Clinical Psychology: Science and Practice

Antidepressant Medication or Short-Term Psychodynamic Psychotherapy for Depression? A Systematic Review and Meta-Analysis of Individual Participant Data

Frederik J. Wienicke, Jack J. M. Dekker, Jaap Peen, Henricus L. Van, Jacques P. Barber, Kevin S. McCarthy, Nili Solomonov, Hasse Karlsson, Jarmo Hietala, Jaime Lopez-Rodriguez, Valerio Villamil-Salcedo, William J. Burk, Jan Spijker, Jos W. R. Twisk, Zachary D. Cohen, Pim Cuijpers, Ellen Driessen

Funded by:

- American Psychoanalytic Association
- National Institute of Mental Health
- Netherlands Organisation for Scientific Research

Antidepressant medication or short-term psychodynamic psychotherapy for depression? A systematic review and meta-analysis of individual participant data.

Date submitted: September 5th, 2024

Abstract

Objective: Antidepressant medication and short-term psychodynamic psychotherapy (STPP) are frequently used to treat depression, but it is unclear which works best for whom. This preregistered individual participant data (IPD) meta-analysis (PROSPERO No.: ###############) examined efficacy and moderators of antidepressants versus STPP for adult depression. **Eligibility criteria**: Randomized comparisons of antidepressants and STPP on standardized measures for depressed adults were included. **Methods of synthesis**: A systematic literature search was performed May 1st, 2024. IPD were requested and analyzed using mixed-effects models. **Results**: Of the six trials identified (n = 472), IPD were obtained for four (n = 310; 65.7%). At post-treatment, antidepressants were slightly more efficacious than STPP regarding depressive symptom levels (d = 0.28, 95% CI [0.03, 0.53], p = .031), but no significant differences were observed at follow-up, nor for self-reported depression or any of the secondary outcomes at post-treatment. Baseline depression severity was found to moderate post-treatment outcomes (d = 0.24, 95% CI [0.08, 0.40], p = .004), such that antidepressants were more efficacious than STPP for participants with higher severity levels rather than those with lower baseline severity. **Conclusions**: This is the first study examining moderators across trials comparing antidepressants and STPP. The findings suggest that baseline severity might be a factor to consider when choosing between these treatments. However, this finding is observational and requires validation in future studies before it can be used to inform treatment selection for depression.

Keywords: depression, antidepressant medication, short-term psychodynamic psychotherapy, individual participant data meta-analysis, moderator

What is the public health significance of this article?

The findings of this study suggest that individuals with depression and their clinicians might expect similar improvements in self-reported depression, anxiety, general psychopathology, and general health after treatment with antidepressants and STPP, and slightly greater improvements in observer-rated depression at the end of treatment with antidepressants rather than STPP. For those showing relatively high baseline depression levels, antidepressants might lead to greater improvements in terms of post-treatment depressive symptom levels compared to STPP. However, these findings are based on a modestly sized sample of predominantly middle-aged women from moderate- to high-income countries.

Antidepressant medication or short-term psychodynamic psychotherapy for depression? A systematic review and meta-analysis of individual participant data.

Since depression is ranked as the single largest contributor to global disability (World Health Organization, 2022), efficacious and efficient treatments are needed. Antidepressant medications and different psychotherapies comprise the major treatment options for depressive disorders (Marcus & Olfson, 2010) and treatment guidelines recommend combining the two for individuals with moderate to severe depression (American Psychological Association [APA], 2019; National Institute for Health and Care Excellence [NICE], 2022). However, a substantial number of patients prefer monotreatment because of personal preferences (McHugh et al., 2013; van Schaik et al., 2004) and concerns about antidepressant side effects (Ferguson, 2001) and long-term use (Bet et al., 2013). Moreover, combined treatment of antidepressants and psychotherapy requires additional investments of time and resources, which is not always feasible for patients and society given the scarcity of available treatment options for depression. For these reasons, it is important to robustly establish the comparative efficacy of antidepressant medication and various psychological treatments for depression, to help patients and clinicians to make well-informed treatment decisions.

Previous research has examined the comparative efficacy of antidepressants and various types of psychotherapy, including cognitive behavioral analysis system of psychotherapy (CBASP; Furukawa et al., 2018), interpersonal therapy (IPT; Cohen et al., 2024), and cognitive behavioral therapy (CBT; Weitz et al., 2015). No significant differences between antidepressants and CBASP or antidepressants and IPT were found regarding post-treatment depressive symptoms (Cohen et al., 2024; Furukawa et al., 2018). However, antidepressants showed a slight advantage over IPT in reducing general psychopathology and

dysfunctional attitudes. In comparison to CBT, antidepressants showed a slight advantage in terms of post-treatment observer-rated depressive symptoms (Weitz et al., 2015).

Short-term psychodynamic psychotherapy (STPP) is another empirically supported treatment for depression (Driessen et al., 2015; Cuijpers et al., 2021) that is frequently applied in clinical practice (Norcross & Rogan, 2013). STPP emerged in the 1970s as a briefer alternative to long-term psychodynamic psychotherapy and typically comprises 12 to 24 weekly sessions (Abbass et al., 2014), focusing on decreasing maladaptive defenses and enhancing insight into unconscious conflicts or avoided emotions stemming from past negative experiences (Charman, 2004; Eppel, 2018). An international survey of over 10,000 psychotherapists revealed that 62% incorporated elements of psychodynamic therapy into their practice (Orlinsky et al., 2020), and about half of 90 globally surveyed therapist training programs have a psychodynamic orientation (Orlinsky et al., 2024). Nevertheless, the comparative efficacy of antidepressants and STPP remains largely understudied.

A conventional meta-analysis comparing the efficacy of antidepressants and STPP reported no significant differences in terms of post-treatment depressive symptoms and interpersonal functioning (Driessen et al., 2015). However, conventional meta-analyses, which are based on study-level information extracted from published trial reports, are limited by their dependence on the quality of the original publications in which treatment effects can be overestimated (Turner et al., 2008). Thus, they can produce biased results (Stewart & Parmar, 1993). Furthermore, there were not enough studies reporting follow-up outcomes to examine the long-term comparative efficacy (Driessen et al., 2015).

Moreover, it remains largely unclear which treatment works best for whom. While STPP has been found more efficacious than antidepressants for males from self-designated minority groups in one study (Barber et al., 2012), these results require replication. An important reason why it is unclear which pre-treatment participant characteristics (so-called

moderators) are associated with differential responses to antidepressants and STPP is the lack of statistical power in individual clinical trials. These typically have sample sizes aimed at identifying intervention effects, while identifying moderator effects requires substantially larger numbers of participants (Brookes et al., 2004).

Conventional meta-analyses have also not been able to answer the question of which depression treatment works best for whom since moderator analyses as part of conventional meta-analysis rely on associations between study-level characteristics and therefore may introduce ecological bias. Ecological bias occurs when relationships found at the study level (e.g., the association between the mean age of the study sample and the average treatment effect) do not necessarily translate to relationships at the participant level (Tudur Smith et al., 2016).

Individual participant data (IPD) meta-analysis is an alternative to conventional meta-analysis that examines treatment effects by combining participant-level data from multiple clinical trials, which increases statistical power (Lambert et al., 2002). IPD meta-analysis also has the advantage over conventional meta-analysis that the same analytic approaches can be used for handling missing data and statistical modeling (Riley et al., 2010). Moreover, results presented in the original studies can be verified and there is the possibility to implement novel, more powerful statistical techniques that were not available at the time of publication in the case of older studies (Riley et al., 2010). For these reasons, IPD meta-analysis can provide more reliable effect estimates (Stewart & Parmar, 1993). Furthermore, IPD meta-analysis allows moderators to be examined at the participant level and thereby can circumvent the risk for ecological bias (Tudur Smith et al., 2016).

We, therefore, conducted a systematic review and IPD meta-analysis to examine the efficacy of antidepressants versus STPP as compared on various outcome measures at post-treatment and follow-up in randomized clinical trials for adults with depression. Furthermore,

given the lack of previous research findings, we explored several participant characteristics as potential moderators of depression outcomes at both time points.

Methods

Study Selection

We searched five bibliographic databases (PubMed, PsycINFO, Embase.com, Web of Science, and Cochrane's Central Register of Controlled Trials), two grey-literature databases (GLIN, UMI ProQuest), and a prospective trial register (http://www.controlled-trials.com) from inception to January 1st, 2024, without applying language or date restrictions (for the search string applied in PubMed see online Supplementary Table S.1). We also searched references of psychodynamic therapy meta-analyses and contacted a listserve of psychodynamic therapy researchers for studies that might have been missed. Finally, in order to identify recent studies, we searched a database of randomized depression psychotherapy trials (www.metapsy.org) from inception to May 1st, 2024. The METAPSY database is developed through comprehensive literature searches in PubMed, PsycINFO, Embase.com, and the Cochrane Library. The exact search terms are available from https://osf.io#######. The METAPSY database has been used in a series of meta-analyses (Cuijpers, Miguel, Harrer, Plessen, Ciharova, Papola, et al., 2023) and is updated every four months.

We included studies if they reported outcomes on standardized measures for depressed adult participants receiving STPP in a randomized comparison with antidepressants. Participants were considered depressed if they met specified diagnostic criteria for major depressive disorder or another unipolar mood disorder, or if they presented an elevated score above the 'no depression' cut-off on a standardized measure of depression. We included studies in which STPP (a) was based on psychoanalytic theories and practices, (b) was time-limited from the treatment onset, and (c) applied language-based techniques.

The screening process consisted of three phases. First, the selection criteria were applied to the citations generated from the searches independently by two raters with a bachelor's degree or higher in psychology. Titles identified as potentially relevant were requested in full text. During the second screening phase, two independent raters with a bachelor's degree or higher in psychology applied the selection criteria to the full-text papers. During the third phase, two expert STPP researcher-clinicians checked the included papers to confirm that the therapies described met the criteria for STPP. In all three phases, disagreements were discussed and resolved by consensus. If consensus could not be reached a third rater was consulted.

Data Collection and Integrity Checks

Authors of the identified studies were contacted using a multi-step contact protocol and invited to contribute their studies' IPD. We requested the anonymized participant-level dataset, including all potential moderators assessed before the start of treatment and all outcome variables assessed during and after treatment. We checked whether all outcome measures and moderators reported in the publication were included in the dataset, whether their values matched, and whether there were any invalid, out-of-range, or inconsistent items. Discrepancies were discussed with the authors.

For each study, we then listed all outcome variables and assessment moments. From the publications, we extracted multiple study sample characteristics (recruitment, target group, depression diagnosis, mean pre-treatment depression score, percentage of women, mean age) and STPP treatment characteristics (duration, format, manual use, therapist training, integrity check). Data extraction was independently conducted by a clinical psychology master's student and Ph.D. student. Disagreements were resolved by consensus. We also rated Cochrane risk of bias items for selection and detection bias based on information in the publications, and attrition bias based on the IPD (Higgins et al., 2011). If

information was not reported in the publications, we asked the authors. We did not rate performance bias as it is considered impossible to blind participants and therapists to treatment in psychotherapy research. Selective reporting bias was considered not applicable, as we requested all outcome measures assessed.

Measures

Depressive symptom levels at post-treatment constituted the pre-specified primary outcome. For each trial, we identified the primary continuous depression outcome as defined by the authors. Follow-up depressive symptoms and other post-treatment or follow-up outcome measures were examined if they were measured in at least two of the included studies. Since different instruments or instrument versions were used to assess these constructs (see online Supplementary Table S.2), we converted outcomes into z-scores within time point and within each study. However, recognizing that standardization of outcomes might obscure meaningful between-study variance, we conducted sensitivity analyses using unstandardized 17-item Hamilton Depression Rating Scale (HAMD) scores as outcome.

Potential moderators included all pre-treatment demographic, clinical, or psychological participant characteristics that were assessed in two or more studies. If these were assessed differently in individual studies, they were also standardized, by converting scores into z-scores (within study) for continuous variables or by recoding variables into similar categories for categorical variables (see online Supplementary Table S.3).

Data Analysis

We conducted one-stage IPD meta-analyses using mixed-effects models with a three-level structure (study, participant, repeated measures) and restricted maximum likelihood estimation. Analyses were based on intent-to-treat samples as far as available, and heterogeneity was assessed with the I^2 statistic. After each analysis, we visually inspected histograms of (standardized) residuals to examine whether the assumption of normality was

not severely violated. All mixed-effects model analyses were performed in R (version 4.4.1; R Core Team, 2020) using the lme4 package (version 1.1-35.5; Bates et al., 2015). Conducting mixed-effects models allowed for participants with only a baseline value but missing post-treatment and/or follow-up assessments to still be included in the analyses (Twisk et al., 2013).

We started with a basic model including a time main effect and a time-by-treatment interaction. This approach is recommended by Twisk and colleagues (2018, equation 2c) because it adequately accounts for differences in baseline values. Time was treated as a categorical variable to facilitate treatment comparison at the different time points. We estimated a random intercept with respect to study and participants (to account for clustering of participants in studies, and repeated measures in participants), and fixed slopes (for the time main effect and time-by-treatment interaction). Based on a -2 log-likelihood change evaluation, we decided whether or not to add a random slope in follow-up analyses for the time-by-treatment interaction on study level. Using this approach, treatment effects can be directly obtained from the regression coefficient of the time-by-treatment interaction (Twisk et al., 2018). For analyses with z-scores as outcome measure, these can be interpreted in the same way as Cohen's *d* effect sizes, where effect sizes ≤0.32 are considered as small, 0.33 to 0.55 moderate, and ≥0.56 large (Lipsey & Wilson, 1993). For analyses with unstandardized scores, the regression coefficients of the time-by-treatment interactions can be interpreted as mean differences. We calculated remission rates for the total sample and per treatment condition based on observed data according to the criteria by Riedel et al. (2010), defining remission as having a post-treatment HAMD-21 total score ≤ 7 or a post-treatment HAMD-17 total score ≤ 6.

We also conducted a two-stage IPD meta-analysis to include studies that provided IPD as well as those that did not (Riley et al., 2007). In the first stage, we calculated effect sizes

separately for both sets of studies. For studies for which IPD were available, effect sizes were calculated using the analytic approach previously described. For studies for which IPD were not available, we extracted relevant summary statistics from the publications and computed the corresponding effect sizes (for more information, see online Supplementary Text S.1). In the second stage, the effect sizes from both IPD-available and IPD-unavailable studies were pooled using a random effects meta-analysis, allowing the true treatment effect and thereby the estimates per study to vary (Riley et al., 2011). These analyses were conducted using the metafor R package (version 4.6-0; Viechtbauer, 2010).

For each potential moderator, we added the moderator's main effect, time-by-moderator interaction, and time-by-moderator-by-treatment 3-way interaction to the basic one-stage IPD model described above. We applied a Bonferonni correction for multiple comparisons and considered a *p* value of <.004 (13 tests) for the 3-way interaction's regression coefficient indicating a statistically significant moderator effect. To avoid the risk of Type II errors (Heo & Leon, 2010), we also reported moderators with an associated *p* value <.05 but interpreted these with caution.

All significant moderators were modeled simultaneously in a final model to determine whether their effects were independent of each other. In addition, we conducted several pre-specified sensitivity analyses to examine the robustness of our findings. We added the risk of bias items, STPP characteristics, and study design characteristics as covariates to the mixed-effects models. For analyses with the HAMD as outcome, we repeated the analyses in studies that administrated the 17-item version, using unstandardized scores as the outcome measure. Additionally, we conducted two post-hoc explorative sensitivity analyses in which we added the number of treatment sessions and follow-up length in months as covariates to the mixed-effects models to account for differences between studies in the number of treatment sessions and follow-up timing. Moderators that appeared robust in the sensitivity analyses were probed

using simple slope analyses (Aiken et al., 1991). To facilitate the interpretability of the simple slopes' graphical representation, we used z-scores standardized across time points.

Transparency and Openness

Institutional review board approval was not required for this project since it comprises secondary data analysis of anonymized data from previously completed treatment trials. The protocol for this study has been published (####### et al., 20##) and the study design and analysis plan has been preregistered in the PROSPERO International prospective register of systematic reviews (registration number: ###############). Deviations from the protocol comprising analyses precluded by unavailable data are listed in online Supplementary Table S.4. Furthermore, this study follows the meta-analysis reporting standard (MARS) guidelines (Appelbaum et al., 2018; for the MARS checklist see online Supplementary Table S.5). The syntax for the analyses and the collective anonymized individual participant database developed for this study, as well as a data dictionary and relevant related documents (e.g., study protocol), are available for use by other researchers. Requests can be made to the corresponding author (###############). Access (with limited investigator support) will be granted after approval of a study proposal by all authors and a signed data access agreement.

Results

Included Studies

The literature search results are described in detail in online Supplementary Figure S.1. Six trials met the inclusion criteria for this study, totaling 472 participants. IPD were obtained for four studies (66.7%; Barber et al., 2012; Dekker et al., 2008; López Rodríguez et al., 2004; Salminen et al., 2008) totaling 310 participants (65.7%). Two studies did not provide their IPD; one was conducted more than 40 years ago and the authors indicated that they no longer had access to the IPD (McLean & Hakstian, 1979). For the second study (Rahmani et al., 2011), the authors did not respond to multiple contact attempts.

The characteristics of the included studies are described in Table 1. In all studies that provided IPD, the target group was adults with depression. Depression inclusion criteria typically consisted of a *Diagnostic and Statistical Manual of Mental Disorders* or *International Classification of Diseases* depression diagnosis combined with an elevated HAMD score, though in one study (25.0%; López Rodríguez et al., 2004) the depression diagnosis constituted the only inclusion criterion for depression. The study samples ranged from 20 to 141 participants. All studies used the HAMD as primary outcome measure, with three studies (75.0%; Barber et al., 2012; Dekker et al., 2008; Salminen et al., 2008) assessing the 17-item version and one study (25.0%; López Rodríguez et al., 2004) assessing the 21-item version. In all studies, STPP was provided individually by trained psychotherapists or psychiatrists for 8 to 20 sessions. STPP was based on the principles described by Luborsky (1984; 1995), de Jonghe (2013), Bellak (1993; 1994), Man (1973), and Malan (1976). In two studies an STPP treatment manual was used (Barber et al., 2012; Dekker et al., 2008), and treatment integrity was assessed by means of supervision and audio- or video-taped treatment sessions in three studies (Barber et al., 2012; Dekker et al., 2008; López Rodríguez et al., 2004). Antidepressants included the selective serotonin reuptake inhibitors fluoxetine and sertraline, and the serotonin noradrenaline reuptake inhibitor venlafaxine. Post-treatment assessments ranged from 8 to 32 weeks. Follow-up assessments were conducted in three studies (75.0%; Barber et al., 2012; López Rodríguez et al., 2004; Salminen et al., 2008), with follow-up periods ranging from 6 to 28 months post-baseline.

The four studies for which IPD were obtained totaled 157 (50.6%) participants in the STPP conditions and 153 (49.4%) in the antidepressant conditions. The majority of the participants (65.5%) were female, with a mean age of 37.6 years (*SD* = 10.6). In terms of diversity, only one study (Barber et al., 2012) assessed participants' ethnicity, with 51.9% of participants reporting to be from a self-designated minority group. No information was

available from any study regarding participants' sexual or gender orientation, physical or cognitive disabilities, social class, or intersectionality (i.e., the interplay of participants' social and demographic factors). One study (López Rodríguez et al., 2004) reported that 92.5% of the participants identified as religious. Socioeconomic status, indicated by education, income, and occupation varied within and across studies (see online Supplementary Table S.6).

Bias Assessments

An overview of the risk of bias assessment is provided in Table 2. Outcome assessors were not blind to treatment allocation in two studies (50.0%; Dekker et al., 2008; Salminen et al., 2008). For two studies (50.0%; Dekker et al., 2008; López Rodríguez et al., 2004) the complete intent-to-treat data were not preserved by the original investigators, and two studies (50.0%; Dekker et al., 2008; Salminen et al., 2008) did not employ adequate allocation concealment procedures. One study (25.0%; Barber et al., 2012) had low risk of bias scores on all four criteria assessed and none of the studies had high risk of bias scores on all four criteria assessed. Descriptively, the two studies for which IPD were not obtained differed from the studies for which IPD were available in their application of other STPP models (Della Selva, 1996; Marmor, 1973; 1975; Wolberg, 1967) and the use of a tricyclic antidepressant in one study (McLean& Hakstian, 1979; Table 1; online Supplementary Table S.7 & S.8).

Treatment Outcomes

Depression treatment outcomes (expressed as HAMD z-scores) of the individual studies are shown in online Supplementary Table S.9. The pooled treatment outcomes of all studies that contributed IPD are presented in Table 3. Antidepressants were found significantly more efficacious than STPP regarding post-treatment depressive symptom levels with a small effect size (d = 0.28, 95% CI [0.03, 0.53], p = .031). Based on observed data (i.e., participants with both a baseline and post-treatment HAMD score; N = 207), remission rates

were 31.4% (65/207) for the total sample; 33.3% (32/96) in the ADM, and 29.7% (33/111) in the STPP conditions.

No significant differences in treatment effects between antidepressants and STPP were found on post-treatment measures of self-reported depression (d = 0.25, 95% CI [-0.02, 0.52], p = .075), anxiety (d = 0.08, 95% CI [-0.22, 0.38], p = .599), general psychopathology (d = 0.17, 95% CI [-0.15, 0.49], p = .303), physical health (d = -0.08, 95% CI [-0.36, 0.20], p = .565), and follow-up depressive symptom levels (d = 0.31, 95% CI [-0.08, 0.72], p = .130). Heterogeneity was low in all these analyses (I^2 = 0).

The two studies with no available IPD both utilized the self-report Beck Depression Inventory (BDI; Beck et al., 1961). Therefore, we combined their effect estimates with the three studies for which IPD were available that also employed a self-report measure of depression in a two-stage IPD meta-analysis. In line with the one-stage IPD meta-analysis based on the studies for which IPD were available, no significant differences were found between antidepressants and STPP regarding post-treatment measures of self-reported depression (k = 5, N = 406, d = 0.17, 95% CI [-0.07, 0.42], p = .164, I^2 = 0).

The patterns of findings were replicated consistently across the sensitivity analyses in which the risk of bias items, STPP characteristics, study design characteristics, number of treatment sessions, and follow-up length in months were added as covariates to the treatment outcome models (see online Supplementary Table S.10). Repeating the analyses in the three studies that administered the 17-item HAMD version using unstandardized scores as outcome also did not change the pattern of results (B = 1.88, 95% CI [0.42, 3.32], p = .013, I^2 = 15.1%).

Moderators

Results of the moderator models are presented in Table 4. Baseline depression severity was found to moderate post-treatment depressive symptom levels (d = 0.24, 95% CI [0.08,

0.40], p = .004), such that antidepressants were more efficacious than STPP for participants with high rather than low baseline depression severity. In addition, general assessment of functioning (GAF) scores moderated depressive symptoms at post-treatment (d = 0.04, 95% CI [0.001, 0.08], p = .045), such that antidepressants were more efficacious relative to STPP for participants reporting high rather than low levels of functioning, but the p value did not reach the Bonferroni corrected significance.

When modeled simultaneously (online Supplementary Table S.11), both baseline depression severity and GAF scores remained significant moderators of post-treatment depressive symptom levels (p = .010 and <.001, respectively). Baseline depression significantly moderated treatment outcomes across all sensitivity analyses (online Supplementary Table S.12), but GAF scores were no longer a significant moderator when adding risk of bias items, STPP characteristics, and study design characteristics to the model (all p's > .05). Probing baseline depression as a moderator (Figure 1) showed that for participants with low baseline depression severity, both treatments led to significant and similar decreases in depressive symptom levels (Panel A). For participants with high baseline depression severity, both treatments also resulted in significant decreases in depressive symptom levels, but decreases were larger in the antidepressants than in the STPP condition (Panel B).

Discussion

We conducted a systematic review and IPD meta-analysis to examine efficacy and moderators of antidepressants versus STPP in the treatment of adult depression. Across the four studies that were identified by a systematic literature search and for which IPD could be obtained, antidepressants were found significantly more efficacious than STPP in terms of post-treatment depressive symptom levels with a small effect size. No significant differences were found between antidepressants and STPP regarding post-treatment measures of self-

reported depression, anxiety, general psychopathology, physical health, and follow-up depressive symptom levels. Baseline depression severity moderated post-treatment depressive symptom levels, such that antidepressants were more efficacious than STPP for participants with more severe baseline depression.

These results are in line with a larger IPD meta-analysis of antidepressants versus CBT for depression including 16 randomized clinical trials totaling 1,700 participants (Weitz et al., 2015). This larger IPD meta-analysis also found antidepressants significantly more efficacious than CBT regarding post-treatment depressive symptom levels (measured by the HAMD) and a nonsignificant trend on the self-reported Beck Depression Inventory (Weitz et al., 2015), suggesting that antidepressants might result in larger decreases in depressive symptoms than psychotherapy at the end of treatment. The differences in findings between clinician-rated and self-reported post-treatment depression outcomes might be accounted for by different facets of depression being measured by the clinician-rated and the self-report depression instruments (Faravelli et al., 1986), with the HAMD focusing more on somatic symptoms and the BDI more on depressive cognitions (Ma et al., 2021). It has to be noted, however, that in the current study, the different efficacy between antidepressants and STPP was mostly driven by one trial, in which outcome assessors were not blinded to treatment allocation (Dekker et al., 2008; see online Supplementary Table S.9 and Table 2). While the current study did not find a significant difference in depressive symptoms at follow-up between patients receiving antidepressants or STPP, the largest trial (Dekker et al., 2008) included in this study did not provide follow-up data thereby limiting the statistical power to find significant differences at follow-up.

Moving beyond average treatment effects, the results of our study suggest that specific patients benefit differently from antidepressants or STPP for their depression. In contrast with what has been reported for antidepressants and CBT (Weitz et al., 2015), our analyses showed

that antidepressants were more efficacious than STPP at the end of treatment for participants with relatively high baseline depression severity. This moderator effect might relate to the ability of severely depressed patients to benefit from psychodynamic interventions. Typical interventions used in STPP such as the exploration of unconscious emotions, defense mechanisms, or interpersonal experiences (Luyten & Blatt, 2012), demand a certain capacity for self-reflection and tolerance to experience discomfort (Rueve & Correl, 2006; Valbak, 2004). This capacity, however, might be initially obstructed in more severely depressed patients.

Overall level of functioning (i.e., GAF scores) also moderated post-treatment depressive symptoms such that antidepressants were more efficacious relative to STPP for participants reporting high rather than low levels of functioning. However, we interpret this finding with caution as the associated *p* value exceeded the Bonferroni correction and was not significant in the sensitivity analyses. We did not find any of the sociodemographic or other clinical variables moderating treatment effects. With regard to gender, this is in line with what has been reported for antidepressants versus CBT (Cuijpers, Weitz, et al., 2014).

Participant Diversity and Generalizability

The studies included in this IPD meta-analysis were conducted in Canada, Finland, Iran, Mexico, the Netherlands, and the USA. Despite this geographical diversity, no low-income countries were represented. The majority of the participants were middle-aged women. Only one study reported on the participants' ethnicity, indicating that 51.9% were of a self-designated minority group. Other aspects of diversity, such as sexual or gender orientation, physical or cognitive disabilities, social class or intersectionality were not reported in any study. Socioeconomic status was inconsistently described across the studies, thereby limiting the ability for comparison. Hence, the generalizability of our findings is

primarily constrained to populations similar to those represented in our analyses, namely middle-aged women from moderate- to high-income countries.

Strengths and Limitations

This study has a number of strengths. First, to the best of our knowledge, this is the first study that examined a range of potential moderators across trials comparing antidepressants and STPP. It is also the first IPD meta-analysis assessing moderators of antidepressants versus a psychological treatment for depression that focuses on a psychotherapy outside the cognitive-behavioral school (Weitz et al., 2015; Cuijpers, Weitz, et al., 2014; Furukawa et al., 2018). Second, the included studies shared similarities in sample recruitment, depression inclusion criteria, treatment format, and primary outcome measure. Third, and most importantly, the IPD meta-analytic methods allowed for examining moderators at the participant level with increased statistical power (Lambert et al., 2002) and for improving the quality of the analyses when compared to conventional meta-analytic methods. For instance, we facilitated standardization across studies by using the same statistical method for data analysis, appropriately adjusting for baseline values in all studies (Twisk et al., 2018).

This study also has a number of limitations. First, the total number of participants included in this meta-analysis is modest. Relatively few trials compared the efficacy of STPP and antidepressants and the most recent study was conducted over ten years ago. We think this reflects that psychodynamic therapy, in general, has been studied less extensively than other forms of psychotherapy for depression, such as CBT (e.g., Cuijpers, Karyotaki, et al., 2014; Cuijpers, Miguel, Harrer, Plessen, Ciharova, Ebert, et al., 2023), which Abbass et al. (2017) have argued might be due to the appraisal of psychodynamic evidence in treatment guidelines and an underrepresentation of psychodynamic researchers on funding committees. Although the current study was sufficiently powered to identify treatment outcome

differences with a small effect size and baseline depression severity as a moderator, it might have lacked the statistical power to identify more subtle treatment effects and weaker moderator relationships, particularly at follow-up. Second, we were not able to obtain IPD for two studies, though including these two studies in a two-stage IPD meta-analysis did not change the pattern of results. Third, due to the limited number of studies examining SNRIs, we were unable to conduct a subgroup analysis to investigate potential treatment differences between antidepressant types. The results of the current study might not be generalizable to comparisons of STPP and tricyclic antidepressants. Related, although the antidepressants in the included studies are still commonly used today, other types of antidepressants are also used and novel types are being developed (Sakurai, 2022), to which the results of this study might not generalize to. Fourth, not all included studies were free from risk of selection, detection, and attrition bias, though the moderating effect of baseline depression severity appeared robust against control for these risks of bias. The number of available studies was too small to reliably assess publication bias. Fifth, the included studies differed in some respects, such as the STPP model applied, the antidepressant type, and follow-up length. Despite potential differences in specific techniques used among the STPP models, they all share the common STPP principles of treating psychopathology by decreasing maladaptive defenses and increasing insight into unconscious conflicts or avoided emotions that arose as a consequence of past negative experiences (Charman, 2004; Eppel, 2018). Regardless of these differences, moderator effects could be identified in the combined studies' data. Sixth, we could not assess all potential moderator variables of interest, because they were not assessed consistently across the trials. For instance, we were not able to examine the moderating effects of minority status in combination with gender reported previously (Barber et al., 2012), because minority status was not assessed in the other studies. Related, not all moderator variables were assessed in all studies. Thus, the individual moderator models can

relate to different subgroups of studies. Seventh, while the primary outcome measure used in all included studies (HAMD) is an observer-rated instrument and frequently applied in depression treatment research, it has also been criticized for its emphasis on anxiety and somatic symptoms of depression and other conceptual and psychometric flaws (Bagby et al., 2004). Eighth, while the primary studies' eligibility criteria did not exclude individuals with very severe depression (i.e., HAMD > 33), such participants were not represented in the current studies sample and therefore it is unknown whether our findings generalize to these individuals. Ninth, a substantial number of participants had missing outcomes, although mixed-effect models are considered rather robust to this (Twisk et al., 2013). Tenth, the included studies did not consistently collect information on whether participants sought additional treatment during the follow-up period, preventing us from accounting for this in the analyses. Eleventh, and most importantly, the moderating effect of baseline depression severity is observational and needs external validation in prospective clinical trials.

Clinical and Research Implications

The findings of this study suggest that individuals with depression and their clinicians might expect similar improvements in self-reported depression, anxiety, general psychopathology, and general health after treatment with antidepressants and STPP, and slightly greater improvements in observer-rated depression at the end of treatment with antidepressants rather than STPP. However, the 1.88-point difference found in the analyses with unstandardized HAMD-17, scores falls short of the 3 to 8-point threshold on the HAMD, which has been argued to represent the minimum for a clinically significant difference (Hengartner & Plöderl, 2022). For those showing relatively high baseline depression severity, antidepressants might lead to greater improvements in terms of post-treatment depressive symptom levels compared to STPP. It is important to note, however, that major depressive disorder (MDD) diagnoses were based on DSM-IV criteria in most studies, which slightly

differ from DSM-5 criteria. The most notable changes are the removal of the bereavement exclusion and the allowance for past manic episodes in DSM-5 (Uher et al., 2013), leading to a slightly more heterogeneous MDD population in current clinical practice. Uher et al. (2013) have argued that the influence of the DSM-5 revision might be limited due to the non-different symptom profiles, family history, and particularly, the similar response to treatment of non-bereavement-related MDD and bereavement-related MDD. As such the findings of the current study will likely generalize to the majority of patients diagnosed using DSM-5 MDD criteria, although it is unclear to what extent they will generalize to those who have experienced depression following bereavement or who have had prior manic episodes. Additionally, our findings need to be interpreted with caution given the limitations of this study previously discussed.

The finding that patients with relatively high baseline depression severity benefit slightly more from antidepressants appears to be in line with current treatment guidelines and clinical practice, where antidepressants are often recommended for the treatment of severely depressed patients (APA, 2019; NICE, 2022). The findings of this study cannot, however, be taken to imply that antidepressant monotherapy should be considered the first-choice treatment for these individuals, as this study does not speak to the comparative efficacy of antidepressants versus other depression treatments, such as CBT or combined treatment of antidepressants and psychotherapy. For instance, recent IPD meta-analyses found combined treatment of antidepressants and STPP significantly more efficacious than antidepressants alone in terms of depressive symptom levels at both post-treatment and follow-up, particularly for more severely depressed patients (Driessen et al., 2020, 2022).

Given the clinical importance of the research question and the limitations of the current study previously mentioned, additional research on the efficacy and moderators of antidepressants versus STPP in the treatment of depression is warranted. Specifically, the

field would benefit from additional large-scale rigorously conducted randomized clinical trials that assess a range of potential moderator and outcome variables, utilize standardized treatment manuals, and monitor therapists' adherence to these manuals. Potential moderator variables should minimally include baseline depression severity, levels of overall functioning (GAF), gender, and minority status. The assessment of other diversity-related participant characteristics is also important to enable comparisons and evaluations of how findings generalize to different patient groups. Data from such future trials are important to validate the current findings. If validated, the findings of this study would suggest that baseline depression severity is a factor to consider when choosing between STPP and antidepressants. Such knowledge can be used in the future to facilitate evidence-based personalized treatment selection for depression (Deisenhofer et al., 2024) and make more efficient use of existing depression treatment resources.

References

References marked with an asterisk indicate studies identified/included in the meta-analysis.

Abbass, A. A., Kisely, S. R., Town, J. M., Leichsenring, F., Driessen, E., De Maat, S., Gerber, A., Dekker, J. J. M., Rabung, S., Rusalovska, S., & Crowe, E. (2014). Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database of Systematic Reviews*, *7*, 1-108. https://doi.org/10.1002/14651858.CD004687.pub4

Abbass, A. A., Luyten, P., Steinert, C., & Leichsenring, F. (2017). Bias toward psychodynamic therapy: Framing the problem and working toward a solution. *Journal of Psychiatric Practice*, *23*(5), 361–365. https://doi.org/10.1097/PRA.0000000000000256

Aiken, L. S., West, S. G., & Reno, R. R. (1991) *Multiple regression: Testing and interpreting interactions.* Sage.

American Psychological Association. (2019). *Clinical practice guideline for the treatment of depression across three age cohorts*. Retrieved from https://www.apa.org/depression-guideline

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Bagby, R. M., Ryder, A. G., Schuller, D. R., & Marshall, M. B. (2004). The Hamilton depression rating scale: Has the gold standard become lead weight? *American Journal of Psychiatry, 161*(12), 2163-2177. https://doi.org/10.1176/appi.ajp.161.12.2163

* Barber, J. P., Barrett, M. S., Gallop, R., Rynn, M. A., & Rickels, K. (2012). Short-term dynamic psychotherapy versus pharmacotherapy for major depressive disorder: A randomized, placebo-controlled trial. *Journal of Clinical Psychiatry, 73*(1), 66-73. https://doi.org/10.4088/JCP.11m06831

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*(6), 561–571.

Bellak, L. (1993). *Manual de psicoterapia breve, intensiva y de urgencia* [Manual of brief, intensive, and emergency psychotherapy]. Manual Moderno.

Bellak, L. (1994). *Manual para la evaluación de las funciones del yo* [Manual for the assessment of the ego functions]. Manual Moderno.

Bet, P. M., Hugtenburg, J. G., Penninx, B. W., & Hoogendijk, W. J. (2013). Side effects of antidepressants during long-term use in a naturalistic setting. *European*

Neuropsychopharmacology, 23(11), 1443–1451. https://doi.org/10.1016/j.euroneuro.2013.05.001

Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A., & Peters, T. J. (2004). Subgroup analyses in randomized trials: Risks of subgroup-specific analyses; power and sample size for the interaction test. *Journal of Clinical Epidemiology, 57*(3), 229-236. https://doi.org/10.3310/hta5330

Charman, P. C. (Ed.). (2004). *Core processes in brief psychodynamic psychotherapy. Advancing effective practice*. Lawrence Erlbaum Associates.

Cohen, Z. D., Breunese, J., Markowitz, J. C., Weitz, E. S., Hollon, S. D., Browne, D. T., Rucci, P., Corda, C., Menchetti, M., Weissman, M. M., Bagby, R. M., Quilty, L. C., Blom, M. B. J., Altamura, M., Zobel, I., Schramm, E., Gois, C., Twisk, J. W. R., Wienicke, F. J., ... Driessen, E. (2024). Comparative efficacy of interpersonal psychotherapy and antidepressant medication for adult depression: A systematic review and individual participant data meta-analysis. *Psychological Medicine, 54*(14), 1–10. https://doi.org/10.1017/S0033291724001788

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders, 159,* 118-126. https://doi.org/10.1016/j.jad.2014.02.026

Cuijpers, P., Miguel, C., Harrer, M., Plessen, C. Y., Ciharova, M., Ebert, D., & Karyotaki, E. (2023). Cognitive behavior therapy vs. control conditions, other psychotherapies, pharmacotherapies, and combined treatment for depression: A comprehensive meta-analysis including 409 trials with 52,702 patients. *World Psychiatry, 22*(1), 105–115. https://doi.org/10.1002/wps.21069

Cuijpers, P., Miguel, C., Harrer, M., Plessen, C. Y., Ciharova, M., Papola, D., Ebert, D., & Karyotaki, E. (2023). Psychological treatment of depression: A systematic overview of a 'meta-analytic research domain'. *Journal of Affective Disorders, 335*, 141-151. https://doi.org/10.1016/j.jad.2023.05.011

Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., Cipriani, A., Cristea, I. A., & Furukawa, T. A. (2021). Psychotherapies for depression: A network meta-analysis covering efficacy, acceptability, and long-term outcomes of all main treatment types. *World Psychiatry, 20*(2), 283–293. https://doi.org/10.1002/wps.20860

Cuijpers, P., Weitz, E., Twisk, J. W. R., Kuehner, C., Cristea, I., David, D., DeRubeis, R. J., Dimidjian, S., Dunlop, B. W., Faramarzi, M., Hegerl, U., Jarrett, R. B., Kennedy, S. H., Kheirkhah, F., Mergl, R., Miranda, J., Mohr, D. C., Segal, Z. V., Siddique, J., ... Hollon, S. D. (2014). Gender as predictor and moderator of outcome in cognitive behavior therapy and pharmacotherapy for adult depression: An "individual patient data" meta-analysis. *Depression and Anxiety, 31*(11), 941-951. https://doi.org/10.1002/da.22328

de Jonghe, F., de Maat, S. C. M., Van, H. L., Hendriksen, M., Kool, S., van Aalst, G., & Dekker, J. J. M. (2013). Short-term psychoanalytic supportive psychotherapy for depression. *Psychoanalytic Inquiry, 33*(6), 614-625. https://doi.org/10.1080/07351690.2013.835184

Deisenhofer, A. K., Barkham, M., Beierl, E. T., Schwartz, B., Aafjes-van Doorn, K., Beevers, C. G., Berwian, I., Blackwell, S., Bockting, C., Brakemeier, E., Brown, G., Buckman, J. E. J., Castonguay, L. G., Cusack, C., Dalgleish, T., de Jong, K., Delgadillo, J., DeRubeis, R., Driessen, E., ..., Cohen, Z. D. (2024). Implementing precision methods in personalizing psychological therapies: Barriers and possible ways forward. *Behaviour Research and Therapy, 172*. https://doi.org/10.1016/j.brat.2023.104443

* Dekker, J. J. M., Koelen, J. A., Van, H. L., Schoevers, R. A., Peen, J., Hendiksen, M., Kool, S., van Aalst, G., de Jonghe, F. (2008). Speed of action: The relative efficacy of short psychodynamic supportive psychotherapy and pharmacotherapy in the first 8 weeks of a treatment algorithm for depression. *Journal of Affective Disorders, 109*(1-2), 183-188. https://doi.org/10.1016/j.jad.2007.10.015

Della Selva, P. C. (1996). *Intensive short-term dynamic psychotherapy: Theory and technique*. John Wiley & Sons.

Driessen, E., Abbass, A. A., Barber, J. P., Connolly Gibbons, M. B., Dekker, J. J. M., Fokkema, M., Fonagy, P., Hollon, S. D., Jansma, E. P., de Maat, S. C. M., Town, J. M., Twisk, J. W. R., Van, H. L., Weitz, E., & Cuijpers, P. (2018). Which patients benefit specifically from short-term psychodynamic psychotherapy (STPP) for depression? Study protocol of a systematic review and meta-analysis of individual participant data. *BMJ Open, 8*(2), e018900. https://doi.org/10.1136/bmjopen-2017-018900

Driessen, E., Dekker, J. J. M., Peen, J., Van, H. L., Maina, G., Rosso, G., Rigardetto, S., Cuniberti, F., Vitriol, V. G., Florenzano, R. U., Adreoli, A., Burnand, Y., López Rodríguez, J., Vilamil Salcedo, V., Twisk, J. W. R., & Cuijpers, P. (2020). The efficacy of adding short-term psychodynamic psychotherapy to antidepressants in the treatment of depression: A systematic review and meta-analysis of individual participant data. *Clinical Psychology Review, 80*. https://doi.org/10.1016/j.cpr.2020.101886

Driessen, E., Fokkema, M., Dekker, J. J., Peen, J., Van, H. L., Maina, G., Rosso, G., Rigardetto, S., Cuniberti, F., Vitriol, V. G., Adreoli, A., Burnand, Y., López Rodríguez, J., Vilamil Salcedo, V., Twisk, J. W. R., Wienicke, F. J., & Cuijpers, P. (2023). Which patients benefit from adding short-term psychodynamic psychotherapy to antidepressants in the treatment of depression? A systematic review and meta-analysis of individual

participant data. *Psychological Medicine, 53,* 6090-6101. https://doi.org/10.1017/S0033291722003270

Driessen, E., Hegelmaier, L. M., Abbass, A. A., Barber, J. P., Dekker, J. J. M., Van, H. L., Jansma, E. P., & Cuijpers, P. (2015). The efficacy of short-term psychodynamic psychotherapy for depression: A meta-analysis update. *Clinical Psychology Review, 42,* 1-15. https://doi.org/10.1016/j.cpr.2015.07.004

Eppel, A. (2018). *Short-term psychodynamic psychotherapy*. Springer. https://doi.org/10.1007/978-3-319-74995-2

Faravelli, C., Albanesi, G., & Poli, E. (1986). Assessment of depression: A comparison of rating scales. *Journal of Affective Disorders, 11*(3), 245-253. https://doi.org/10.1016/0165-0327(86)90076-5

Ferguson, J. M. (2001). SSRI antidepressant medications: Adverse effects and tolerability. *Primary Care Companion to the Journal of Clinical Psychiatry, 3*(1), 22. https://doi.org/10.4088/pcc.v03n0105

Furukawa, T. A., Efthimiou, O., Weitz, E. S., Cipriani, A., Keller, M. B., Kocsis, J. H., Klein, D. N., Michalak, J., Salanti, G., Cuijpers, P., & Schramm, E. (2018). Cognitive-behavioral analysis system of psychotherapy, drug, or their combination for persistent depressive disorder: Personalizing the treatment choice using individual participant data network metaregression. *Psychotherapy and Psychosomatics, 87*(3), 140-153. https://doi.org/10.1159/000489227

Furukawa, T. A., Reijnders, M., Kishimoto, S., Sakata, M., DeRubeis, R. J., Dimidjian, S., Dozois, D. J. A., Hegerl, U., Hollon, S. D., Jarrett, R. B., Lespérance, F., Segal, Z. V., Mohr, D. C., Simons, A. D., Quilty, L. C., Reynolds III, C. F., Gentili, C., Leucht, S., Engel, R. R., & Cuijpers, P. (2020). Translating the BDI and BDI-II into the HAMD and

vice versa with equipercentile linking. *Epidemiology and Psychiatric Sciences, 29*, e24. https://doi.org/10.1017/S2045796019000088

Hengartner, M. P., & Plöderl, M. (2022). Estimates of the minimal important difference to evaluate the clinical significance of antidepressants in the acute treatment of moderate-to-severe depression. *BMJ Evidence-Based Medicine, 27*(2), 69-73. https://doi.org/10.1136/bmjebm-2020-111600

Heo, M., & Leon, A. C. (2010). Sample sizes required to detect two-way and three-way interactions involving slope differences in mixed-effects linear models. *Journal of Biopharmaceutical Statistics, 20*(4), 787-802. https://doi.org/10.1080/10543401003618819

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343*(7829), 1-9. https://doi.org/10.1136/bmj.d5928

Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society, Series A, Statics in Society, 172*(1), 137-159. https://doi.org/10.1111/j.1467-985X.2008.00552.x

Lambert, P. C., Sutton, A. J., Abrams, K. R., & Jones, D. R. (2002). A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of Clinical Epidemiology, 55*(1), 86-94. https://doi.org/10.1016/s0895-4356(01)00414-0

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*(12). https://doi.org/10.1037/0003-066X.48.12.1181

* López Rodríguez, J., López Butrón, M. A., Vargas Terrez, B. E., & Villamil Salcedo, V. (2004). Estudio doble ciego con antidepresivo, psicoterapia breve y placebo en pacientes

con depresión leve a moderada [Double-blind study with antidepressant, brief psychotherapy and placebo in patients with mild to moderate depression]. *Salud Mental, 27*(5), 53–61.

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy: Manual for supportive-expressive treatment*. Basic Books Inc Publishers.

Luborsky, L., Mark, D., Hole, A. V., Popp, C., Goldsmith, B., & Cacciola, J. (1995). Supportive-expressive dynamic psychotherapy of depression: A time-limited version. In J. P. Barber, & P. Crits-Christoph (Eds.), *Psychodynamic psychotherapies for psychiatric disorders (Axis I)* (pp. 13–42). Basic Books.

Luyten, P., & Blatt, S. J. (2012). Psychodynamic treatment of depression. *Psychiatric Clinics of North America, 35*(1), 111-129.https://doi.org/10.1016/j.psc.2012.01.001

Ma, S., Kang, L., Guo, X., Liu, H., Yao, L., Bai, H., Chen, C., Hu, M., Du, L., Du, H., Ai, C., Wang, F., Wang, G., Li, R., & Liu, Z. (2021). Discrepancies between self-rated depression and observed depression severity: The effects of personality and dysfunctional attitudes. *General Hospital Psychiatry, 70*, 25-30. https://doi.org/10.1016/j.genhosppsych.2020.11.016

Malan, D. H. (1976). *The frontier of brief psychotherapy: An example of the convergence of research and clinical practice*. Plenum Medical Book Co.

Mann, J. (1973). *Time-limited psychotherapy*. Harvard University Press.

Marcus, S. C., & Olfson, M. (2010). National trends in the treatment of depression from 1998 to 2007. *Archives of General Psychiatry, 67*(12), 1265-1273. https://doi.org/10.1001/archgenpsychiatry.2010.151

Marmor, J. (1973). *Psychiatry in transition.* Brunner/Mazel.

Marmor, J. (1975). The nature of the psychotherapeutic process' revisited. *The Canadian Psychiatric Association Journal, 20*(8), 557-565.

McHugh, R. K., Whitton, S. W., Peckham, A. D., Welge, J. A., & Otto, M. W. (2013). Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: A meta-analytic review. *Journal of Clinical Psychiatry, 74*(6), 13979. https://doi.org/10.4088/JCP.12r07757

* McLean, P. D., & Hakstian, A. R. (1979). Clinical depression: Comparative efficacy of outpatient treatments. *Journal of Consulting and Clinical Psychology, 47*(5), 818-836. https://doi.org/10.1037/0022-006X.47.5.818

National Institute for Health and Care Excellence (2022, June). *Depression in adults: Treatment and management*. Retrieved from https://www.nice.org.uk/guidance/ng222

Norcross, J. C., & Rogan, J. D. (2013). Psychologists conducting psychotherapy in 2012: Current practices and historical trends among division 29 members. *Psychotherapy, 50*(4), 490–495. https://doi.org/10.1037/a0033512

Orlinsky, D. E., Messina, I., Hartmann, A., Willutzki, U., Heinonen, E., Rønnestad, M. H., Löffler-Stastka, H., & Schröder, T. (2024). Ninety psychotherapy training programmes across the globe: Variations and commonalities in an international context. *Counselling & Psychotherapy Research, 24*(2), 558–571. https://doi.org/10.1002/capr.12690

Orlinsky, D. E., Rønnestad, M. H., Hartmann, A., Heinonen, E., & Willutzki, U. (2020). The personal self of psychotherapists: Dimensions, correlates, and relations with clients. *Journal of Clinical Psychology, 76*(3), 461–475. https://doi.org/10.1002/jclp.22876

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

* Rahmani, F., Rezaei, F., & Nasuri, M. (2011). Major depressive disorder: A comparison of a focused psychodynamic psychotherapy and pharmacotherapy. *Iranian Journal of Psychiatry and Behavioral Sciences, 50*(1), 37-44.

Riedel, M., Möller, H. J., Obermeier, M., Schennach-Wolff, R., Bauer, M., Adli, M., Kronmüller, K., Nickel, T., Brieger, P., Laux, G., Bender, W., Heuser, I., Zeiler, J., Gaebel, W., & Seemüller, F. (2010). Response and remission criteria in major depression – A validation of current practice. *Journal of Psychiatric Research, 44*(15), 1063-1068. https://doi.org/10.1016/j.jpsychires.2010.03.006

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ, 342*. https://doi.org/10.1136/bmj.d549

Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ, 340*, c221. https://doi.org/10.1136/bmj.c221

Riley, R. D., Simmonds, M. C., & Look, M. P. (2007). Evidence synthesis combining individual patient data and aggregate data: A systematic review identified current practice and possible methods. *Journal of Clinical Epidemiology, 60*(5), 431–439. https://doi.org/10.1016/j.jclinepi.2006.09.009

Rueve, M. E., & Correll, T. L. (2006). The art of psychotherapy: Selecting patients for psychodynamic psychotherapy. *Psychiatry (Edgmont), 3*(11), 44-55.

Sakurai, H., Yonezawa, K., Tani, H., Mimura, M., Bauer, M., & Uchida, H. (2022). Novel antidepressants in the pipeline (phase II and III): A systematic review of the US clinical trials registry. *Pharmacopsychiatry, 55*(4), 193–202. https://doi.org/10.1055/a-1714-9097

* Salminen, J. K., Karlsson, H., Hietala, J., Kajander, J., Aalto, S., Markkula, J., Rasi Hakala, H., & Toikka, T. (2008). Short-term psychodynamic psychotherapy and fluoxetine in major depressive disorder: A randomized comparative study. *Psychotherapy and Psychosomatics, 77*(6), 351-357. https://doi.org/10.1159/000151388

Stewart, L. A., & Parmar, M. K. B. (1993). Meta-analysis of the literature or of individual patient data: Is there a difference? *Lancet, 341*(8842), 418-422. https://doi.org/10.1016/0140-6736(93)93004-k

Targum, S. D., Wedel, P. C., Robinson, J., Daniel, D. G., Busner, J., Bleicher, L. S., Rauh, P., & Barlow, C. (2013). A comparative analysis between site-based and centralized ratings and patient self-ratings in a clinical trial of major depressive disorder. *Journal of Psychiatric Research, 47*(7), 944–954. https://doi.org/10.1016/j.jpsychires.2013.02.016

Tudur Smith, C., Marcucci, M., Nolan, S. J., Iorio, A., Sudell, M., Riley, R., Rovers, M. M., & Williamson, P. R. (2016). Individual participant data meta-analyses compared with meta-analyses based on aggregate data. *Cochrane Database of Systematic Reviews,* (9). https://doi.org/10.1002/14651858.MR000007.pub3

Turner, E. H., Matthews, A. A., Linardatos, E., Tell R. A., & Rosenthal R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine, 358*(3), 252–260. https://doi.org/10.1056/NEJMsa065779

Twisk, J. W. R., Bosman, L., Hoekstra, T., Rijnhart, J., Welten, M., & Heymans, M. (2018). Different ways to estimate treatment effects in randomised controlled trials. *Contemporary Clinical Trials Communications, 10,* 80-85. https://doi.org/10.1016/j.conctc.2018.03.008

Twisk, J. W. R., de Boer, M., de Vente, W., & Heymans, M. (2013). Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. *Journal of Clinical Epidemiology, 66*(9), 1022-1028. https://doi.org/10.1016/j.jclinepi.2013.03.017

Uher, R., Payne, J. L., Pavlova, B., & Perlis, R. H. (2013). Major depressive disorder in DSM-5: Implications for clinical practice and research of changes from DSM-IV. *Depression and Anxiety, 31*(6), 459-471. https://doi.org/10.1002/da.22217

Valbak, K. (2004). Suitability for psychoanalytic psychotherapy: A review. *Acta Psychiatrica Scandinavica, 109*(3), 164-178. https://doi.org/10.1046/j.1600-0447.2003.00248.x

van Schaik, D. J. F., Klijn, A. F. J., van Hout, H. P. J., van Marwijk, H. W. J., Beekman, A. T. F., de Haan, M., & van Dyck, R. (2004). Patients' preferences in the treatment of

depressive disorder in primary care. *General Hospital Psychiatry, 26*(3), 184–189. https://doi.org/10.1016/j.genhosppsych.2003.12.001

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Weitz, E. S., Hollon, S. D., Twisk, J., van Straten, A., Huibers, M. J., David, D., DeRubeis, R. J., Dimidjian, S., Dunlop, B. W., Cristea, I. A., Faramarzi, M., Hegerl, U., Jarrett, R. B., Kheirkhah, F., Kennedy, S. H., Mergl, R., Miranda, J., Mohr, D. C., Rush, A. J., ... Cuijpers, P. (2015). Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: an individual patient data meta-analysis. *JAMA Psychiatry, 72*(11), 1102-1109. https://doi.org/10.1001/jamapsychiatry.2015.1516

Wolberg, L. R. (1967). *Short-term psychotherapy*. Grune & Stratton.

World Health Organization. (2022). *World mental health report: Transforming mental health for all*. Geneva: World Health Organization

Table 1

Characteristics of Identified Studies

Study	Country	Depression diagnosis	*N*	% Female	M_{Age}	N_{SE}	STPP model	HAMD version	Antidepressant type and dosage [a]	Post-treatment	Follow-up
					IPD available:						
Barber et al., 2012	USA	Major Depressive Disorder (DSM-IV); HAMD ≥ 14	106	60.8	36.2	20	Luborksy, 1984, 1995	17-item	1. Sertraline max. 200 mg/d (SSRI), 2. Venlafaxine max. 375 mg/d (SNRI)	16 weeks	28 months
Dekker et al., 2008	NL	Depressive Episode (DSM-IV); HAMD = 12-26	141	73.8	37.0	8	de Jonghe et al., 2005, 2013	17-item	Venlafaxine max. 225 mg/d (SNRI)	8 weeks	-
López Rodríguez et al., 2004	Mexico	Mild to moderate depression (DSM-IV and ICD-10)	20	70	32.0	20	Bellak, 1993, 1994	21-item	Fluoxetine 20 mg/day (SSRI)	20 weeks	10 months
Salminen et al., 2008	Finland	Major Depressive Disorder (DSM-IV); HAMD ≥ 15	51	68.8	42.4	16	Mann, 1973; Malan, 1976,	17-item	Fluoxetine 20-40 mg/day (SSRI)	16 weeks	12 months
					IPD not available:						
McLean & Hakstian, 1979	Canada	Primary depression (RDC)	104	72 [b]	39.2 [b]	8-12	Marmor, 1973, 1975; Wolberg, 1967	-	Amitriptyline 150 mg/day (TCA)	10 weeks	6 months
Rahmani et al., 2011	Iran	Major depressive episode (DSM-IV); MMPI-2 depression = 70-100; BDI = 17-63	50	63.6 [c]	- [d]	15	Della Selva, 1996	-	Citalopram 40 mg/day (SSRI)	15 weeks	-

Note. BDI = Beck Depression Inventory; DSM = Diagnostic and Statistical Manual; HAMD = Hamilton Depression Rating Scale; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, 10th edition; IPD = individual participant data; MMPI-2 = Minnesota Multiphasic Personality Inventory-2; *N*se = number of STPP sessions; RDC = Research Diagnostic Criteria; SNRI = Serotonine Noradrenaline Reuptake Inhibitors; SSRI = Selective Serotonin Reuptake Inhibitor; STPP = short-term psychodynamic psychotherapy; TCA = Tricyclic Antidepressant.

[a] Numbers indicate sequential steps in the antidepressants protocol.

[b] Demographics were only reported for the whole study sample, including participants from a behavior therapy and control condition.

[c] The percentage of females in the sample was only reported for the completer sample.

[d] The age of the sample was provided in age categories for only the completer sample.

Table 2

Risk of Bias Ratings of Identified Studies

Study	Selection bias		Detection bias	Attrition bias
	Random sequence generation	Allocation concealment	Blinding of outcome assessment	Complete outcome data
Barber et al., 2012	+	+	+	+
Dekker et al., 2008	+	-	-	-
López Rodríguez et al., 2004	+	+	+	-
Salminen et al., 2008	?	-	-	+
McLean & Hakstian, 1979	-	-	n/a	-
Rahmani et al., 2011	?	?	n/a	-

Note. + = low risk of bias, - = high risk of bias, ? = Information not retrievable, n/a = not applicable (study assessed only self-report instruments).

Table 3

Treatment Effects of Antidepressants Versus STPP at Post-treatment and Follow-up

Assessment moment	Outcome	*k*	*N*	*d*	95% CI	*p*	I^2
Post-treatment	Depressive symptom levels	**4**	**310**	**0.28**	**0.03, 0.53**	**.031**	**0**
	Self-reported depression	3	271	0.25	-0.02, 0.52	.075	0
	Anxiety	2	214	0.08	-0.22, 0.38	.599	0
	General psychopathology	3	250	0.17	-0.15, 0.49	.303	0
	Physical Health	3	243	-0.08	-0.36, 0.20	.565	0
Follow-up	Depressive symptom levels	3	177	0.31	-0.08, 0.72	.130	0

Note. STPP = Short-term psychodynamic psychotherapy.

Positive signs indicate better outcomes in the antidepressant than in the STPP condition and negative signs indicate better outcomes in the STPP than in the antidepressant condition.

Statistical significance ($p < .05$) is indicated by bold printed numbers.

Table 4

Cohen's d Effect Sizes on Depressive Symptom Levels of Antidepressants versus STPP for the Different Moderator Levels

Moderator	Post-treatment					Follow-up				
	k	*N*	*d*	95%CI	*p*	*k*	*N*	*d*	95%CI	*p*
Gender	4	310				3	177			
Male			0.26	-0.18, 0.70	.254			0.38	-0.33, 1.12	.302
Female			0.28	-0.02, 0.59	.079			0.28	-0.23, 0.78	.263
Marital Status	4	302				3	176			
Single, divorced, widowed			0.30	-0.02, 0.63	.070			0.14	-0.44, 0.72	.633
Married/cohabiting			0.26	-0.15, 0.67	.218			0.45	-0.11, 1.02	.467
Education	4	277				3	158			
High-school non-graduate			0.16	-0.41, 0.73	.589			0.43	-0.21, 1.10	.191
High-school graduate			0.33	0.03, 0.62	.035			0.22	-0.30, 0.74	.411
Employment Status	3	174				3	174			
Unemployed, student, retired, homemaker			0.05	-0.50, 0.60	.860			0.11	-0.63, 0.82	.759
Employed			0.19	-0.25, 0.64	.413			0.42	-0.08, 0.93	.100
Episode duration	2	201								
< 4 weeks			0.19	-0.37, 0.73	.495			-	-	-
4 weeks to 1 year			0.58	0.09, 1.08	.019			-	-	-
1 year to 2 years			0.30	-0.43, 1.06	.433			-	-	-
> 2 years			0.19	-0.36, 0.74	.495			-	-	-
Age	4	310				3	177			
Average			0.28	0.02, 0.54	.032			0.29	-0.11, 0.70	.165
Per year increase			0.01	-0.02, 0.03	.578			0.01	-0.03, 0.04	.542
Baseline depression severity	4	310				3	177			
Average			**0.28**	**0.13, 0.44**	**<.001**			0.28	-0.002, 0.58	.061

Per SD increase			**0.24**	**0.08, 0.40**	**.004**			0.17	-0.14, 0.49	.286
Baseline self-reported depression symptoms	3	267				2	150			
Average			0.30	0.03, 0.55	.027			0.32	-0.14, 0.79	.176
Per SD increase			-0.01	-0.27, 0.25	.962			0.03	-0.45, 0.52	.906
Baseline anxiety symptoms	2	210								
Average			0.37	0.08, 0.67	.015			-	-	-
Per SD increase			0.04	-0.25, 0.32	.809			-	-	-
Baseline general psychopathology level	3	247				2	146			
Average			0.25	-0.02, 0.52	.070			0.35	-0.13, 0.82	.159
Per SD increase			0.20	-0.08, 0.48	.163			0.12	-0.33, 0.55	.598
Baseline physical health level	3	224				2	140			
Average			0.24	-0.06, 0.52	.117			0.30	-0.25, 0.82	.269
Per SD increase			-0.20	-0.50, 0.09	.183			0.06	-0.46, 0.59	.824
Baseline comorbid PD symptoms	2	165								
Average			0.24	-0.08, 0.56	.157			-	-	-
Per SD increase			-0.22	-0.54, 0.11	.194			-	-	-
GAF	3	252				2	155			
Average			**0.21**	**-0.05, 0.48**	**.122**			0.17	-0.31, 0.65	.496
Per point increase			**0.04**	**0.001, 0.08**	**.045**			0.04	-0.03, 0.11	.222

Note. GAF = Global Assessment of Functioning Scale; PD = Personality disorder; STPP = Short-term psychodynamic psychotherapy.

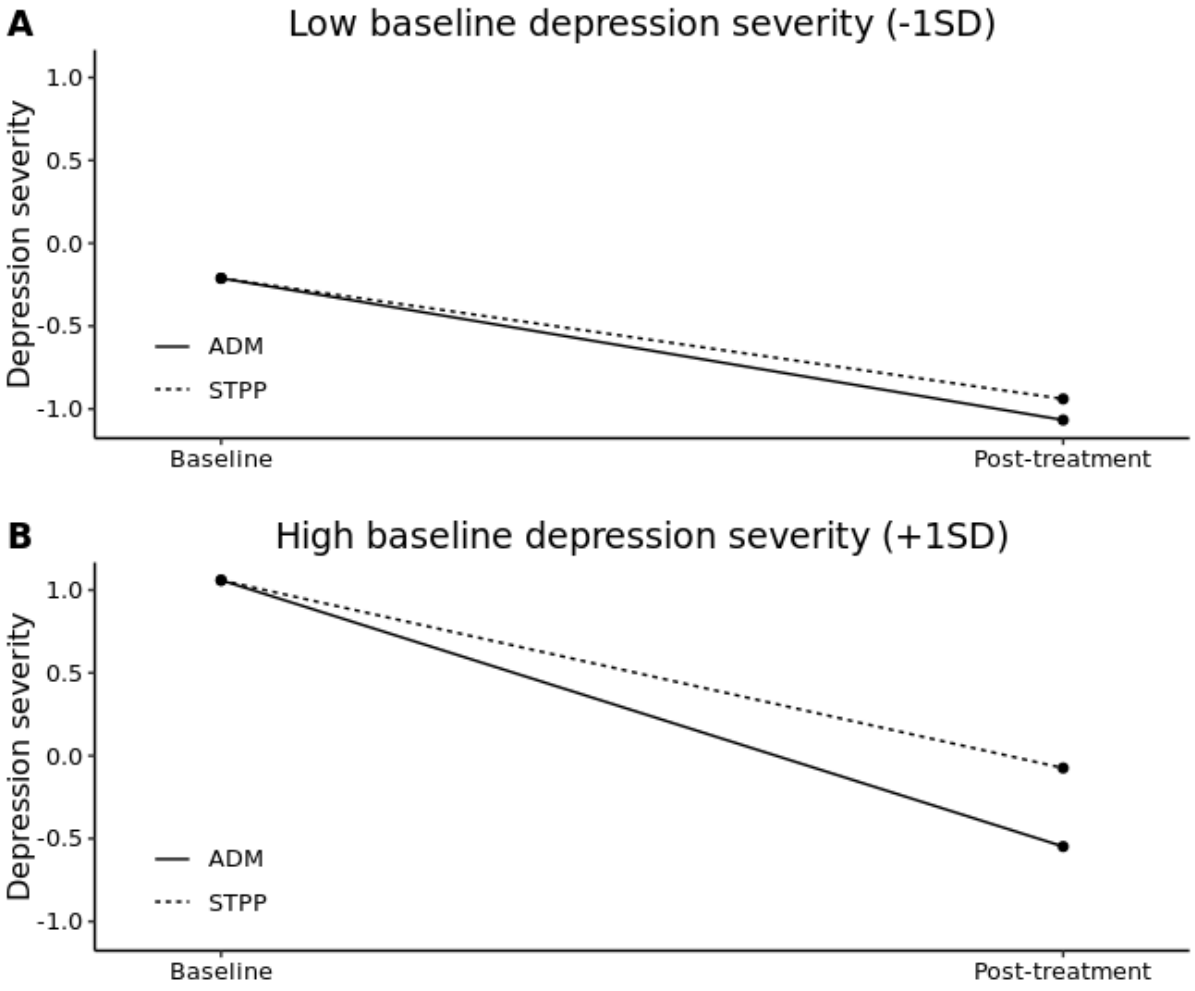
Positive signs indicate better outcomes in the antidepressant than in the STPP condition and negative signs indicate better outcomes in the STPP than in the antidepressant condition.

Bold printed numbers indicate statistical significance ($p < .05$) of the time-by-moderator-by-treatment 3-way interaction. The significance of continuous moderators shows that there is an increasing differential treatment efficacy for each unit increase in baseline values (indicated by "Per … increase"), which is added to the treatment effect for participants with average scores of the study sample (indicated by "Average").

"-" = insufficient data to conduct analyses.

Figure 1

Graphical Illustration of the Baseline Depression Severity Moderation Finding



Note. ADM = Antidepressant medication, STPP = Short-term psychodynamic psychotherapy. Depression severity expressed by z-scores.

Slope estimates for low baseline depression severity: ADM (b = −0.86, *SE* = 0.11, $p < .001$), STPP (b = −0.73, *SE* = 0.10, $p < .001$). Slope estimates for high baseline depression severity: ADM (b = −1.60, *SE* = 0.11, $p < .001$), STPP (b = −1.13, *SE* = 0.10, $p < .001$).