Effectiveness of Schema Therapy versus Cognitive Behavioral Therapy versus Supportive Therapy for Depression in Inpatient and Day Clinic Settings: A Randomized Clinical Trial

Johannes Kopf-Beck, Celina L. Müller, Jeanette Tamm, Julia Fietz, Nils Rek, Leah Just, Zoe Ilona Spock, Katharina Wewecke, Keisuke Takano, Martin Rein, Martin E. Keck, Samy Egli

Max Planck Institute of Psychiatry, Munich, Germany; Department of Psychology, LMU Munich, Munich, Germany; International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany; Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK; Human Informatics and Interaction Research Institute (HIIRI), National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan; Oberberg Tagesklinik München Bogenhausen, Munich, Germany

Keywords
- Schema therapy
- Cognitive behavioral therapy
- Depression
- Psychotherapy research
- Randomized controlled trial

Abstract

Introduction: Schema therapy (ST) reduces depressive symptoms, but clinical trials have not investigated its effectiveness for patients suffering from severe forms of depression and high rates of comorbidities. There is high demand for exploring and improving treatments for this patient group. The objective of the current study was to evaluate whether ST is more effective than individual supportive therapy (IST) and noninferior compared with cognitive behavioral therapy (CBT) in treating depression.

Methods: For this clinical trial, medicated patients were recruited in inpatient and day clinic settings. The major inclusion criteria were age between 18 and 75 years and primary diagnosis of depression without psychotic symptoms. A total of 292 participants were randomized to ST, CBT, or IST and received 7 weeks of psychotherapy (up to 14 individual and 14 group sessions). The primary outcome was change in depression severity after treatment measured by Beck Depression Inventory-II. Primary test for efficacy was superiority of ST over IST. Secondary test was noninferiority of ST compared with CBT. Multilevel modeling was conducted. The results at 6-month follow-up were explored.

Results: Across treatment, ST was not superior to IST. Secondary outcome analyses and completer analyses showed similar results. However, ST showed clinically relevant noninferiority compared with CBT.

Conclusion: ST for depression as part of a psychiatric care program showed clinical noninferiority compared to CBT, without being superior to IST. ST represents a potentially useful addition to the therapeutic repertoire for the treatment of depression but its efficacy, including long-term efficacy, should be evaluated further.
Introduction

Depressive disorders are highly prevalent around the globe [1] and are one of the leading causes for years lived with disability [2]. They are associated with increased mortality [3], individual suffering, and massive economic costs [4].

Besides antidepressant medication, psychotherapy belongs to the most important strategies for treating depressive disorders and has shown its overall effectiveness across different treatment types [5]. Nevertheless, the high rates of treatment resistance [6], relapse [7], chronic course [8], and the complexity of depression regarding symptom heterogeneity [9] and comorbidities like anxiety disorders [10] or personality disorders [11] call for further development of psychotherapeutic approaches and concepts.

The newly developed treatments follow integrative approaches and address the heterogeneity of depression. A particularly promising one is the schema therapy (ST), originally developed for nonresponders of traditional cognitive therapy, chronic courses of disorders like depression, and patients suffering from personality disorders [12]. Recently, ST has been extended to several other disorders, such as anxiety disorders [13], eating disorders [14], and post-traumatic stress disorders [15]. A growing body of research supports ST’s overall effectiveness, albeit the vast majority of randomized controlled trials (RCTs) on ST are focused on personality disorders [16]. Nevertheless, the first evidence suggests that ST also reduces depressive symptoms [17–19].

Besides behavioral and cognitive strategies, ST applies emotion-focused techniques and focuses on therapeutic alliance [12]. It addresses the so-called early maladaptive schemas and early learning experiences with the aim to modify them to fulfill emotional core needs.

Referring to cognitive perspectives on schemas and depression [20], ST concepts postulate that early maladaptive schemas are generally linked with the development of psychopathological symptoms [12]. Beyond that, they are considered to be a risk factor for the emergence, maintenance, and chronification of depression [21–23]. Since ST addresses early maladaptive schemas as the central treatment target, we assume that it is particularly useful and effective in the treatment of severe forms of depression, which often go along with high rates of comorbidities (e.g., personality disorders) [11] and typically occur in psychiatric inpatient and day clinic settings. However, empirical evidence from RCTs investigating the clinical effectiveness of ST for this patient group and setting is missing.

The present study aimed to test the effectiveness of a high-intensity ST for severe forms of depression (primary diagnosis) in inpatient and day clinic settings. Since ST addresses fundamental processes responsible for the emergence and maintenance of depression, we hypothesized that ST would be superior to individual supportive therapy (IST; i.e., nonspecific supportive therapy) in reducing depression symptoms (hypothesis H1) [24].

ST has further advantages over other treatments, such as high acceptance among patients [16]. Since statistically significant differences in treatment effects are not necessarily clinically relevant [25], we additionally tested ST against the first-line treatment of severe depression, cognitive behavioral therapy (CBT) [26]. A previous research has revealed comparable effectiveness of ST and CBT in the treatment of depression when it was applied as 1-year treatment in an outpatient setting [17]. Thus, we assumed noninferiority of high-intensity ST compared with CBT regarding the decrease of depression at the end of treatment (hypothesis H2). The noninferiority margin was based on empirical estimations of what can be considered as the minimal clinically important difference in depression in the form of percentage change from baseline [27].

All therapy approaches in the current study were applied in psychiatric inpatient and day clinic settings as a 7-week high-intensity-combined single-group program [28]. Long-term effectiveness at 6-month follow-up was explored. The trial reports followed methodological recommendations for psychotherapy trials [29] (for CONSORT check list see online suppl. Supplement 4; for all online suppl. material, see https://doi.org/10.1159/000535492).

Materials and Methods

Study Design and Participants

The OPTIMA trial (OPtimized Treatment Identification at the MAx Plank Institute) was conducted as monocentric, rater-blinded, prospective, parallel-group, block-RCT with a nonspecific (IST) and specific active control condition (CBT) at the psychiatric clinic of the Max Planck Institute of Psychiatry in Munich, Germany [30]. Minor subsets of the trial data have been used in prior analyses, yet focusing on other aspects than treatment differences [31, 32].

Patients aged between 18 and 75 were recruited between September 2017 and October 2020. Inclusion criteria were a primary diagnosis of depression (moderate or severe) without psychotic symptoms represented by ICD-10 diagnoses (F32.1, F32.2, F33.1, or F33.2) and corresponding to a Beck Depression Inventory (BDI-II) score ≥20 [33] or Montgomery-Åsberg Depression Rating Scale (MADRS) score ≥20 [34]. Depression was
diagnosed by clinical assessment and by the Munich-Composite International Diagnostic Interview (M-CIDI) [35], a computerized, fully standardized German version of the World Mental Health Composite International Diagnostic Interview [36]. Exclusion criteria were psychotic symptoms (F32.3, F33.3), acute suicidality, lifetime history of any psychotic or bipolar disorder, severe neurological or internal concomitant or past diseases, an IQ <80 and/or severe learning disability, current alcohol or any illicit drug withdrawal syndrome, concomitant organic mental disorder, severe mutism or stupor, mental disorder secondary to a medical condition or substance use disorder, pregnancy and lactation period, and missing eligibility for psychotherapy due to language barriers. The Institutional Ethics Committee of LMU, Munich approved the study. Written informed consent was provided by all participants prior inclusion.

Randomization and Masking
Participants were randomized to the 3 treatment conditions ST, CBT, and IST with an overall 1:1:1 allocation ratio (applied to the complete study center over the complete course of recruitment time) using computer-generated numbers in a block randomization technique (block size: 24) stratified by treatment units (wards or day clinics). Such a three-arm trial including an active control treatment (IST) and a first-line treatment (CBT) allows for answering different research questions and additionally for further sensitivity analysis [37]. The randomization sequence was administered by a study coordinator who was not involved in recruitment, therapy delivery, or assessment procedures. Involved therapists received allocation information after eligibility was assured and baseline assessments were finalized. All interviews and clinical ratings were conducted by raters blind to treatment condition. In case of unintentional unblinding, ratings were conducted by another rating team member. Inter-rater reliability of raters was assessed regularly and additional training was administered as needed.

Interventions
All treatments followed manualized protocols and were designed as 14 twice-weekly group- and 14 twice-weekly single-sessions, resulting in a potential maximum of 28 sessions. The applied ST manual [38] integrates cognitive and behavioral techniques from approaches like gestalt therapy, psychodynamic therapy, and ego-state therapy [12]. It addresses changing and distancing schemas by using so-called mode dialogues on chairs, imagery rescripting, and limited reparenting to form therapeutic alliance as key tools. ST applies emotion-focused and experimental techniques and focuses on the unmet core needs to overcome psychopathology. In the manual used, techniques were combined in a semi-structured way including 3 different phases: (1) an initial exploration phase to familiarize the patient with the concept (psychoeducation about schemas, modes, basic psychological needs, and emotions), to explore the patients’ schemas and modes in individual conversations, e.g., by discussing the results of questionnaires, and to establish therapy goals and therapeutic alliance (e.g., through validation of emotions, self-disclosure of the therapist, and empathic confrontation within the framework of limited reparenting); (2) a second modification phase using the above techniques (mode dialogues on chairs and imagery rescripting) to modify and distance early maladaptive schemas; and (3) a third transfer and relapse prevention phase to consolidate previous gains, including learning how to transfer gains from psychotherapy settings to everyday social life. For details, see the full manual [38].

For the CBT condition, we adapted and extended widely used treatment strategies based on Beck’s theory of depression [20, 39]. The manual included 5 highly structured phases dealing with (1) psycho-education (symptoms and risk factors of the onset and maintenance of depression, individual pathogenesis model), (2) behavioral activation (self-observation, mood diary, identification and planning of positive activities, self-reinforcement strategies), (3) cognitive modification (modification of dysfunctional attitudes), (4) social competence training (psychoeducation on competent social behavior, role plays for difficult social situations such as dealing with relationship difficulties, gaining sympathy, or coping with the disorder), and (5) relapse prevention (early warning signals of depression, coping strategies, crisis plan, etc.).

IST was provided as a third, nonspecific, nonstructured, manualized active control condition based on the common factors of psychotherapy [40]. Comparable approaches have been used in previous studies [24, 41]. Thus, IST referred to 3 main principles: (1) to activate the patient’s individual resources in the conversation and reinforce pre-existing competencies, strengths, positive attitudes, etc.; (2) to foster and use the therapeutic alliance by validating the patient and applying the principles of unconditional positive regard, authenticity, and congruence; and (3) to encourage the patient to acknowledge, verbalize, and give space to emotions by helping the patient to identify and specify emotions and by validating and normalizing them. Therapists were instructed not to structure the treatment and not to use CBT or ST techniques, but to apply a supportive, nonjudgmental, and empathic communication style.

Procedures
Adherence and Treatment Integrity
All participating therapists (N=total = 50, nST = 29, nCBT = 28, nIST = 31) were trained clinical psychologists or psychiatrists. To qualify as a study therapist, treatment-specific training workshops had to be completed. Therapists delivered IST either in addition to CBT or in addition to ST but were not allowed to deliver both active treatment conditions (ST and CBT) at the same time. They were allowed to switch between ST and CBT after 1 year.

To ensure treatment adherence, therapists received expert supervision (monthly) in groups of four therapists, separated by the treatment condition. The trainers and supervisors were all certified therapists in either CBT and ST or CBT alone, with several years of clinical experience. Supervision and intervision sessions were based on psychotherapy sessions presented by the psychotherapist (videos of individual and/or group sessions). In addition, all sessions (individual and group) were videotaped. The adherence ratings of a random subsample of videos showed high to very high treatment adherence and highly significant differences between treatment arms regarding the execution of specific interventions. For details, see online supplementary 1A.

Outcomes
The primary outcome was defined as depression severity over the course of 7 weeks of treatment, assessed weekly by BDI-II [33]. The secondary outcomes were change over treatment in depression severity (change over baseline), treatment response rate (defined as ≥50% decrease in depressive severity), and relapse prevention (time to relapse [33, 34]). Secondary outcomes included cognitive and behavioral functioning, social functioning, self-awareness and self-confidence, and therapist alliance (rated by the patients at the end of treatment) [38]. All outcomes were measured in the complete study center over the complete course of recruitment and therapy delivery (7 weeks).

Effectiveness of Schema Therapy for Depression

Psychoter Psychosom
DOI: 10.1159/000535492
severity assessed as clinical rating (MADRS) [34], general psychopathology assessed by the Brief Symptom Inventory (BSI) [42], global functioning assessed as clinical rating by the World Health Organization Disability Assessment Schedule (WHODAS) [43], and quality of life assessed by the World Health Organization Quality Of Life (WHOQOL) [44]. Additionally, response to treatment, defined as 50% decrease in the BDI-II sum score from baseline, remission, defined as the BDI-II sum score below ≤12, and change in M-CIDI diagnosis [35] were investigated. For further exploratory analysis, all primary and secondary outcomes were assessed at 6-month follow-up. Clinical ratings (MADRS and WHODAS) were conducted by trained and blinded raters. Inter-rater reliability (intraclass coefficient) was calculated on a regular basis and showed excellent correspondence (MADRS: M = 0.930, SD = 0.040; WHODAS: M = 0.998, SD = 0.004).

**Sample Size Calculation**

Sample size calculation was based on the estimates by Button and colleagues [27], who defined a 17.5% reduction in BDI-II scores from baseline as the minimal clinically important difference. Based on pilot data, defining the target-sample as moderately to severely depressed, we calculated a required sample size of N = 300, when setting power at 0.80 and α at 0.05 in post-treatment differences in BDI-II [30], which is also sufficient for linear mixed models.

**Statistical Analysis**

Data were analyzed between August 2021 and April 2022 based on the intention-to-treat sample. In order to test the first hypothesis (superiority of ST over IST), two-level multilevel modeling was conducted for the primary outcome (BDI-II sum scores) and the continuous secondary outcomes (i.e., MADRS, BSI, WHODAS, WHOQOL). Treatment effects over time were examined using interaction terms of contrast-coded treatment variables (DifferenceIST to ST, DifferenceST to CBT) by week (time). Time was added as an interaction term in the multilevel model were provided for post-treatment and follow-up. Logistic regressions were conducted for response and remission. For sensitivity analyses, all results were conducted based on the completer sample (n = 207). Analyses were performed in R Statistics with a significance level of 0.05/7.

In order to test the second hypothesis (noninferiority of ST compared to CBT), the noninferiority margin of a minimal clinically important difference was defined as 17.5% change from baseline severity (reported BDI-II sum scores) in accordance with empirical estimations [27] and the study protocol [30]. For claiming noninferiority, the pre-post-percentage change score of ST was not allowed to be lower than 17.5 percentage points than the CBT equivalent.

**Results**

**Sample**

We assessed 1,370 patients for eligibility, of which 292 met the inclusion criteria and were randomized to ST (95), CBT (101), or IST (96). Treatment was completed by 207 (nST = 62, nCBT = 80, and nIST = 65) and follow-up assessment by 159 patients (nST = 46, nCBT, = 60, and nIST = 53). For details, see CONSORT flowchart in Figure 1.

The mean age was M = 43.53 (SD = 13.60), 162 patients (55.48%) were female, 145 (49.66%) were married or cohabitating, 177 (40.07%) reported high education level, 150 (51.35%) were employed, and 241 (82.51%) were holding German citizenship. Most patients (n = 238; 82.52%) were treated in an inpatient setting and 72 (24.66%) underwent treatment under pandemic-related hygienic conditions, such as wearing masks. Patients reported an average amount of M = 2.36 (SD = 2.08) previous treatments (psychotherapy or pharmacotherapy). Mean level of depression at baseline was severe (BDI-II: M = 32.02, SD = 8.61) and many patients suffered from comorbidities such as anxiety disorders (n = 171, 60.00%), substance disorders (n = 76, 26.67%), somatoform disorders (n = 65, 22.81%), obsessive compulsive disorder (n = 39, 22.80%), and post-traumatic stress disorder (n = 37, 12.98%) assessed by M-CIDI [35]. None of these numbers differed significantly between treatment conditions. For details, see Table 1.

A tendency of more patients dropping out from the ST (n = 33, 34.74%) and IST condition (n = 31, 32.29%) than CBT condition (n = 21, 20.79%) (p = 0.06) was found. Reasons for dropout were discharged from hospital (n = 45), being overloaded with study procedures/missing motivation (n = 13), missing more than 6 of the offered treatment sessions (corresponds to 22%, n = 8), and receiving treatment not in accordance with study regulations (n = 7). Participants who dropped out of the study did not differ significantly from those who continued in the study in terms of BDI-II sum scores at the time of dropout (Mnondropout = 23.60, Mdropout = 24.43, p = 0.50).

**Interventions and Concomitant Care**

The number of sessions received by patients was M = 20.62 (SD = 8.57) in ST, M = 22.76 (SD = 6.14) in CBT, and M = 17.10 (SD = 8.57) in IST (F(2, 289) = 13.47, p ≤ 0.001). To investigate the dose effect on the effect of treatment arms, an interaction term (dose by treatment difference score) was added to the multilevel model. The analysis did not show an interaction effect between psychotherapy dose and treatment arms on BDI-II (IST-ST: B = 1.10, 95%
Fig. 1. CONSORT flowchart. Trial design and flow of patients throughout the measurement points at baseline, after 7 weeks of treatment, and 6-month follow-up.
Table 1. Baseline characteristics of the sample

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
<th>ST (n= 95)</th>
<th>CBT (n=101)</th>
<th>IST (n= 96)</th>
<th>total (N=292)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), years</td>
<td>M= 43.58 (12.65)</td>
<td>M= 42.22 (14.04)</td>
<td>M= 44.85 (14.03)</td>
<td>M= 43.53 (13.60)</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Gender (female)a, n (%)</td>
<td>54 (56.84)</td>
<td>50 (49.50)</td>
<td>58 (60.41)</td>
<td>162 (55.48)</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>Family status, n (%)</td>
<td>Married or cohabitating: 53 (55.79)</td>
<td>52 (51.49)</td>
<td>40 (41.67)</td>
<td>145 (49.66)</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single: 26 (27.39)</td>
<td>26 (25.74)</td>
<td>34 (35.42)</td>
<td>86 (29.45)</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Separated, divorced, or widowed: 6 (6.32)</td>
<td>15 (14.85)</td>
<td>17 (17.71)</td>
<td>38 (13.01)</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Educational levelc, n (%)</td>
<td>Low: 18 (18.95)</td>
<td>15 (14.85)</td>
<td>15 (15.63)</td>
<td>48 (16.44)</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium: 34 (35.79)</td>
<td>31 (30.69)</td>
<td>32 (33.33)</td>
<td>97 (33.22)</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High: 33 (34.74)</td>
<td>44 (43.56)</td>
<td>40 (41.67)</td>
<td>117 (40.07)</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Employment (yes)</td>
<td>47 (49.47)</td>
<td>57 (56.44)</td>
<td>46 (47.92)</td>
<td>150 (49.66)</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>German nationality</td>
<td>77 (81.05)</td>
<td>87 (86.14)</td>
<td>77 (80.21)</td>
<td>241 (82.52)</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Inpatient setting</td>
<td>76 (80.00)</td>
<td>80 (79.21)</td>
<td>82 (85.42)</td>
<td>238 (81.51)</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>Treatment during COVID pandemicd</td>
<td>27 (28.42)</td>
<td>26 (25.74)</td>
<td>19 (19.79)</td>
<td>72 (24.66)</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>No. of previous treatmentsg</td>
<td>M= 2.20</td>
<td>M= 2.31</td>
<td>M= 2.51</td>
<td>M= 2.36</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Depression severity at baselineh</td>
<td>(SD = 2.18)</td>
<td>(SD = 2.04)</td>
<td>(SD = 2.01)</td>
<td>(SD = 2.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbiditiesi, n (%)</td>
<td>Anxiety disorders: 57 (62.64)</td>
<td>60 (60.00)</td>
<td>54 (57.44)</td>
<td>171 (60.00)</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Substance disorder: 25 (27.39)</td>
<td>22 (22.00)</td>
<td>29 (30.85)</td>
<td>76 (26.67)</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Somatoform disorder: 25 (27.36)</td>
<td>17 (17.00)</td>
<td>22 (23.40)</td>
<td>65 (22.81)</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OCD: 12 (13.19)</td>
<td>12 (12.00)</td>
<td>15 (15.96)</td>
<td>39 (22.80)</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PTSD: 14 (15.38)</td>
<td>14 (14.00)</td>
<td>9 (9.57)</td>
<td>38 (13.01)</td>
<td>0.47</td>
<td></td>
</tr>
</tbody>
</table>

ST, schema therapy; CBT, cognitive behavioral therapy; IST, individual supportive therapy; OCD, obsessive compulsory disorder; PTSD, post-traumatic stress disorder. aData are shown as numbers (percentage) unless specified differently. bParticipants were given the opportunity to report their gender. Race or ethnicity was not ascertained. cLow corresponds to 9 years of education or less; Medium corresponds to 10–12 years of education; High corresponds to 13 years of education. dTreatment was conducted during the corona pandemic including hygiene rules like wearing a mask and keeping distance. eTreatment was defined as a period when perceiving medication or psychotherapy or an inpatient or day clinic stay. fAccording to reported BDI-II sum scores. gAccording to Munich-Composite International Diagnostic Interview (M-CIDI); N=285; nST = 91, nCBT = 100, nIST = 94; only comorbid disorders are shown that were reported to be more than 5% of the ITT sample. hIncludes generalized anxiety disorder, specific phobia, social phobia, agoraphobia, and panic disorder; OCD; PTSD.

CI: =−4.62 to 6.82, p = 0.71; IST-CBT: B = 6.49, 95% CI: =−1.76 to 14.74, p = 0.12; and CBT-ST: B = −5.39, 95% CI: =−14.18 to 3.39, p = 0.23). Additionally, patients received a comprehensive psychiatric care program, including pharmacotherapy and concomitant care. Mostly used medication at baseline were antidepressants, taken by 246 patients (84.42%), neuroleptics (n = 85, 29.11%), and mood stabilizers (n = 41, 14.04%). Over the course of 7 weeks of treatment, 276 patients (94.52%) took antidepressants, 115 (39.38%) neuroleptics, and 56 (19.19%) mood stabilizers. Most of the patients participated at least once in offered ergotherapy treatment (n = 223, 76.37%), sports (n = 217, 74.32%), case management (n = 202, 69.18%), relaxation training (n = 153, 52.40%), or cognitive training (n = 99, 33.91%). Neither medication (baseline or during trial) nor concomitant care (single interventions or overall dose) differed significantly between treatment arms (online suppl. 1B/1C).

Primary Outcome: Severity of Depression
Depressive symptoms decreased significantly throughout therapy (B = -1.87; 95% CI: = −2.24 to −1.50, p < 0.001). Depression severity (reported BDI-II sum scores) over time can be taken from Figure 2 and online supplementary 1D.

Figure 2 shows depression severity over time. Multilevel modeling indicated no significant difference scores.
main effects between IST and ST (B = 0.35, 95% CI: = −0.85 to 1.55, p = 0.57) nor interaction between IST and ST difference scores and time in BDI-II sum scores (B = −0.05, 95% CI, −0.58 to 0.47, p = 0.84; see Table 2). So, hypothesis H1 was not confirmed. The exploratory analysis on change in BDI-II scores including the follow-up timepoint (t8; online suppl. 2A) and the sensitivity analysis in the completer sample (n = 207) demonstrated the same effects (online suppl. 3A). The estimated marginal means for the week by treatment interaction in the multilevel model showed no significant differences in BDI-II scores between ST and IST at the end of treatment (ST: M = 17.63, 95% CI: = 15.02–20.23; IST: M = 17.65, 95% CI: = 15.05–20.25) and at follow-up (ST: M = 15.88, 95% CI: = 13.09–18.64; IST: M = 16.51, 95% CI: = 13.74–19.28; for details see online suppl. 1E).

Exploratory sensitivity analysis [37] revealed no difference scores main effects between IST and CBT (B = 0.15, 95% CI: = −1.03 to 1.32, p = 0.81), but interaction effects between IST and CBT difference scores and time (B = −0.60, 95% CI: = −1.10 to −0.09, p = 0.02). These effects were also shown at the follow-up (t8; online suppl. 2A) and on a marginal significant level in the completer sample (online suppl. 3A), and demonstrate the adequacy of the trial design and the robustness of the findings from the primary comparison (IST-ST).

The exploratory analysis on the differences between CBT and ST showed the same pattern: no main effect investigating the difference between CBT and ST (B = 0.21, 95% CI: = −0.98 to 1.39, p = 0.73), but an interaction effect of CBT-ST difference scores with time (B = 0.54, 95% CI: = 0.04–1.05, p = 0.04; see online suppl. 1F). The estimated marginal means for the week by treatment interaction in the multilevel model also indicated these differences in BDI-II scores between conditions at the end of treatment (ST: M = 17.63, 95% CI: = 15.02–20.23; CBT: M = 13.68, 95% CI: = 11.21–16.04) but not at follow-up (ST: M = 15.88, 95% CI: = 13.09–18.64; CBT: M = 11.92, 95% CI: = 9.35–14.48; for details see online suppl. 1E). The exploratory analysis on change in BDI-II scores including the follow-up timepoint (B = 0.45, 95% CI: = −0.04 to 0.94, p = 0.07) and the sensitivity analysis in the completer sample (n = 207) replicated the effects (B = 0.30, 95% CI: = −0.20 to 0.80, p = 0.24) without suggesting superiority of CBT over ST in interaction with time.

Secondary Outcomes
A significant decrease of MADRS, BSI, and WHODAS scores and a significant increase of WHOQOL scores were found throughout the treatment (online suppl. 1G). Neither of these effects was moderated by the difference between IST and ST (online suppl. 1H). These effects were also shown in the exploratory analysis on secondary outcomes including the follow-up timepoint (online suppl. 2B/2C). Likewise, the results of the main predictors were also demonstrated in the sensitivity analysis in the completer sample (online suppl. 3B).

Response and Remission
53% of patients (n = 110) responded to treatment at t7 defined as a BDI-II reduction above 50% from baseline (t0). Response to treatment was neither significantly predicted by the IST-ST differences (H1) nor by the IST-CBT differences (replicated in the exploratory follow-up analysis; see online suppl. 11 and 2C).

38% of patients (n = 79) remitted from depression at the end of treatment (t7) as defined by a BDI-II score below or equal to 12. Remission of depression was not

---

**Table 2.** Results of multilevel model investigating the primary outcome (BDI-II sum scores) from pre- to posttreatment

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>B (95% CI)</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p value</th>
<th>d (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>30.85 (30.01–31.70)</td>
<td>0.43</td>
<td>1536</td>
<td>71.58</td>
<td>&lt;0.001*</td>
<td>–</td>
</tr>
<tr>
<td>BDI-II baseline</td>
<td>−0.86 (−0.90–0.81)</td>
<td>0.03</td>
<td>286</td>
<td>30.01</td>
<td>&lt;0.001*</td>
<td>3.55 (3.08–4.02)</td>
</tr>
<tr>
<td>Week</td>
<td>−1.87 (−2.24 to −1.50)</td>
<td>0.19</td>
<td>1536</td>
<td>−9.83</td>
<td>&lt;0.001*</td>
<td>−0.50 (−0.74 to −0.26)</td>
</tr>
<tr>
<td>DifferenceIST to CBT</td>
<td>0.15 (−0.13–1.32)</td>
<td>0.60</td>
<td>286</td>
<td>0.24</td>
<td>0.81</td>
<td>0.03 (−0.20–0.26)</td>
</tr>
<tr>
<td>DifferenceIST to ST</td>
<td>0.35 (−0.85–1.55)</td>
<td>0.61</td>
<td>286</td>
<td>0.57</td>
<td>0.57</td>
<td>0.07 (−0.16–0.30)</td>
</tr>
<tr>
<td>Week × DifferenceIST to CBT</td>
<td>−0.60 (−1.10 to −0.09)</td>
<td>0.26</td>
<td>1536</td>
<td>−2.30</td>
<td>0.022*</td>
<td>−0.12 (−0.35–0.11)</td>
</tr>
<tr>
<td>Week × DifferenceIST to ST</td>
<td>−0.05 (−0.58–0.47)</td>
<td>0.27</td>
<td>1536</td>
<td>−0.20</td>
<td>0.84</td>
<td>−0.01 (−0.24–0.22)</td>
</tr>
</tbody>
</table>

Variables DifferenceIST to CBT and DifferenceIST to ST are the contrast-coded treatment variables with IST as the reference category (coded 0) as compared to CBT or ST (coded as 1). Model characteristics: random slope: week; covariance structure: autoregressive (1). *Indicates significance at α = 0.05.
significantly predicted by the IST-ST or CBT-IST differences (online suppl. 1I). Both the effects were replicated in the exploratory follow-up analysis (online suppl. 2C). Change in depression diagnosis according to M-CIDI was not predicted by differences between any treatment condition (online suppl. 1I). Ten patients (4.8%) reported a deterioration defined as an increase in BDI-II scores from baseline to post-treatment ($n_{ST} = 4$, percentage change: $M_{ST} = 6.08\%$, $SD_{ST} = 3.36$; $n_{CBT} = 2$, $M_{CBT} = 12.8\%$, $SD_{CBT} = 1.22$; and $n_{IST} = 4$, $M_{IST} = 16.0\%$, $SD_{IST} = 9.65$).

**Noninferiority**

In order to further investigate the clinical relevance of ST, it was tested for noninferiority against the gold-standard for the treatment of depression, CBT. The used noninferiority margin of what can be considered the minimal clinically important difference in BDI-II was defined as 17.5% change from baseline severity [27, 30]. Thus, the pre-post-percentage change score of ST was not allowed to be lower than 17.5 percentage points than the CBT equivalent to claim noninferiority. The analysis revealed a pre-post-percentage change score of $M = -44.37$ (95% CI: $-51.67$ to $-37.08$) corresponding to $-14.13$ in absolute BDI-II change scores for ST and $M = -52.01$ (95% CI: $-57.72$ to $-46.30$) corresponding to $-16.49$ BDI-II change scores for CBT. Thus, the ST change score lay within the tolerable margin of 17.5% for noninferiority (cut off: $M_2 = 34.51$) and hypothesis $H_2$ was confirmed. Percentage changes in manifesting BDI-II sum scores and noninferiority margins are reported in Figure 3.

As a further exploratory sensitivity analysis, we examined the effect of IST ($M = -39.58$, 95% CI: $-46.57$ to $-32.59$) corresponding to $-13.53$ BDI-II change scores, which was not within the margin and thus did not show noninferiority compared to CBT. This finding, again, stresses the assay sensitivity of the trial [37] and the robustness of the results of the primary comparison (ST-CBT).

**Safety**

Three adverse events, defined as “acute suicidality,” occurred during the study ($n = 2$ in ST and $n = 1$ in CBT). Indicated psychiatric care was provided immediately and participants were excluded from the study. There was no identifiable relation to the study procedures. Serious adverse events (suicide attempts) did not occur during the study.

**Discussion**

In this RCT, a specific, semi-structured ST was investigated in psychiatric inpatient and day clinic settings in a sample of severely depressed highly

---

**Fig. 2.** Reported depression severity over time. Development of BDI-II scores over time. ST, schema therapy; CBT, cognitive behavioral therapy; IST, individual supportive therapy; x-axis represents study weeks: 0, baseline; 1, study week 1/start of treatment; 7, study week 7/end of treatment; 8, 6-month follow-up.
comorbid patients. This was the first trial examining the use of ST for depression in inpatient and day clinic settings as a high-intensity short-term treatment.

ST versus IST. The first aim of the current trial was to test the superiority of ST over IST. The applied multilevel modeling revealed no significant differences between ST compared to the nonspecific, nonstructured supportive therapy (IST) in symptom severity over the course of treatment. These findings were consistent throughout sensitivity analyses and were supported by the estimated marginal means after treatment and at follow-up. They also applied to the level of functioning, general psychopathology, quality of life, response rate, or remission.

ST versus CBT. At the same time, ST showed non-inferiority in a clinically relevant sense compared to a specific and highly structured first-line CBT program. These results are in line with the previous findings for long-term ST for depression in outpatient settings [17], suggesting similar effectiveness of ST and CBT. Nevertheless, the subsequent exploratory analysis revealed less consistent results. Multilevel modeling showed a small but significant interaction effect of ST-CBT difference scores over the course of treatment. This finding was supported by EMMs differences at end of treatment but disappeared at follow-up. The secondary outcomes did not show any significant difference between ST and CBT throughout treatment. Thus, the results of the exploratory multilevel model indicate a potential short-term advantage for CBT in the given setting.

Overall, the treatment conditions in this trial performed similarly. Therefore, ST may be a useful addition to the therapeutic repertoire for the treatment of depression in a psychiatric setting. Restraints regarding the effectiveness of a high-intensity and short-term treatment might be related to the missing opportunity to play-off ST-strengths targeting underlying mechanisms in the short run. Therefore, it remains to be investigated how ST unfolds in terms of long-term effects or whether some patients (e.g., patients with chronic depression or a comorbid personality disorder) particularly benefit from ST compared to others.

Moreover, these results contribute to the debate on the role of common factors of psychotherapy versus specific interventions for depression. On the one hand, missing superiority of ST toward IST underlines previous findings about the therapeutic power of such supportive treatments as active control condition in clinical trials [41, 48]. Forms of supportive therapy may be a less expensive and easier-to-implement option in psychiatric care that should be explored more in detail in the future. On the other hand, the differences between CBT (highly structured, depression-specific) and IST (nonstructured, common factor based) and the lack of noninferiority of IST compared to CBT underline that not all therapies are equally effective. The results suggest that common factors play a crucial role in the treatment of highly comorbid and severely depressed patients, but that disorder-specific ingredients and a structured delivery in a short-term treatment have an additional effect not to be neglected. Although the active control condition IST is a manualized treatment based on common factors of psychotherapy, a
certain overlap with some forms of supportive counseling cannot be ruled out. The latter has been shown to be inferior compared to other psychotherapy approaches of depression [5, 49]. Therefore, the inferiority of IST compared to CBT may partly be attributed to the shared concept of supportive counseling, which would be in line with previous findings.

The overall dropout rate of about 29% is above average for psychotherapy trials for depression [50], but within the normal range, given the higher dropout rates in trials using medication and/or naturalistic settings [51]. Our results show a trend toward lower dropout in the CBT condition. Since dropout in psychotherapy trials has often been associated with patient dissatisfaction with treatment, lack of motivation, problems in therapeutic alliance, and poorer treatment outcomes [52], this could be interpreted as a slight indication of superiority of the highly structured CBT program compared to the other treatment conditions. However, this might not be easily transferred to inpatient trials. Individual costs of hospitalization, such as being away from home or family, or sharing a room with strangers may lead to increased motivation to leave hospital once a certain level of improvement has been achieved, as “got-what-they-needed” dropouts [53]. The lack of differences in symptom severity between dropouts and nondropouts in the current study underlines the difficulty of disentangling reasons for dropout in inpatient trials and comparing them with those in outpatient psychotherapy trials. Therefore, the (only marginally significant) results should be interpreted with caution.

The OPTIMA trial has several strengths by fulfilling important requirements for high-quality psychotherapy trials [29] such as active and effective control groups, a broad range of outcome measures, the combination of self-reports and clinical ratings by blinded raters, standardized and manualized treatments, a highly co-morbid, medicated, heterogeneous, and thus externally valid sample, and the implementation in a comprehensive psychiatric care program which represents clinical reality. Apart from that, the study is not without limitations. Naturally, participants and therapists in psychotherapy trials are not blinded, so expectancy effects cannot be ruled out. Regarding diagnosis, the longitudinal course of the disorders was considered by assessing the number of previous treatments (Table 1), but more systematic information according to staging models [54] would have given more insight. Another limitation regarding diagnostics is the use of M-CIDI based on the DSM-IV criteria. This instrument was chosen since at the timepoint of recruitment start, no German version of DSM-V-based clinical interviews was available. In addition, the noninferiority margin was informed by previous empirical research [27] and takes into account patients’ subjective perspectives as well as individual differences in symptom severity. Yet, we acknowledge that there are controversies in the definition of noninferiority margins [55], and recognize the resulting pitfall for interpretations of noninferiority analyses.

Furthermore, some setting-related limitations must be taken into account. For different reasons, we were not able to deliver the same treatment dose consistently, yet to control for dose effects (same applies for potential COVID-related effects). For ethical reasons, we could not add a medication-free treatment condition to the study design. Therefore, our conclusions are limited to the treatment differences in a combined pharmaco-psychotherapy treatment embedded in a psychiatric care program. Possible confounding effects of medication cannot be ruled out but can only be statistically controlled for. In addition to medication, patients received a comprehensive psychiatric care program including various interventions such as occupational therapy, case management, etc., which may have led to a ceiling effect and treatment oversaturation. Additionally, the demanding study procedures might have resulted in a selection bias. The combination of self-reports and clinical ratings is a methodological strength of the trial [56], and BDI-II is one of the most widely used assessment tools in depression research. However, the use of a self-report measure as the primary outcome is a limitation due to its limitations to assessing clinical change [57]. We used a slow-open design to our group format (clinical reality in many settings), resulting in varying patient constellations, which might have limited group cohesion. Finally, due to COVID-related limitations, we were only able to conduct follow-up assessments using online questionnaires. This resulted in a relatively low assessment rate and must be taken into account when interpreting the effects at follow-up.

**Conclusion**

In conclusion, the OPTIMA trial found that high-intensity ST for depression as a part of a psychiatric care program showed clinical noninferiority compared to CBT, without being superior to IST. Regarding short-term treatment effects, the trial found that CBT displayed some advantages over ST, possibly because the cognitive behavior program was more structured. Overall, ST
represents a potentially useful addition to the therapeutic repertoire for the treatment of depression, but its efficacy, including long-term efficacy, should be evaluated further.

Acknowledgments

We would like to thank Seth Levine for revising the manuscript and Tanja Brückl for advice; Gabi Kohl, Anna Hartl, and Melanie Huber for the administration of the day-to-day business and conduct of measures, all contributors to this research project including physicians, psychologists, study nurses, research assistants, and of course patients of the hospital of the Max Planck Institute of Psychiatry in Munich.

Statement of Ethics

The trial is registered at clinicaltrials.gov (NCT03287362). It was approved by the Institutional Ethics Committee of the Faculty of Medicine of Ludwig Maximilian University, Munich (project number 17-395). Modifications of the protocol were reported to the committee. All participants were informed in oral and written form about the study, treatments, assessments, and duration of the study. Written informed consent was obligatory prior to enrollment.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

References
