.6), (Δ=2.8, 95% CI -9.85 to 15.44) when replacing data with the MI method, both underlining the robustness of the primary analysis. Even when applying a non-inferiority margin of -9.9 (an effect size equal to 30% of the standard deviation of 32.9 instead of 50%), HT was still non inferior to CBT in the both the ITT and the CPP sample.

Table 1
Characteristics of the Trial Sample (ITT, *n*=152)

Variables	CBT (*n*=78)	HT (*n*=74)	Total (*n*=152)
Age, mean (SD)	37.6 (14.5)	41.5 (14.4)	39.5 (14.6)
Hypnotic susceptibility[1], mean (SD)	6.2 (2.3)	5.7 (2.5)	6.0 (2.4)
Sex, No. (%) Female	51 (65.4)	49 (66.2)	100 (65.8)
Highest educational level, No. (%)			
High school or higher	62 (79.5)	52 (70.3)	114 (75.0)
No high school degree	16 (20.5)	22 (29.7)	38 (25.0)
Antidepressant medication (AD), No. (%)	25 (32.1)	31 (41.9)	56 (36.8)
SSRI	10 (12.8)	13 (17.6)	23 (15.1)
SNRI	6 (7.7)	3 (4.1)	9 (5.9)
Tetracyclic AD	4 (5.1)	5 (6.8)	9 (5.9)
Tricyclic AD	3 (3.8)	0 (0.0)	3 (2.0)
NDRI	1 (1.3)	3 (4.1)	4 (2.6)
Phytomedical AD	1 (1.3)	7 (9.5)	8 (5.3)
MD Subtype, No. (%)			
Recurrent	53 (67.9)	64 (86.5)	117 (77.0)
Single episode	25 (32.1)	10 (13.5)	35 (23.0)
Comorbidity, No. (%)	32 (41.0)	36 (48.6)	68 (44.7)
Current	17 (21.8)	23 (31.1)	40 (26.3)
PDD	10 (12.8)	7 (9.5)	17 (11.2)
Anxiety disorder	7 (9.0)	12 (16.2)	19 (12.5)
Personality disorder	0 (0.0)	4 (5.4)	4 (2.6)
Only lifetime	15 (19.2)	13 (17.6)	28 (18.4)
Substance abuse	5 (6.4)	3 (4.1)	8 (5.3)
Eating disorder	3 (3.8)	2 (2.7)	5 (3.3)
PDD	5 (6.4)	7 (9.5)	12 (7.9)
other	2 (2.6)	1 (1.4)	3 (2.0)
Treatment preferences[2], No. (%)			
No preference	34 (43.6)	28 (37.8)	62 (40.8)
Preferring HT	18 (23.1)	19 (26.7)	37 (24.3)
Preferring CBT	24 (30.8)	24 (32.4)	48 (31.6)

HT=Hypnotherapy; CBT=Cognitive Behavioral Therapy; SSRI = Selective serotonine reuptake inhibitor; SNRI = Serotonine- norepiphrene reuptake inhibitor; NDRI = Norepiphrene-dopamine reuptake inhibitor; MD=Major Depression; PDD = Persistent Depressive Disorder;

[1] assessed with the Harvard Group Scale of Hypnotic Susceptibility with CBT: *n*=75 and HT: *n*=69.

[2] with *n*=5 missing data.

3.3. Secondary outcomes

The difference in the mean percentage reduction of depressive symptoms in the MADRS score between baseline (t1) and the first follow-up six months after the end of the treatment (t3) in the ITT sample between the HT (M=46.35, SD=49.09) and CBT (M=47.19, SD=53.98) was -0.84 (95% CI -14.71 to 13.03). The difference in the mean percentage reduction of depressive symptoms in the MADRS score between baseline (t1) and the second follow-up twelve months after the end of the treatment (t4) in the ITT sample between the HT (M=54.37, SD=42.81) and CBT (M=50.36, SD=42.60) was 4.01 (95% CI -7.46 to 15.48). Concerning the pre-specified non-inferiority margin of -16.4 both results support the non-inferiority of HT compared to CBT, also for the follow-ups. Means and standard deviations are displayed in Table 2.

Concerning the treatment response rates we found a higher relative proportion of patients with a reduction of at least 50% in the MADRS

Table 2
Primary and secondary outcomes in the PP and ITT samples

Variables	PP sample (n=134) CBT (*n*=67)	HT (*n*=67)	Total (*n*=134)	Statistics
MADRS pre	20.2 (5.7)	21. 8 (5.3)	21.0 (5.5)	
MADRS post	12.8 (9.1)	12.8 (9.9)	12.8 (9.5)	
MADRS percentage improvement pre post	35.8 (46.1)	39.8 (46.6)	37.8 (46.2)	Δ=4.0 (95%CI -9.27 - 17.27)
	ITT sample (n=152, MI)			
	CBT (*n*=78)	**HT (*n*=74)**	**Total (*n*=152)**	**Statistics**
MADRS pre	20.4 (5.6)	21.8 (5.3)	21.1 (5.5)	
MADRS post	12.8 (9.1)	13.0 (10.0)	12.9 (9.5)	
MADRS t3	10.6 (9.3)	11.7 (10.3)	11.2 (9.8	
MADRS t4	10.5 (9.1)	10.4 (9.1)	10.5 (9.1)	
MADRS percentage improvement pre post	36.3 (46.6)	39.1 (47.6)	37.7 (47.0)	Δ=2.8 (95%CI -9.85 to 15.44)
MADRS percentage improvement pre/t3	47.2 (54.0)	46.4 (49.1)	46.8 (51.6)	Δ=-0.84 (95%CI -14.71 to 13.03).
MADRS percentage improvement pre/t4	50.4 (42.6)	54.4 (42.8)	52.4 (42.7)	Δ = 4.01 (95%CI -7.46 to 15.48)

Data expressed as means (standard deviation). MADRS=Montgomery Asperg Depression Scale, HT=Hypnotherapy; CBT=Cognitive Behavioural Therapy, PP = per protocol; ITT = intention to treat; MI = multiple imputation (5 datasets); t3 = six months after end of treatment; t4 = 12 months after end of treatment.

between pre and post for HT (n=33, 44.6%) compared to CBT (n=30, 38.5%) in the ITT sample (RR=1.16, 95% CI 0.79 to 1.69). In the PP sample, the proportion was 49.3% in HT and 44.8% in CBT (RR=1.10, 95% CI 0.77 to 1.58). At post treatment, n=23 (31.1%) of HT and n=18 (23.1%) of CBT patients in the ITT sample showed MADRS scores in the symptom free range (0-6), thus considered to be in a remission (RR=1.35, 95% CI 0.79 to 2.29). In the PP sample, the numbers were 34.3% in HT and 26.9% in CBT (RR=1.28, 95% CI 0.76 to 2.14).

3.4. Post hoc analyses

Exploratively, we calculated Cohen's d effect sizes in both samples. Depressive symptoms between pre and post improved in the CBT patients with an effect of d= -1.14 (95% CI -1.49 to -0.79) and in HT patients with d= -1.41 (95% CI -1.77 to -1.05) in the PP sample. The effect sizes of the difference between CBT and HT in the MADRS post score was almost zero (d<.001, 95% CI -0.48 to 0.48). In the ITT sample, we found that the symptoms in the CBT patients improved from pre to post with an effect size of d= -1.17 (95% CI -1.51 to -0.83) and in the HT patients with about d= -1.37 (95% CI -1.73 to -1.01). The effect size of the difference in the post MADRS scores between both groups was d= 0.02 (95% CI -0.30 to 0.34).

3.5. Safety

There are no established standards regarding the documentation and assessment of potential side effects or adverse events in psychotherapy trials. We also considered negative developments of the illness (suicidality as a clinical symptom of depression or clinical deterioration) but also changes in the personal, professional, or social functioning (e.g. new job, move to a new city, separation of the partner, death of a family member) of the patient as potential side effects (SEs) or AEs additionally to the commonly used SAEs in drug trials (Linden, 2013). A total rate of 97 SEs/AEs was reported for 56 patients from the trial from pre to post treatment with a range of 1-9 events per patient, see also Table 3. Overall, 37 of those were serious adverse events (SAEs) that occurred in 16 patients, six CBT (7.7%) and ten HT patients (13.5%). Moreover, there were nine SAEs in the CBT and 28 SAEs in the HT patients. Only in seven cases, SAEs occurred that are commonly used in drug trials (hospitalization). In seven cases, the SEs/AEs were reported as potentially treatment-related (six of them received HT). Only one SAE was rated as being potentially treatment-related: One HT-patient reported suicidal thoughts during therapy. However, the patient was credibly distant from committing suicide and could continue study treatment.

4. Discussion

To our knowledge, this is the first study to compare the efficacy of hypnotherapy in mild to moderate MD with that of evidence-based psychotherapy. More than that, we provide safety data on both CBT and HT.

HT showed to be not inferior to CBT in the mean percentage symptom reduction according to the MADRS from baseline to the end of treatment. The results of the follow-ups support the non-inferiority of HT compared to CBT up to twelve months after the end of the treatment. Another trial comparing psychodynamic treatment to CBT reported a non-inferiority margin of Cohen's d = 0.30 (Driessen et al., 2013). Thus, we decided to run a post hoc sensitivity analysis assuming a smaller non-inferiority margin of 0.3 instead of the a-priori planned margin of 0.5. The results support the non-inferiority of HT compared to CBT also

Table 3
Safety results

Variables	CBT (*n*=78)	HT (*n*=74)	Total sample (*n*=152)
Adverse events (AEs)/ Side effects (SEs)			
97 AEs/SEs were reported in…	36 AEs in 26 pat. (33.3%,1-5 pp)	61 AES in 30 pat. (40.5%,1-9 pp)	97 AEs in 56 pat. (36.8%,1-9 pp)
60 were AEs/SEs[b] in…	27 AEs in 22 pat. (28.2%,1-2 pp)	33 AEs in 24 pat. (35.1%, 1-3 pp)	62 AEs in 48 patients (31.6%,1-3)
24 with new symptoms	11111 AEs in 10 pat.	13 AEs in 97 pat.	24 AEs in 19 pat.
16 with deterioration of depression	5 AEs in 4 pat.	11 AEs in 9 pat.	16 AEs in 13 pat.
7 with changes in social life	5 AEs in 4 pat.	2 AEs in 2 pat.	7 AEs in 6 pat.
10 with changes in professional life	3 AEs in 3 pat.	7 AEs in 7 pat.	10 AEs in 10 pat.
6 other events	2 AEs in 2 pat.	4 AEs in 4 pat.	6 AEs in 6 pat.
Serious adverse events (SAEs)			
37 were SAEs[1] in…	9 SAEs in 6 pat. (7.7%, 1-3 pp)	28 SAEs in 10 pat. (13.5%,1-9 pp)	37 SAEs in 16 pat. (11.2%,1-9 pp)
2 with serious deterioration of depression	1 SAE in 1 patient	1 SAE in 1 patient	2 SAEs in 2 pat.
29 with suicidal thoughts	8 SAEs in 5 pat.	21 SAEs in 6 pat.	29 SAEs in 11 pat.
2 with hospitalization because of physical conditions	0	2 SAEs in 2 pat.	2 SAEs in 2 pat.
6 with hospitalization because of the MD	1 SAE in 1 pat.	5 SAEs in 4 pat.	6 SAEs in 5 pat.

HT=Hypnotherapy; CBT=Cognitive Behavioral Therapy.; 97 overall AEs were reported in 56 patients. In 96 patients, no (S)AE was reported; Pat. = Patient; pp = per patient. Most of the patients had more than one (S)AE.

[1] SAE could consist of one or two symptoms (e.g. deterioration of depression, suicidal thoughts).

[b] AE could consist of one or two symptoms (e.g. deterioration of depression, changes in social life).

for an alternative endpoint with a more strict non-inferiority margin. Furthermore, the complementary within-group effect sizes of both treatments were very large with Cohen's d > .80, whereas between-group effect sizes at the end of the treatment seemed to be small with Cohen's d < 0.03. Results were thus similar to another trial comparing CBT and Interpersonal Psychotherapy (IPT) in the treatment of depression (Lemmens et al., 2015). Nevertheless, future research should definitely replicate results in a larger trial and investigate differential response for HT as in Nemeroff et al. (2003).

Within the primary outcome of mean percentage symptom reduction, we found a high variation in our sample. Overall symptom reduction was 38% and rates of patients with more than 50% symptom reduction were only about 30%. This result also indicated that there was still a considerable number of patients without improvement or even symptom deterioration. The low symptom improvement and small remission rates may have been due to the fact that we randomly assigned patients to two very different psychotherapeutic approaches like in the study that compared CBT with psychodynamic psychotherapy (Driessen et al., 2013). Remission rates in the study of Driessen et al. (2013) were comparable to the results of our trial. Even if patients agreed to be randomly assigned, and although less than one third of patients actually preferred a treatment with HT in our trial, the expectations regarding CBT and HT might be very different and could have reduced the efficacy and internal validity of the overall trial. This is supported by a study which demonstrated that higher symptom improvement was found in patient-preference matches, compared to a mismatch (Kwan et al., 2010).

Techniques of HT focused mainly on emotional processing of experiences associated with the depression. Results of brain-imaging studies indicate that somatic and emotional control can be achieved during hypnosis in healthy subjects by an increased connectivity between the dorsolateral prefrontal cortex and the insula (Jiang et al., 2017). In CBT, an increase of prefrontal functioning was found which was interpreted as an increase of cognitive control (Karlsson, 2011). We therefore hypothesize that different mechanisms are addressed in HT compared to CBT. Future studies should investigate if different brain networks are associated with a positive treatment outcome of HT or CBT.

4.1. Limitations

Although confirming the non-inferiority of HT, it can only be indirectly concluded that both of the offered treatments were effective, as a control group without treatment was waived for ethical reasons. Compared to Luty et al. (2007) who reported a 55% improvement (mean) in symptoms after 16 weeks treatment with either CBT or Interpersonal Psychotherapy, we only found a mean improvement of 38% in the whole sample. The smaller symptom improvement might be related to the circumstance that the patients in the trial of Luty et al. (2007) were more severely depressed and therefore had a higher probability to improve during treatment, whereas we had excluded patients with severe depressive symptoms. Another limitation of the study can be certainly seen in the artificial limitation of the number of sessions to 20 appointments. In the real-world therapeutic setting in Germany, longer courses of treatment are often planned even for CBT, with individual reference to the symptoms of the patients. The number of about 16-20 sessions in a six months period though was comparable with the trials of Driessen et al. (2013) and Lemmens et al. (2015) and with the 16-weeks treatment period of Alladin and Alibhai (2007). However, dropout rate was very low. Given the manualized and standardized planning, the study is nevertheless able to provide information about the comparable efficacy of the two therapeutic methods based on parallelized treatment length and dose. Future studies though should first confirm these findings and secondly examine whether HT might reduce treatment length and costs. Treatment fidelity might have been enhanced by a constant feedback loop between raters of fidelity, supervisors, and therapists whenever ratings of fidelity were completed. We found a higher rate of SAEs in HT compared to CBT. Most of the SAEs were related to the depression (suicidal thoughts or worsening of symptoms). Positively, not many cases reported those kinds of SAEs that are defined for pharmacological studies (e.g. hospitalization). However, our rates of SAEs occurred in 7.7% and 13.5% of the cases which is similar to the rates of 2-15% discussed in Linden (2013) and thus are also comparable to pharmacological trials. The high rate of depression-related SAEs especially in HT patients could be explained by the fact that some of the HT modules focused on childhood memories and thus could have activated negative feelings. Unfortunately, we did not use additional measures of potential negative effects of psychotherapy (Herzog et al., 2019). Further studies on HT should evaluate risks and side effects compared to other psychological treatments that also focus on past experiences, such like CBASP.

Beyond the limitations, this study has a number of methodological strengths: it is characterized by a manualized therapeutic procedure and included a methodical planning analogous to a pharmacotherapy study-setting: In addition to randomized allocation, external monitoring and blinded evaluation, the evaluations were carried out on both an ITT and a PP sample.

The results may be interpreted as a first indication that HT might be a psychotherapeutic method that extends the number of available therapies in the treatment of mild to moderate MD. This study should be considered as a first clinical indication of the applicability of HT in patients with mild to moderate depressive episodes. The findings accomplish results of a previous trial (Alladin and Alibhai, 2007) which found additive effects for HT combined with CBT compared to CBT-alone. Further studies should be necessary to confirm the results and may also provide data regarding a differential indication for either therapeutic approach. Suggestibility, patient preferences, age and gender, number and length of episodes, but most of all subset of symptoms or etiologically significant factors might be considered as predictors of outcome, depending on the treatment method. Baseline characteristics and treatment moderators could help to assign patients to the more promising treatment. As Wallace et al. (2013) described, the interpretation of effectiveness should take into account that different subsets of populations can react differently to treatment conditions. Further analysis will help us to identify potential predictors and moderators for treatment response. Furthermore, the expectancy of patients regarding the specific treatment and its efficacy on treatment adherence and results should be evaluated.

5. Data Statement

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

Contributors

AB (Principal Investigator) wrote the manuscript and the study protocol and received the grant. CM (Statistician) was involved in the sample size, the power calculation, the management of the external database, the randomization procedure, the statistical analysis plan, and the data analysis. KF (Coordinator) finalized the study protocol, analyzed the data, and wrote a first draft and revision of the manuscript. AB, CM, and KF had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

CS, AB, BC, JH, JJ, MP, AS, and CZ were all involved in conducting the study treatment (Therapists) and gave feedback to the current form of the manuscript.

All authors have approved the final article.

Declaration of Competing Interest

AB (the last author) has received the grant from the Milton Erickson

Gesellschaft (MEG; Milton H. Erickson Society for Clinical Hypnosis) for this study. AB has received honorary (altogether less than 1000 €) and travel reimbursements for talks at the meetings of the MEG, AB is responsible for education in cognitive behavioral therapy.

KF has received travel reimbursements for talks at the meetings of the MEG.

CS is a member of the Milton Erickson Gesellschaft (MEG; Milton H. Erickson Society for Clinical Hypnosis.

CM, AB (third author), BC, JH, JJ MP, AS and CZ have no conflicts of interest to declare.

Role of the Funding Source

This work was supported by the Milton Erickson Gesellschaft (MEG), Waisenhausstraße 55, 80637 Munich, Germany. The sponsor was not involved in the study design, the preparation and conduction of the trial, the preparation and analysis of date, and the manuscript.

Acknowledgements

We want to thank all the patients who participated in our study. Further, we thank Clara Huber, Theresa Trauner, Johanna Schreiber, Debora Gerdemann, Julia Nannt, and Katharina Laux for helping us with the blind follow-up interviews, and the data entry and management of the current study.

We thank Dr. Colleen Dockery-Alami for proofreading the manuscript.

Previous oral presentation: Annual meeting of Milton H. Erickson Society for Clinical Hypnosis, Bad Kissingen, Germany, 21st March 2019

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2021.02.069.

References

Alladin, A., 2009. Evidence-based hypnotherapy for depression. Contemp. Hypn. 26 (4), 245–262. https://doi.org/10.1002/ch.391.

Alladin, A., Alibhai, A., 2007. Cognitive hypnotherapy for depression: An empirical investigation. Int. J. Clin. Exp. Hypn. 55, 147–166. https://doi.org/10.1080/00207140601177897.

American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5®). American Psychiatric Pub.

Beck, A.T., 2002. Cognitive models of depression. In: Leahy, R.L., Dowd, E.T. (Eds.), Clinical advances in cognitive psychotherapy: Theory and application. Springer Publishing Company, New York, pp. 29–61.

Bongartz, W., 1985. German Norms for the Harvard Roup Scale of Hypnotic Susceptibility, Form A. Int. J. Clin. Exp. Hypn. 3 (2), 131–139. https://doi.org/10.1080/00207148508406643.

Driessen, E., Van, H.L., Don, F.J., Peen, J., Kool, S., Westra, D., Hendriksen, M., Schoevers, R.A., Cuijpers, P., Twisk, J.W.R., Dekker, J.J.M., 2013. The Efficacy of Cognitive-Behavioral Therapy and Psychodynamic Therapy in the Outpatient Treatment of Major Depression: A Randomized Clinical Trial. Am. J. Psychiatry. 170 (9), 1041–1050. https://doi.org/10.1176/appi.ajp.2013.12070899.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1997. Structured Clinical Interview for DSM-IV Axis I disorders (SCID I). Biometric Research Department, New York, New York.

Fuhr, K., Schweizer, C., Meisner, C., Batra, A., 2017. Efficacy of hypnotherapy compared to cognitive-behavioural therapy for mild-to-moderate depression: study protocol of a randomised-controlled rater-blind trial (WIKI-D). BMJ Open 7 (11), e016978. https://doi.org/10.1136/bmjopen-2017-016978.

Global Health Estimates, 2016. Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2015. World Health Organization, Geneva.

Hautzinger, M.Kognitive, 2013. Verhaltenstherapie bei Depressionen, 7th ed. Beltz Verlag, Weinheim.

Herzog, P., Lauff, S., Rief, W., Brakemeier, E.L., 2019. Assessing the unwanted: A systematic review of instruments used to assess negative effects of psychotherapy. Brain. Behav. 9 (12), e01447. https://doi.org/10.1002/brb3.1447.

Hollon, S.D., DeRubeis, R.J., Shelton, R.C., Amsterdam, J.D., Salomon, R.M., O'Reardon, J.P., Lovett, M.L., Young, P.R., Haman, K.L., Freeman, B.B., Gallop, R., 2005. Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. Arch. Gen. Psychiat. 62 (4), 417–422. https://doi.org/10.1001/archpsyc.62.4.417.

Jiang, H., White, M.P., Greicius, M.D., Waelde, L.C., Spiegel, D., 2017. Brain activity and functional connectivity associated with hypnosis. Cerebral cortex 27 (8), 4083–4093. https://doi.org/10.1093/cercor/bhw220.

Karlsson, H., 2011. How Psychotherapy changes the Brain. Psychiat. Times. 28 (8), 21–23.

Katz, D., Baptista, J., Azen, S.P., Pike, M.C., 1978. Obtaining confidence intervals for the risk ratio in cohort studies. Biometrics 469–474. https://doi.org/10.1002/bimj.4710370808.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, J.A., Walters, E.E., Wang, P.S., National Comorbidity Survey Replication, 2003. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 289 (23), 3095–3105. https://doi.org/10.1001/jama.289.23.3095.

Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., Brown, W.A., 2012. A systematic review of comparative efficacy of treatments and controls for depression. PLoS One 7 (7), e41778. https://doi.org/10.1371/journal.pone.0041778.

Kwan, B.M., Dimidjian, S., Rizvi, S.L., 2010. Treatment preference, engagement, and clinical improvement in pharmacotherapy versus psychotherapy for depression. Behav. Res. Ther. 48 (8), 799–804. https://doi.org/10.1016/j.brat.2010.04.003.

Lemmens, L.H.J.M., Arntz, A., Peeters, F., Hollon, S.D., Roefs, A., Huibers, M.J.H., 2015. Clinical effectiveness of cognitive therapy v. interpersonal psychotherapy for depression: results of a randomized controlled trial. Psychol. Med. 45 (10), 2095–2110. https://doi.org/10.1017/S0033291715000033.

Lewinsohn, P.M., 1974. A behavioral approach to depression. In: Coyne, J.C. (Ed.), Essential papers on depression. University Press, New York, pp. 150–172.

Linden, M., 2013. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. Clin. Psychol. Psychot. 20 (4), 286–296. https://doi.org/10.1002/cpp.1765.

Luty, S.E., Carter, J.D., McKenzie, J.M., Rae, A.M., Frampton, C.M., Mulder, R.T., Joyce, P.R., 2007. Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression. Br. J. Psychiatry. 190, 496–502. https://doi.org/10.1192/bjp.bp.106.024729.

Milling, L.S., Valentine, K.E., McCarley, H.S., LoStimolo, L.M., 2019. A Meta-Analysis of Hypnotic Interventions for Depression Symptoms: High Hopes for Hypnosis? Am. J. Clin. Hypn. 61 (3), 227–243. https://doi.org/10.1080/00029157.2018.1489777.

Montgomery, S.A., Asberg, M., 1979. A new depression scale designed to be sensitive to change. The Brit. J. Psychiat. 134 (4), 382–389. https://doi.org/10.1192/bjp.134.4.382.

National Collaborating Centre for Mental Health (NICE), 2016. Depression: The treatment and management of depression in adults. National Clinical Practice Guideline 90. The British Psychological Society & The Royal College of Psychiatrists, London. https://www.nice.org.uk/guidance/cg90. Accessed January 18, 2017.

Nemeroff, C.B., Heim, C.M., Thase, M.E., Klein, D.N., Rush, J.A., Schatzberg, A.F., Ninan, P.T., McCullough, J.P., Weiss, P.M., Dunner, D.L., Rothbaum, B.O., Kornstein, S., Keitner, G., Keller, M.B., 2003. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. PNAS 100 (24), 14293–14296. https://doi.org/10.1073/pnas.2336126100.

Rudolf, G., 2013. Strukturbezogene Psychotherapie – Leitfaden zur psychodynamischen Therapie struktureller Störungen, 3th ed. Schattauer, Stuttgart.

US Food and Drug Administration, 2016. Non-inferiority clinical trials to establish effectiveness: guidance for industry. US Food and Drug Adminstration, Silverspring, MD.

Wallace, M.L., Frank, E., Kraemer, H.C., 2013. A Novel Approach for Developing and Interpreting Treatment Moderator Profiles in Randomized Clinical Trials. JAMA Psychiatry 70 (11), 1241–1247. https://doi.org/10.1001/jamapsychiatry.2013.1960.

Whorwell, P.J., Prior, A., Faragher, E.B., 1984. Controlled trial of hypnotherapy in the treatment of severe refractory irritable-bowel syndrome. Lancet 324 (8414), 1232–1234. https://doi.org/10.1016/S0140-6736(84)92793-4.

Wiersma, J.E., Van Schaik, D.J.F., Hoogendorn, A.W., Dekker, J.J., Van, H.L., Schoevers, R.A., Blom, M.B.J., Maas, K., Smit, J.H., McCullough, J.P., Beekman, A.T.F., Van Oppen, P., 2014. The Effectiveness of the Cognitive Behavioral Analysis System of Psychotherapy for Chronic Depression: A Randomized Controlled Trial. Psychother. Psychosom. 83, 263–269. https://doi.org/10.1159/000360795.

Wilhelm-Goessling, C., Schweizer, C., Duerr, C., Fuhr, K., Revenstorf, D., 2020. Hypnotherapie bei Depressionen. Ein Manual für Psychotherapeuten. Kohlhammer Verlag, Stuttgart.