# Psychotherapy Versus Second-Generation Antidepressants in the Treatment of Depression

## *A Meta-Analysis*

*Glen I. Spielmans, PhD,* Margit I. Berman, PhD,† and Ashley N. Usitalo, BA**

**Abstract:** Most meta-analyses have concluded that psychotherapy and pharmacotherapy yield roughly similar efficacy in the short-term treatment of depression, with psychotherapy showing some advantage at long-term follow-up. However, a recent meta-analysis found that selective serotonin reuptake inhibitors medications were superior to psychotherapy in the short-term treatment of depression. To incorporate results of several recent trials into the meta-analytic literature, we conducted a meta-analysis of trials which directly compared psychotherapy to second-generation antidepressants (SGAs). Variables potentially moderating the quality of psychotherapy or medication delivery were also examined, to allow the highest quality comparison of both types of intervention. Bona fide psychotherapies showed equivalent efficacy in the short-term and slightly better efficacy on depression rating scales at follow-up relative to SGA. Non-bona fide therapies had significantly worse short-term outcomes than medication ($d$ = 0.58). No significant differences emerged between treatments in terms of response or remission rates, but non-bona fide therapies had significantly lower rates of study completion than medication (odds ratio = 0.55). Bona fide psychotherapy appears as effective as SGAs in the short-term treatment of depression, and likely somewhat more effective than SGAs in the longer-term management of depressive symptoms.

**Key Words:** Antidepressants, depression, meta-analysis, psychotherapy, selective serotonin reuptake inhibitors.

(*J Nerv Ment Dis* 2011;199: 142–149)

Identifying the most effective treatments for depressive disorders has important public health and economic implications both within the United States and globally. The Global Burden of Disease study concluded that unipolar depression was the fourth most important disease-related cause of disability worldwide, and projected that it would be the second leading disease-related cause of disability by 2020, second only to heart disease (Murray and Lopez, 1996). Treatment of depression is also increasing, primarily due to dramatically increased use of antidepressant medications, especially newer selective serotonin reuptake inhibitor (SSRI) and serotonin/norepinephrine reputake inhibitors (SNRI) medications (Olfson et al., 2002).

Antidepressants are currently the most frequently prescribed class of drugs in ambulatory care settings in the United States (Cherry et al., 2008; Raofi and Schappert, 2006). By 1997, nearly 60% of all depression treatment was via SSRI medication (Olfson et al., 2002). The compound growth rate of psychotropic drug use from 1996 to 2001 was nearly 20%, and more than 50% of the growth in spending on psychotropic medication during that period was due exclusively to second-generation antidepressants (SGAs) (Zuvekas, 2005). A survey completed from February 2001 to April 2003 found that over 80% of persons who had taken antidepressants in the past year had taken an SGA (Mojtabi and Olfson, 2008). Further, from 1996 to 2005, the number of people receiving antidepressants in the United States had doubled and over 90% of people receiving antidepressants received SGAs in 2005 (Olfson and Marcus, 2009). Despite the general increase in depression treatment, the proportion of individuals treated for depression with psychotherapy has declined over time, although psychotherapy continues to be frequently used (Olfson et al., 2002; Marcus and Olfson, 2010).

In studies, which have directly compared psychotherapy and antidepressant medication, several meta-analyses and mega-analyses have found that the 2 treatment modalities yield roughly equivalent benefits at the end point of short-term trials (De Maat et al., 2006; DeRubeis et al., 1999; Gaffan et al., 1995; Gloaguen et al., 1998; Imel et al., 2008; Robinson et al., 1990). Two such recent meta-analyses found that psychotherapy and antidepressants provided similar efficacy in the short-term, but that psychotherapy had superior results in the longer term (DeMaat et al., 2006; Imel et al., 2008). However, the aforementioned meta-analytic comparisons of psychotherapy and medication have consisted exclusively or mostly of trials using tricyclic (TCA) or monoamine oxidase-inhibiting (MAOI) antidepressant medications.

Compared with other antidepressants, SGAs have generally demonstrated equivalent efficacy, though there is evidence that they may be somewhat better tolerated (Anderson, 2000; Anderson and Tomenson, 1995). One might therefore expect that they would perform similarly to other antidepressants in comparison to psychotherapy. However, one meta-analysis (Cuijpers et al., 2008) found that SSRI medications, but not older antidepressants, were more effective than psychotherapy in the treatment of depressive disorders. Although the difference in favor of SSRIs in this study was small in magnitude, it illustrates that relying on indirect comparisons for treatment efficacy data is problematic. Trials comparing SGAs to TCAs/MAOIs and SGAs to psychotherapy may differ from one another significantly in terms of researcher allegiance, qualifications of study therapists, trial duration, concomitant treatment, and other variables that may influence trial outcome. Instead of using indirect comparisons, examining the results of trials which have directly compared treatments is preferable (Shadish and Sweeney, 1991; Spielmans et al., 2010).

Just as different classes of antidepressant medication may have different efficacy in comparison to psychotherapy, differences in the quality of psychotherapy also likely affect treatment outcome. Previous comparisons of different types of psychotherapy, for example, have found that non-bona fide psychotherapeutic interventions—those that are not actually intended to be therapeutic—are less effective than bona fide psychotherapies (Spielmans et al., 2007;

*Department of Psychology, Metropolitan State University, St. Paul, MN; and †Department of Psychiatry, Dartmouth Medical School, Lebanon, NH.

A prior version of this paper was presented at the American Psychological Association's annual convention in Toronto in August 2009.

Glen I. Spielmans has holdings of less than $10,000 in a mutual fund (Vanguard Health Care), which invests nearly exclusively in pharmaceutical companies. Dr. Berman and Ms. Usitalo report no competing interests.

Send reprint requests to Glen I. Spielmans, PhD, Department of Psychology, Metropolitan State University, 1450 Energy Park Drive, St. Paul, MN 55108. E-mail: glen.spielmans@metrostate.edu.



ISSN: 0022-3018/11/19903-0142
DOI: 10.1097/NMD.0b013e31820caefb

Wampold et al., 2002). Unfortunately, non-bona fide psychotherapies—those that are not based on any recognized psychological principles or treatment manuals, are not tailored to clients' issues, do not refer to any active ingredients in the therapeutic intervention, and/or use inadequately trained therapists—are sometimes tested in randomized controlled trials (RCTs) that compare therapy to medication. Only one meta-analysis of which we are aware (Imel et al., 2008) specifically coded studies to ensure that psychotherapies compared with antidepressants were in fact bona fide.

The use of non-bona fide psychotherapies in comparative treatment research is just one example of inadequate treatment quality in RCTs and associated meta-analyses of treatment outcomes; in general, meta-analyses comparing antidepressant medication with psychotherapy have been criticized on methodological grounds, as have the studies which comprise them (Klein, 2000). Besides including non-bona fide therapies, comparative treatment meta-analyses have not consistently examined the influence of variables that affect treatment quality, such as dosage of medication of frequency of study visits, or tapering or switching of treatments. Other variables that may negatively affect treatment quality by introducing bias, such as funding source or allegiance of study authors to the treatment, have also typically not been addressed. As with bona fide psychotherapies, the allegiance of the authors to the treatments under investigation appears important to treatment outcome, as one meta-analysis found that effect size differences between psychotherapy and pharmacotherapy were associated with the allegiance of the researchers (Gaffan et al., 1995). However, this meta-analysis did not examine SGAs.

As noted previously, only one meta-analysis (Cuijpers et al., 2008) has specifically compared psychotherapy to second-generation antidepressants, and since its publication several additional comparative trials of psychotherapy and second-generation antidepressants have been published. The sparse meta-analytic research that compares psychotherapy and newer antidepressant medications thus represents a major gap in the literature. The current meta-analysis therefore updates and extends the investigation of the comparative efficacy of SGAs and psychotherapy, while examining the potential influence of moderator variables that affect the treatment quality of medication or psychotherapy. In addition, we sought to replicate Imel et al.'s finding that psychotherapy was superior to medication at treatment follow-up when comparisons were limited to psychotherapy versus SGAs (Imel et al., 2008).

# METHOD

## Selection of Studies

We included all studies used in the 2 most recent meta-analyses comparing psychotherapy and antidepressants provided that they also met our inclusion criteria described below (Cuijpers et al., 2008; Imel et al., 2008). Cuijpers et al.'s meta-analysis consisted of studies retrieved from a thorough literature search of trials indexed between 1966 and May 2007 (Cuijpers et al., 2008). Their search retrieved 149 studies which examined the psychological treatment of depression in adults; all trials were randomized and included some sort of comparison group (Cuijpers et al., 2008). From this database, we located 9 trials which compared SGA treatment to psychotherapy and met our other inclusion criteria.

Our supplementary search included searches of both Medline and the Cochrane Central Register of Clinical Trials to locate any trials published after 2005 (and therefore potentially not included in the above 2 meta-analyses) that had directly compared an SGA to a psychotherapy condition. We completed our search on November 26, 2009. Search terms included the word "depression" as well as the generic names of all SGAs (i.e., bupropion, citalopram, duloxetine, escitalopram, fluvoxamine, fluoxetine, milnacipran, mirtazapine, nefazodone, paroxetine, sertraline, and venlafaxine). Our supplementary search located 6 additional trials. Therefore, we located a total of 15 studies eligible for inclusion in the present meta-analyses, including 19 comparisons and 1975 participants.

To be included in our analysis, studies were required to meet the following inclusion criteria.

1. The study compared an SGA medication and psychotherapy.
2. Participants were diagnosed with major depressive disorder.
3. Authors included sufficient information to calculate effect sizes. One of our included studies provided no numerical data at study end point; however, a figure in the publication clearly denoted nearly no mean difference between groups (Rodríguez et al., 2004). To maximize sample size, the study was included with an effect size coded as zero. This study was small, and including it led to no meaningful changes in our results.
4. Participants were aged 18 or over.
5. The study did not augment SGA treatment with medication that was not an SGA (e.g., a study could not have added lithium, an MAOI, or a TCA to patients who did not respond to the initial course of medication).

## Data Extraction

The first and third authors (an associate professor of psychology and an advanced undergraduate student) independently coded all moderator variables and extracted efficacy and tolerability data for 13 studies, which were then checked for inconsistencies. The first author also coded 2 additional studies, which were independently coded by the second author, a psychologist. Prior to coding, several studies not included in the current analysis were coded to establish interrater consistency. All disagreements were resolved through consensus conference after rereading the study in question. For the moderator variables, interrater agreement was calculated via kappa coefficient for categorical variables and 2-way random effects intraclass correlation coefficient based on absolute agreement for continuous variables.

## Moderator Variables

### Bona Fide Psychotherapy

Each psychotherapy condition was coded to determine the extent to which it was bona fide, using criteria modified slightly from Wampold et al., (1997). To qualify as bona fide, a treatment must have met criteria I, II, and III.

(I) The therapist was trained to provide the therapy and held at least a master's degree or was enrolled in a graduate program in a relevant field, such as clinical or counseling psychology or social work.

(II) The therapist developed a relationship based on face-to-face meetings with the client and the treatment was individualized (i.e., the treatment did not merely deliver a rigidly standardized set of procedures, such as prerecorded relaxation instructions, to each client).

(III) Treatments contained psychologically valid components, as evidenced by at least 2 of the following 4 criteria: (*a*) a citation was made to an established school of or approach to psychotherapy; (*b*) a description of the therapy was given in the article and the description contained a reference to a psychological process (e.g., operant conditioning); (*c*) a manual for the treatment was used to guide the delivery of the treatment; and (*d*) the active ingredients of the treatment were identified and citations provided for those ingredients. Interrater agreement for categorizing treatments into bona fide or not yielded a kappa coefficient of 0.77.

### Treatment Quality Characteristics

Each trial was examined for several treatment quality characteristics. The potential moderating effect of each variable was only

examined within comparisons between bona fide psychotherapy and medication because this was the fairest comparison between treatments (Spielmans et al., 2010):

1. Medication dosage: Was medication dosed within the Food and Drug Administration (FDA)-approved range? This was determined by examination of prescribing information for each drug. Interrater agreement was 100%. No kappa coefficient could be calculated because there was no variance—all studies used dosages in the FDA-approved ranges.
2. Medication-psychotherapy dosage: Was at least one psychotherapy session given per every week that the drug group received treatment? $\kappa = 0.34$. Given the relatively low interrater agreement on this variable, extensive discussion and recoding of this variable occurred to ensure that coding was accurate.
3. Drug type: Was the drug an SSRI, SNRI, or other (bupropion, nefazodone)? $\kappa = 1.0$.
4. Dose escalation: Was either treatment allowed to use dose escalation or was dosage fixed? $\kappa = 0.87$ for medication dose escalation; $\kappa = 0.77$ for psychotherapy dose escalation.
5. Funding of trial: Was trial funded by the manufacturer of a medication used in the trial? $\kappa = 1.0$.
6. Switching: Was switching to another drug or psychotherapy allowed if treatment was not resulting in adequate change? Drug group allowed to switch $\kappa = 1.0$. Interrater agreement for whether switching was allowed in the psychotherapy group was 100%, but $\kappa$ could not be calculated because no trial allowed switching in the psychotherapy group.
7. Psychotherapy allegiance: Did one or more authors of the trial develop the psychotherapy used in the trial? $\kappa = 0.54$. Due to the relatively low interrater reliability, recoding and discussion of each study occurred to ensure accuracy.
8. Length of trial: How many weeks in duration was the trial? The intraclass correlation coefficient could not be calculated because interrater agreement was 100%.
9. Drug group visits: How many visits to the study medical staff were completed by participants in the medication group? Intraclass correlation coefficient = 0.99.

Table 1 shows the treatment quality characteristics of all included trials.

## Outcome Measures

Effect sizes were calculated for all depression measures. Measures such as Clinical Global Impressions and anxiety scales were excluded from analysis, as the focus of the current meta-analysis focused on depression specifically. Separate analyses were conducted for continuous measures and for categorical measures, including treatment response and remission, as defined in each study. Rates of treatment discontinuation were also extracted.

## Statistical Analysis

### Calculation of Effect Sizes

Outcomes across all relevant outcome variables were pooled within studies to provide one omnibus effect size for each study. Overall effect size of each study was weighted by its inverse variance to provide a pooled effect size estimate ($d+$) that most accurately represented the true population effect size (Hedges and Olkin, 1985). All effect sizes were calculated using Comprehensive Meta-Analysis software (Biostat, 2006).

In some studies, 3 or more treatment groups were evaluated. If 2 psychotherapies were compared with a medication in the same study, then both psychological treatments were computed versus the same medication group, creating a structural dependency, which may result in inaccurate analyses (Gleser and Olkin, 1994), though a much larger meta-analysis found that failing to correct for dependencies made little difference in the final results (Wampold et al., 1997). Pooling effects within studies reduces power to detect differences between psychotherapy and medication. Analyses were thus conducted in which dependencies were eliminated through pooling the average effect (i.e., pooling psychotherapy A versus medication and psychotherapy B versus medication into a single average effect size in the same study) and where the assumption of independence was violated through including both effects separately. Our results showed very little difference when dependent comparisons were included, so results are reported which include dependent comparisons, as this maximizes the ability to detect the impact of moderator variables. Effect sizes were calculated separately for study end point (immediately after treatment was concluded) and at the final follow-up visit.

### Homogeneity

Tests of homogeneity examine whether all studies within a particular classification are estimating the same effect. Thus, a test of homogeneity, using $Q$, was performed for each analysis. However, the $Q$ test of homogeneity often lacks power to detect heterogeneity when the number of studies in a meta-analysis is small (Hardy and Thompson, 1998; Hedges and Pigott, 2001). Therefore, in addition to using $Q$, we used $I^2$, which assesses the percentage of variation between studies due to heterogeneity as opposed to chance and is not dependent on the number of studies in a meta-analysis. Guidelines suggest that an $I^2$ value of greater than 50% indicates significant heterogeneity (Higgins et al., 2003).

### Statistical Model

To maximize the generalizability of our findings, random effects analyses were used (Hedges, 1994).

# RESULTS

## Study Characteristics

Fifteen studies that reported data on a total of 1014 patients receiving psychotherapy and 961 patients receiving pharmacotherapy were included. Of the 19 psychotherapy interventions compared with medication, 14 were cognitive and/or behavioral, whereas the remainder were psychodynamic, interpersonal, or supportive in nature. Eight trials explicitly excluded suicidal patients and 7 trials did not mention whether suicidal participants were included. One study displayed the potential confound of excluding prior nonresponders to the medication used in the trial (paroxetine) while failing to also exclude participants who had previously responded poorly to the psychotherapy method used in the trial (cognitive-behavioral therapy—DeRubeis et al., 2005). All trials used medication doses in the FDA-approved range. Trials generally included participants with a wide range of depressive symptom severity; only one study exclusively included participants with severe depression (Shamsaei et al., 2008).

### Efficacy and Tolerability

At trial end point, psychotherapy and pharmacotherapy were not significantly different in their impact on either continuous depression scales (Table 2) or on categorical measures, including response, remission, and study completion rates (Table 3). The omnibus effect across continuous outcome measures for each individual study and the overall result at study end point are graphically portrayed in Figure 1. Whether psychotherapy was bona fide was not a significant moderator of treatment efficacy (continuous outcomes) in comparison to medication when using a random effect model: $Q_B = 3.30$, $p = 0.07$. But when using a fixed effects model, the

**TABLE 1.** Selected Characteristics of Included Studies

| Study | Psychotherapy | Drug | Dose Escalation | Chronic Depression | Bona fide Psychotherapy[a] | Funding Source | Treatment Switching | Drug Group Visits | Psychotherapy Allegiance | Trial Duration (Follow-up) | N[b] | Effect Size (Follow-up) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| David et al., 2008 | Cognitive | Fluoxetine | Medication | ? | Both | Government | No | 14 | No | 14 (26) | 170 | 0.11 (0.39) |
| | Rational-emotive behavioral | | | | | | | | | | | 0.27 (0.64) |
| Dekker et al., 2008 | Psychodynamic | Venlafaxine | Medication | ? | No[c] | Drug sponsor | No | 4 | Yes | 8 | 101 | −0.35 |
| DeRubeis et al., 2005 | CT | Paroxetine | Both | ? | Yes | Government | No | 6 | Yes | 8[d] | 180 | −0.16 |
| Dimidjian et al., 2006 | BA | Paroxetine | Medication | ? | Both | Government | No | 10 | Yes | 16 | 188 | 0.00 |
| | CBT | | | | | | | | | | | −0.46 |
| Faramarzi et al., 2008 | CBT | Fluoxetine | No | ? | Yes | Government | No | ? | No | 12 | 59 | 0.50 |
| Keller et al., 2000 | CBASP | Nefazodone | Both | Yes | Yes | Drug sponsor | No | 12 | Yes | 12 | 436 | −0.06 |
| Konarski et al., 2009 | CBT | Venlafaxine | Medication | ? | Yes | Government and drug sponsor | No | ? | No | 16 | 24 | −1.34 |
| Martin et al., 2001 | IPT | Venlafaxine | No | ? | No[c] | Drug sponsor | No | 3 | No | 6 | 28 | −0.54 |
| Miranda et al., 2003 | CBT | Paroxetine | Both | ? | Yes | Government | Drug only | ? | Yes | 26 (26) | 178 | −0.36 (0.05) |
| Mohr et al., 2001 | CBT | Sertraline | No | ? | Both | Government; nonprofit foundation | No | 4 | No | 16 (26) | 55 | 0.22 (0.40) |
| | Supportive-expressive | | | | | | | | | | | −0.40 (0.08) |
| Mynors-Wallis et al., 2000 | GP problem-solving | Fluvoxamine or paroxetine | Medication | ? | Neither[c] | Government | No | 6 | Yes | 12 (40) | 104 | −0.20 (0.16) |
| | Nurse problem-solving | | | | | | | | | | | −0.20 (0.14) |
| Rodriguez et al., 2004 | Bellak's | Fluoxetine | No | ? | Yes | ? | No | 20 | No | 20 (18)[d] | 20 | 0.00 (0.00) |
| Salminen et al., 2008 | Psychodynamic | Fluoxetine | Medication | ? | Yes | Government; nonprofit foundation | No | 6 | No | 16 | 51 | −0.12 |
| Segal et al., 2006 | CBT | Fluvoxamine or paroxetine or venlafaxine | Medication | ? | Yes | Government | Drug only | 10–13 | No | 26 | 301 | 0.37 |
| Shamsaei et al., 2008 | CBT | Citalopram | Medication | No | No[c] | Government | No | ? | ? | 8 | 80 | −1.63 |

Positive effect sizes represent an advantage for psychotherapy.

[a]"Yes" means that the trial used bona fide psychotherapy; "no" means it did not. "Both" means 2 types of psychotherapies were tested in the trial and both were bona fide; "neither" means 2 types were tested and none of them were bona fide. Note that there were a total of 5 non-bona fide psychotherapies tested across 3 trials.

[b]Sample size refers to number of participants included in continuous efficacy analyses.

[c]The non-bona fide psychotherapies were not considered fully therapeutic because either (*a*) they were delivered by nurses or general practice physicians, whose academic training does not include the substantial theoretical and practical training typical of psychotherapists or (*b*) the qualifications of the therapists were unclear (Dekker et al., 2008).

[d]In this study, pharmacotherapy participants who did not respond at 8 wk were given adjunctive lithium or desipramine. Thus, we only examined data from the first 8 wk of the trial.

[e]In this study, treatment lasted for 6 mo; outcome was assessed weekly through week 20 (not week 26), then again at follow-up.

? indicates that relevant information to code this variable was not clearly provided in the study.

**TABLE 2.** Composite Effect Sizes and Homogeneity of Effects for Psychotherapy Versus Pharmacotherapy: Continuous Outcomes

| Comparison | *k* | *d* | *Z* | *p* | *Q* | *p* | $I^2$ |
|---|---|---|---|---|---|---|---|
| End point | | | | | | | |
| All psychotherapies | 19 | −0.19 | 1.89 | 0.06 | 82.94 | <0.001 | 78.30% |
| Non-bona fide therapies | 5 | −0.58 | 2.12 | 0.03 | 23.21 | <0.001 | 82.76% |
| Bona fide therapies | 14 | −0.05 | 0.56 | 0.57 | 41.03 | <0.001 | 68.32% |
| Follow-up | | | | | | | |
| All psychotherapies | 8 | 0.26 | 3.01 | 0.003 | 7.60 | 0.37 | 7.88% |
| Non-bona fide therapies | 2 | 0.15 | 0.80 | 0.42 | 0.00 | 0.95 | 0% |
| Bona fide therapies | 6 | 0.29 | 2.52 | 0.01 | 7.17 | 0.21 | 30.29% |

Positive values represent an advantage for psychotherapy.

**TABLE 3.** Composite Effect Sizes and Homogeneity of Effects for CBT and Other Therapies by Type of Disorder: Categorical Outcomes

| Comparison | *k* | OR | *Z* | *p* | *Q* | *p* | $I^2$ |
|---|---|---|---|---|---|---|---|
| Response: end point | | | | | | | |
| All psychotherapies | 12 | 1.12 | 0.91 | 0.36 | 14.33 | 0.22 | 23.21% |
| Non-bona fide therapies | 2 | 0.56 | 1.57 | 0.12 | 0.06 | 0.80 | 0% |
| Bona fide therapies | 10 | 1.21 | 1.59 | 0.11 | 10.31 | 0.33 | 12.71% |
| Remission: end point | | | | | | | |
| All psychotherapies | 13 | 1.07 | 0.41 | 0.68 | 23.06 | 0.03 | 47.96% |
| Non-bona fide therapies | 2 | 0.55 | 1.77 | 0.08 | 0.02 | 0.89 | 0% |
| Bona fide therapies | 11 | 1.18 | 1.04 | 0.30 | 18.22 | 0.05 | 45.11% |
| Study completion: end point | | | | | | | |
| All psychotherapies | 16 | 1.17 | 0.61 | 0.54 | 72.43 | <0.001 | 79.29% |
| Non-bona fide therapies | 3 | 0.55 | 2.00 | 0.05 | 0.86 | 0.65 | 0% |
| Bona fide therapies | 13 | 1.41 | 1.16 | 0.25 | 63.97 | <0.001 | 81.24% |
| Response: follow-up | | | | | | | |
| All psychotherapies | 4 | 0.99 | 0.04 | 0.97 | 0.98 | 0.81 | 0% |
| Non-bona fide therapies | 2 | 0.78 | 0.64 | 0.52 | 0.01 | 0.93 | 0% |
| Bona fide therapies | 2 | 1.16 | 0.46 | 0.65 | 0.35 | 0.55 | 0% |
| Remission: follow-up | | | | | | | |
| All psychotherapies | 5 | 1.30 | 1.50 | 0.14 | 0.59 | 0.96 | 0% |
| Non-bona fide therapies | 2 | 1.14 | 0.40 | 0.69 | 0.12 | 0.73 | 0% |
| Bona fide therapies | 3 | 1.38 | 1.52 | 0.13 | 0.24 | 0.89 | 0% |

ORs greater than one represent an advantage for psychotherapy.

extent to which psychotherapy was bona fide was a significant moderator: $Q_B = 18.70$, $p < 0.001$. At end point, medication was superior to non-bona fide psychotherapy on continuous outcome measures by a moderate effect size. At follow-up, bona fide psychotherapy was superior to medication by a small effect size. This effect was only marginally significant when dependent comparisons were eliminated, though the effect size was similar to when dependent comparisons were included: $d = .23$, $p = .10$. On response and remission, there were no significant differences between treatments. However, at end point, bona fide psychotherapies yielded significantly higher response rates relative to medication than did non-bona fide psychotherapies: $Q_B = 3.94$, $p = 0.05$. Bona fide psychotherapies showed higher remission rates relative to medication than did non-bona fide psychotherapies: $Q_B = 4.18$, $p = 0.04$. Study discontinuation rates were significantly higher among participants receiving non-bona fide psychotherapy than bona fide psychotherapy relative to medication: $Q_B = 5.02$, $p = 0.03$.

The differences between medication and psychotherapy were often heterogeneous, even when only considering comparisons between bona fide therapy and pharmacotherapy. However, apart from whether the therapy was bona fide, none of our treatment quality moderators explained the variance in effects, including author allegiance, study duration, number of drug visits, type of medication, dose escalation, medication-psychotherapy dosage, trial funding source, and if drug switching was allowed. A post hoc analysis was undertaken to examine if the type of bona fide psychotherapy (cognitive-behavioral or a variant thereof vs. non-CBT) was related to outcome. The difference between types of therapy was not significant: $Q_B = 0.47$, $p = 0.49$. At follow-up, bona fide CBT showed a significant advantage over pharmacotherapy ($k = 4$, $d = 0.35$, $p = 0.02$), while bona fide non-CBT did not outperform medication ($k = 2$, $d = 0.05$, $p = 0.85$); the effects of CBT versus medication showed signs of heterogeneity ($Q = 6.32$, $I^2 = 52.38\%$).

## DISCUSSION

Psychotherapy and pharmacotherapy yielded quite similar results in the short-term treatment of depressive symptoms. Tolerability of treatments, in terms of dropout rates, was comparable. In the short-term treatment of depression, there was much heterogeneity among the effect size differences between medication and psychotherapy that was largely unexplained by our treatment quality moderator variables. On continuous measures, whether psychotherapy was bona fide was a highly significant moderator using a fixed effects model, but it was only of borderline significance when using a random effects model. However, using a random effects model, non-bona fide psychotherapy was statistically significantly inferior to medication by a moderate effect size at treatment end point whereas bona fide psychotherapy was equivalent to medication. For both response and remission, whether psychotherapy was bona fide was a significant moderator. At follow-up, bona fide psychotherapy was superior to medication by a small margin.

However, our efforts to explain this heterogeneity using moderator variables uncovered important methodological problems in the evidence base that compares medications with psychotherapy. Of greatest methodological concern is continued comparative research using psychotherapies that are not bona fide, generally by virtue of lacking adequately trained practitioners. Three studies included in our sample used nurses or general practice physicians with minimal or unclear levels of training and experience in psychotherapy (Martin et al., 2001; Mynors-Wallis et al., 2000; Shamsaei et al., 2008). Another study offered little data about the qualifications of the therapists to provide the therapy (Dekker et al., 2008). (We attempted to contact the authors of 3 studies (Dekker et al., 2008; Konarski et al., 2009; Miranda et al., 2003) to gather further information regarding the qualifications and training of therapists used in their trials. We received additional information regarding the Konarski et al., (2009) and Miranda et al., (2003) studies but not the Dekker et al., (2008) study.) In contrast (and by necessity), all studies used qualified, licensed prescribers to provide medication conditions. Although the relationship of therapist train-

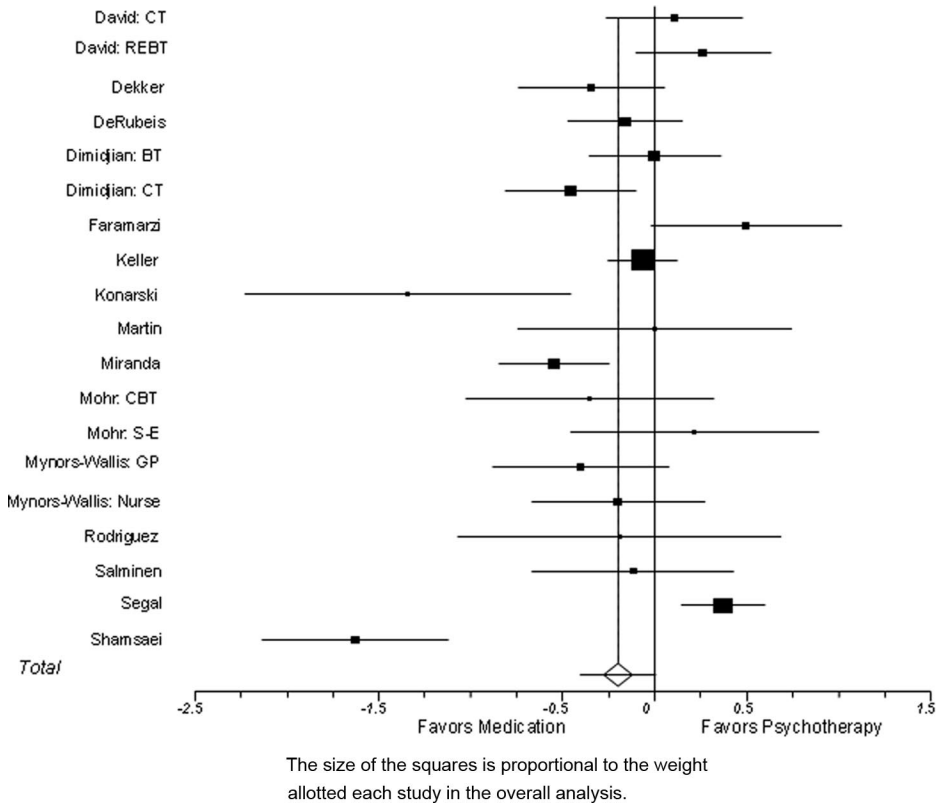


**FIGURE 1.** Forest plot of effect size differences and confidence intervals for individual trials and overall.

ing and experience to outcome is complex (e.g., Beutler, 1997), a fair comparison of psychotherapy to medication requires licensed and qualified providers of psychotherapy, or at least trainees in professional fields that provide psychotherapy, to deliver therapy in research.

A second related methodological concern affecting several included studies deals with the "dosages" of psychotherapy and the frequent face-to-face medication management received by participants in the medication groups. The provision of adequate dosages of SGAs constituted a methodological strength of the research analyzed here. All included studies prescribed antidepressant medications in FDA-approved ranges for efficacious treatment of depression; the majority also allowed increases in medication dosage based on clinical response. The "dose" provided of psychotherapy, on the other hand, was often strikingly small: Five studies included in this meta-analysis offered 10 or fewer sessions of psychotherapy, and 4 of these brief interventions were delivered by therapists with unclear or non-bona fide qualifications.

If the dose of psychotherapy provided participants in these studies was often strikingly small, the dose of face-to-face contact with medication prescribers was often much larger than in realistic clinical practice settings. In 3 studies, medication was actually offered for a longer duration than psychotherapy, a clear confound. In over half of our included studies, participants had visits with the medical staff during at least half of the weeks of the study, and in 2 studies, participants met weekly for medication management. Although we did not find that frequency of medication management visits moderated outcome, the lack of correspondence between this aspect of research design and usual treatment may limit the generalizability of these comparative trials to clinical practice.

Other common methodological problems made it difficult to adequately uncover the source of heterogeneity in our results. We hoped to test moderator variables related to participant characteristics as well as treatment quality (e.g., severity or chronicity of depression). However, researchers generally offered incomplete information such variables, making them impossible to reliably code. For example, few studies offered complete information about the severity of depression among their participants, making it difficult to establish if more severe patients benefit more from medication or psychotherapy. While the mean baseline level of depression was reported in each trial, the actual number of participants experiencing severe depression was rarely provided. Many trials excluded the most severely distressed persons: for example, most trials excluded individuals with Borderline Personality Disorder, and 8 trials excluded at least some types of suicidal patients, generally those judged to be at imminent risk of suicide and in need of hospitalization. However, researchers offered no information about how suicidality was assessed and what specific criteria were used to exclude patients. Trials also varied substantially in what information they elicited about their participants' prior history of depression and chronicity: only 4 trials included participants' mean number of prior major depressive episodes, whereas 6 included the mean duration of the current major depressive episodes. Researchers should make efforts to include the most severely distressed individuals with depression in comparative trials, as the ethical risk of their inclusion is counterbalanced by the need to identify the most effective treatment for them.

Researchers also should pay careful attention to potential confounds, such as nonrandom assignment, concurrent or prior use of psychiatric medications and/or psychotherapy among participants, and allowing changes based on nonresponse to one form of treatment but not the other. Eleven studies permitted dose increases to optimize response to antidepressant medication; only 3 of these allowed a similar increase in intensity for participants who appeared to need more psychotherapy. Similarly, 3 studies allowed clinicians to change the antidepressant medication given to participants if they

did not respond to the first medication; none of the studies permitted similar changes in psychotherapy, although some outcome studies in psychotherapy have included therapy-switching for nonresponders in their designs (e.g. Agras et al., 1995). Only 4 trials attempted to exclude participants who had failed a previous adequate trial of either the study medication or study psychotherapy. One trial displayed the potential confound of excluding prior nonresponders to one mode of treatment used in a trial (paroxetine) while failing to also exclude participants who had previously responded poorly to another form of treatment (cognitive therapy) used in a trial (DeRubeis et al., 2005). Two trials did not assign participants randomly, although in one trial this was due to one modality being a group intervention, and a blocking strategy was used to attempt to address the limitation created by forming groups out of consecutive referrals (Mohr et al., 2001).

Despite these methodological flaws, however, we were able to uncover some interesting findings with respect to the moderators that affect treatment outcome. Our finding that psychotherapy allegiance, for example, was not linked to better outcomes for psychotherapy is surprising, given that several meta-analyses have found that researcher allegiance has a notable and consistent relationship with outcome (Luborsky et al., 1999; Weisz et al., 2006). In comparative psychotherapy trials, where study therapists may be quite aware of which therapy is expected to show superior results, such expectations may lead them to deliver the favored therapy with a great deal of enthusiasm and to deliver the less favored therapy in a less enthusiastic manner. While researcher allegiance in psychotherapy-pharmacotherapy trials may lead to very competent delivery of psychotherapy, it seems less likely to result in poor delivery of pharmacotherapy interventions.

At follow-up, psychotherapy appears to offer a small advantage over medication, particularly when considering only bona fide therapies. However, our finding that bona fide psychotherapy appeared to perform better than non-bona fide therapy relative to medication must be considered in the context of a small sample size; only 4 trials used psychotherapy that was not clearly bona fide (Dekker et al., 2008; Martin et al., 2001; Mynors-Wallis et al., 2000; Shamsaei et al., 2008). In addition, trials varied in terms of how much contact they permitted with therapists and prescribers during the follow-up period. One study permitted an equivalent number of "booster sessions" of psychotherapy and medication management, while permitting participants to continue their medication (David et al., 2008). Some studies discontinued treatment in both therapy and medication conditions (Miranda et al., 2003; Mohr et al., 2001). One study permitted continued medication use during follow-up, but not therapy (Mynors-Wallis et al., 2000). The extent to which participants sought treatment outside the treatment protocol after the initial phase was not mentioned in any study except for Mohr et al., 2001. In sum, while psychotherapy outcomes were superior to pharmacotherapy outcomes at follow-up, the widely divergent designs of the included studies make it difficult to determine under which circumstances either psychotherapy or pharmacotherapy might exert maximal relative efficacy in the long-term.

The most notable remaining limitation of this meta-analysis is a small sample size, which limits the ability to detect differences between types of treatment. Despite this limitation, we were able to identify some important differences. It appears that bona fide psychotherapy is as effective as medication in the short-term treatment of depression. Psychotherapy may also offer an advantage over SGAs in the longer-term management of depressive symptoms, though this finding should be considered tentative in light of the small sample of trials addressing this important topic.

## ACKNOWLEDGMENT

*The authors thank Maria Isabel Jimenez for her timely assistance in translating a study for use in the present meta-analysis (Rodríguez et al., 2004).*

## REFERENCES

Agras WS, Telch CF, Arnow B, Eldredge K, Detzer MJ, Henderson J, Marnell M (1995) Does interpersonal therapy help patients with binge eating disorder who fail to respond to cognitive-behavioral therapy? *J Consult Clin Psychol*. 63:356–360.

Anderson IM (2000) Selective serotonin reuptake inhibitors versus tricyclic antidepressants: A meta-analysis of efficacy and tolerability. *J Affect Disord*. 58:19–36.

Anderson IM, Tomenson BM (1995) Treatment discontinuation with selective serotonin reuptake inhibitors compared with tricyclic antidepressants: A meta-analysis. *BMJ*. 310:1433–1438.

Beutler LE (1997) The psychotherapist as a neglected variable in psychotherapy: An illustration by reference to the role of therapist experience and training. *Clin Psychol Sci Pract*. 4:44–52.

Biostat (2006) *Comprehensive Meta-Analysis* [Computer software]. Version 2.2.030.

Cherry DK, Hing E, Woodwell DA, Rechsteiner EA (2008) National Ambulatory Medical Care Survey: 2006 Summary. Hyatsville (MD): National Center for Health Statistics.

Cuijpers P, van Straten A, van Oppen P, Andersson G (2008) Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? A meta-analysis of comparative studies. *J Clin Psychiatry*. 69:1675–1685.

Cuijpers P, van Straten A, Warmerdam L, Andersson G (2008) Psychological treatment of depression: A meta-analytic database of randomized studies. *BMC Psychiatry*. 8:36.

David D, Szentagotai A, Lupu V, Cosman D (2008) Rational emotive behavior therapy, cognitive therapy, and medication in the treatment of major depressive disorder: A randomized clinical trial, posttreatment outcomes, and six-month follow-up. *J Clin Psychol*. 64:728–746.

De Maat S, Dekker J, Schoevers R, De Jonghe F (2006) Relative efficacy of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Psychother Res*. 16:566–578.

Dekker JJ, Koelen JA, Van HL, Schoevers RA, Peen J, Hendriksen 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. *J Affect Disord*. 109:183–188.

DeRubeis RJ, Gelfand LA, Zang TZ, Simons AD (1999) Medications versus cognitive behavior therapy for severely depressed outpatients: Mega-analysis of four randomized comparisons. *Am J Psychiatry*. 156:1007–1013.

DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, O'Reardon JP, Lovett ML, Gladis MM, Brown LL, Gallop R (2005) Cognitive therapy versus medications in the treatment of moderate to severe depression. *Arch Gen Psychiatry*. 62:409–416.

Dimidjian S, Hollon SD, Dobson KS, Schmaling KB, Kohlenberg RJ, Addis ME, Gallop R, McGlinchey JB, Markley DK, Gollan JK, Atkins DC, Dunner DL, Jacobson NS (2006) Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *J Consult Clin Psychol*. 74:658–670.

Faramarzi M, Alipor A, Esmaelzadeh S, Kheirkhah F, Poladi K, Pash H (2008) Treatment of depression and anxiety in infertile women: Cognitive behavioral therapy versus fluoxetine. *J Affect Disord*. 108:159–164.

Gaffan EA, Tsaousis I, Kemp-Wheeler SM (1995) Researcher allegiance and meta-analysis: The case of cognitive therapy for depression. *J Consult Clin Psychol*. 63:966–980.

Gleser LJ, Olkin I (1994) Stochastically dependent effect sizes. In Cooper H, Hedges LV The Handbook of Research Synthesis (pp 339–356). New York (NY): Russell Sage.

Gloaguen V, Cottraux J, Cucherat M, Blackburn I (1998) A meta-analysis of the effects of cognitive therapy in depressed patients. *J Affect Disord*. 49:59–72.

Hardy RJ, Thompson SG (1998) Detecting and describing heterogeneity in meta-analysis. *Stat Med*. 17:841–856.

Hedges LV (1994) Statistical considerations. In Cooper H, Hedges LV The Handbook of Research Synthesis (pp 29–38). New York (NY): Russell Sage.

Hedges LV, Olkin I (1985) . Statistical Methods for Meta-Analysis. San Diego (CA): Academic Press.

Hedges LV, Pigott TD (2001) The power of statistical tests in meta-analysis. *Psycholol Methods*. 6:203–217.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *BMJ*. 327:557–560.

Imel ZE, Malterer MB, McKay KM, Wampold BE (2008) A meta-analysis of psychotherapy and medication in unipolar depression and dysthymia. *J Affect Disord*. 110:197–206.

Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ, Markowitz JC, Nemeroff CB, Russell JM, Thase ME, Trivedi MH, Zajecka J (2000) A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *New Engl J Med*. 342:1462–1470.

Klein DF (2000) Flawed meta-analyses comparing psychotherapy with pharmacotherapy. *Am J Psychiatry*. 157:1204–1211.

Konarski JZ, Kennedy SH, Segal ZV, Lau MA, Bieling PJ, McIntyre RS, Mayberg HS (2009) Predictors of nonresponse to cognitive behavioural therapy or venlafaxine using glucose metabolism in major depressive disorder. *J Psychiatry Neurosci*. 34:175–180.

Luborsky L, Diguer L, Seligman DA, Rosenthal R, Krause ED, Johnson S, Halperin G, Bishop M, Berman JS, Schweizer E (1999) The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clin Psychol Sci Pract*. 6:95–106.

Marcus SC, Olfson M (2010) National trends in the treatment for depression from 1998 to 2007. *Arch Gen Psychiatry*. 67:1265–1273.

Martin SD, Martin E, Rai SS, Richardson MA, Royall R (2001) Brain blood flow changes in depressed patients treated with interpersonal psychotherapy or venlafaxine hydrochloride: Preliminary findings. *Arch Gen Psychiatry*. 58:641–648.

Miranda J, Chung JY, Green BL, Krupnick J, Siddique J, Revicki DA, Belin T (2003) Treating depression in predominantly low-income young minority women: A randomized controlled trial. *JAMA*. 290:57–65.

Mohr DC, Boudewyn AC, Goodkin DE, Bostrom A, Epstein L (2001) Comparative outcomes for individual cognitive-behavior therapy, supportive-expressive group psychotherapy, and sertraline for the treatment of depression in multiple sclerosis. *J Consult Clin Psychol*. 69:942–949.

Mojtabi R, Olfson M (2008) National patterns in antidepressant treatment by psychiatrists and general medical providers: results from the national comorbidity survey replication. *J Clin Psychiatry*. 69:1064–1074.

Murray CJ, Lopez AD (1996) Evidence-based health policy–lessons from the global burden of disease study. *Science*. 274:740–743.

Mynors-Wallis LM, Gath DH, Day A, Baker F (2000) Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *BMJ*. 320:26–30.

Olfson M, Marcus SC (2009) National patterns in antidepressant medication treatment. *Arch Gen Psychiatry*. 66:848–856.

Olfson M, Marcus SC, Druss B, Elinson L, Tanielian T, Pincus HA (2002) National trends in the outpatient treatment of depression. *JAMA*. 287:203–209.

Raofi S, Schappert SM (2006) . Medication Therapy in Ambulatory Medical Care: United States, 2003–04. Hyatsville (MD): National Center for Health Statistics.

Robinson LA, Berman JS, Neimeyer RA (1990) Psychotherapy for the treatment of depression: A comprehensive review of controlled outcome research. *Psychol Bull*. 108:30–49.

Rodríguez JL, Butrón MAL, Terrez BEV, Salcedo VV (2004) Estudio doble ciego con antidepresivo, psicoterapia reve y placebo en pacientes con depresión leve a moderada. *Salud Mental*. 27:53–61.

Salminen JK, 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. *Psychother Psychosom*. 77:351–357.

Segal ZV, Kennedy S, Gemar M, Hood K, Pedersen R, Buis T (2006) Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Arch Gen Psychiatry*. 63:749–755.

Shadish WR, Sweeney RB (1991) Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *J Consult Clin Psychol*. 59:883–893.

Shamsaei F, Rahimi A, Zarabian MK, Sedehi M (2008) Efficacy of pharmacotherapy and cognitive therapy, alone and in combination in major depressive disorder. *Hong Kong J Psychiatry*. 18:76–80.

Spielmans GI, Gatlin ET, McFall JP (2010) The efficacy of evidence-based psychotherapies versus usual care for youth: Controlling confounds in a meta-reanalysis. *Psychother Res*. 20:234–246.

Spielmans GI, Pasek LF, McFall JP (2007) What are the active ingredients in cognitive and behavioral psychotherapy for anxious and depressed children? A meta-analytic review. *Clin Psychol Rev*. 27:642–654.

Wampold BE, Minami T, Baskin TW, Callen Tierney S (2002) A meta-(re)analysis of the effects of cognitive therapy versus "other therapies" for depression. *J Affect Disord*. 68:159–165.

Wampold BE, Mondin GW, Moody M, Stich F, Benson K, Ahn H (1997) A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially, "all must have prizes." *Psychol Bull.* 122:203–215.

Weisz JR, Jensen-Doss A, Hawley KM (2006) Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *Am Psychol*. 61:671–689.

Zuvekas SH (2005) Prescription drugs and the changing patterns of treatment for mental disorders, 1996–2001. *Health Affairs*. 24:195–205.