VU VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

Short-term psychotherapy for depression

Driessen, E.

2013

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Driessen, E. (2013). *Short-term psychotherapy for depression: Broadening the field of efficacy research*. GVO drukkers & vormgevers BV- Ponsen & Looijen.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 26. Oct. 2021

3

Publication bias inflates the apparent efficacy of psychological treatment for major depressive disorder: a systematic review and meta-analysis of US National Institute of Health-funded trials

This chapter is submitted as:
Driessen, E., Hollon, S. D., Bockting, C., Cuijpers, P., & Turner, E. H. (2012). Publication bias inflates the apparent efficacy of psychological treatment for major depressive disorder: A systematic review and meta-analysis of National Institute of Health-funded trials. *Manuscript submitted for publication.*

Abstract

Background: It has been shown empirically that the efficacy of antidepressant medication is overestimated due to publication bias but this has only been inferred with regard to psychological treatment for depression. We aimed to directly assess the extent of study publication bias in trials examining the efficacy of psychological treatment for major depressive disorder.

Methods and Findings: We identified US National Institute of Health grants intended to fund randomized controlled trials comparing psychological treatment to controls or other treatments in patients diagnosed with major depressive disorder for the period 1972-2000 and determined whether those grants led to publications. For studies that were not published, data were requested from investigators and included in the meta-analyses. The proportion of grants with unpublished findings was compared with the reported proportion of unpublished antidepressant studies. Seven of the 36 funded grants that began trials did not result in publications and two others never started. Effect sizes were significantly lower (p=0.04) in unpublished (g=0.16; -0.12-0.43) than in published (g=0.50; 0.35-0.64) comparisons to control conditions, resulting in a 14% decrease in effect size when the two were pooled (g=0.43; 0.28-0.57). The proportion of unpublished psychological treatment trials (7/36=19%) did not differ from antidepressant medications (23/74=31%) (p=0.26). Our findings may overestimate the "true" effect of psychological treatment for depression as outcome reporting bias could not be examined quantitatively.

Conclusion: The efficacy of psychological interventions for depression has been overestimated in the published literature, just as it has been for pharmacotherapy. Both are efficacious, but not to the extent that the published literature would suggest. Funding agencies or journals should archive both original protocols and raw data from any treatment trial to facilitate the assessment of outcome reporting bias. Clinicians, guidelines developers, and decision makers should be aware of overestimated effects of the predominant treatments for depression.

Introduction

Publication bias has been defined as the tendency for authors to submit or journals to accept manuscripts for publication based on the direction or strength of the study's findings (Dickersin, 1990). It has been long recognized that publication bias can lead to an overestimation of (treatment) effects (Rosenthal, 1979) and, therefore, can represent a serious threat to the validity of evidence-based decisions.

Major depression is a highly prevalent and disabling disorder associated with major personal and societal costs (Kessler, 2012; Kessler et al., 2003; Kupfer, Frank, & Phillips, 2012). It is the fourth leading cause of disease burden worldwide and it is expected to rank first in high-income countries by the year 2030 (Mathers & Loncar, 2006). Depression responds to medication and most depressed patients are now treated in primary care (Marcus & Olfson, 2010). However, the efficacy of pharmacotherapy for depression has been overestimated due to publication bias. Turner, Matthews, Linardatos, Tell, and Rosenthal (2008) found that publication was strongly linked to outcome in an inception cohort (a cohort of research that is followed from the beginning [Song et al., 2009]) of 74 placebo-controlled antidepressant studies submitted to the US Food and Drug Administration. All but one of the 38 studies viewed by the Food and Drug Administration as having positive results were published in a way that agreed with the Food and Drug Administration. By contrast, all but 3 of the 36 studies viewed by the Food and Drug Administration as having negative or questionable results were subjected to publication bias in one of two forms: (1) the studies were not published (61%), referred to as study publication bias, or (2) negative results were published as if they were positive (31%), also known as outcome reporting bias. Selective reporting of sites and subjects (especially dropouts) were among the methods employed to produce outcome reporting bias. Comparing the published literature with the original Food and Drug Administration data resulted in a 25% reduction in mean effect size for pharmacotherapy versus pill-placebo from 0.41 (0.36-0.45) to 0.31 (0.27-0.35). These findings, along with indications that antidepressants only separate from placebo among patients with more severe depressions (Fournier et al., 2010), have led some to conclude that psychotherapy should be preferred over medications in depression (Hollon, 2011).

The specific efficacy of psychotherapy might also be limited to more severe depressions (Driessen, Cuijpers, Hollon, & Dekker, 2010) and the question can be raised as to whether the effects of psychological treatments for depression might be overestimated due to publication bias too. Cuijpers, Smit, Bohlmeijer, Hollon, and Andersson (2010) examined 117 published trials including 175 comparisons between psychological treatments and control conditions and found strong indications of publication bias. Using Duval and Tweedie's (2000a; 2000b) trim and fill procedure, they found that the effects for 51 "missing" studies had to be imputed to adjust for publication bias, reducing the overall effect size for psychological treatment relative to controls by 37% from 0.67 (0.60-0.75) to 0.42 (0.33-0.51). This suggests that psychological treatment may not be as efficacious as the published literature would indicate, much as was found for medication treatment.

However, this conclusion rests on the assumption that the association between sample size and effect size in a meta-analysis, which produces an asymmetric funnel plot, necessarily stems from a failure to publish small studies with small effect sizes. However, small studies may disproportionately show large effects for reasons other than publication bias (Borenstein, Hedges, Higgins, & Rothstein, 2009; Sterne, Egger, & Smith, 2001). Therefore, Borenstein et al. (2009) state: "the only true test for publication bias is to compare effects in the published studies formally with effects in the unpublished studies" (p. 280).

We aimed to do such a "true" test for study publication bias in an inception cohort (Song et al., 2009) of trials based on US National Institute of Health grant approval, examining depressive symptom reduction of psychological treatments compared to control or alternative interventions in randomized clinical trials for adult or geriatric patients with major depressive disorder. We ascertained the frequency with which conducted studies were not published and sought a better estimate of psychological treatment's effect on major depressive disorder by pooling unpublished and published findings.

Methods

Defining the set of relevant funded grants

We searched US National Institute of Health databases for grants funding psychological treatment randomized clinical trials on major depressive disorder for the years 1972-2000. US National Institute of Health records do not go back before 1972 and we stopped after the year 2000 in order to allow adequate time for completion of the trial and publication of its results. When we started our study in 2010, we conservatively estimated up to seven years for the grant period plus three additional years for the trial to be written up and published. We used http://crisp.cit.nih.gov/crisp_query.generate_screen to search for relevant grants through 1984 and switched to http://projectreporter.nih.gov/reporter.cfm from 1985 on. These databases provide summary information on the grants and often include abstracts.

We searched using all possible combinations of the two following sets of terms for "*depression*" (depression, depressive, major depressive disorder, mood disorder, affective disorder, melancholic, melancholia) and "*psychological treatment*" (cognitive therapy, behavior therapy, behavioral therapy, interpersonal therapy, psychodynamic therapy, dynamic therapy, humanistic therapy, supportive therapy, experiential therapy, [self-]control therapy, [problem-]solving therapy, [supportive-]expressive therapy, family therapy, group therapy, marital therapy, couples therapy, aversive therapy, exposure therapy, psychotherapy, psychotherapies, psychotherapeutic, counseling, disease management, psychoanalytic, behavioral activation, cognitive behavioral analysis system, desensitization, relaxation techniques, and progressive muscle relaxation). We deleted duplicate listings and multiple years for the same grants.

We then screened the grant titles and available abstracts for possible inclusion, following the methods used by Turner et al. (2008) as closely as possible in order to facilitate comparison of study publication bias with the pharmacological literature. We included all grants that proposed to conduct 1) a randomized clinical trial examining 2) psychological treatment for 3) acute depression in 4) adults. Following Cuijpers et al. (2010), we defined psychological treatment as any intervention in which verbal communication between therapist and client was the core element or bibliotherapy was supplemented with personal support from a therapist. We only included studies in which the differential efficacy of psychological treatment could be ascertained, excluding studies in which all patients received the same psychological treatment or not all patients in a condition received psychological treatment.

Depression was defined as meeting diagnostic criteria for major depressive disorder or its equivalent in earlier nosologies (e.g., Research Diagnostic Criteria [Spitzer, Endicott, & Robins, 1978] or Feighner criteria [Feighner et al., 1972]). We included studies with mixed diagnoses only if outcomes were reported separately for major depressive disorder patients. We excluded studies on children and adolescents younger than 18, but included studies with geriatric patients since more recent studies do and psychological treatment efficacy does not differ across adulthood as a function of age (Cuijpers, van Straten, Smit, & Andersson, 2009).

Matching grants to published articles

We consulted an existing database of 243 randomized clinical trials (accessed in 2010) regarding psychological treatment for adult depression to identify published articles (Cuijpers, van Straten, Warmerdam, & Andersson, 2008). This continuously updated database was developed through a comprehensive literature search and is publicly available (www.evidencebasedpsychotherapies.org). We considered this database as defining the published literature on psychological treatment randomized clinical trials for depression.

We used this database to find published articles that could be matched to the grants based on the grant number acknowledged in the paper, checking to be sure that they matched with respect to investigator(s), type of psychological treatment, and comparator. When we could not match a grant with a published article, we searched PubMed for articles published by the principal investigator. Failing that, we contacted the investigator to determine whether findings ever had been published.

We focused on those articles that reported acute outcome in terms of measures of depression. When multiple sequential articles were found reporting acute response on the same measures, we selected the most complete in terms of sample size. We excluded articles that reported secondary analyses when another article reported the primary data.

We examined the completeness of our grant search strategy by checking all the published studies in the database just described for grant acknowledgements. Whenever we found an article that acknowledged a grant that was not in our list, we contacted the investigator to resolve the discrepancy. All published articles that matched to the US National Institute of Health grants were read independently by two

of the authors to ensure that they met inclusion criteria. Disagreements were resolved by consensus.

Assessment of study publication bias

We first calculated: 1) the proportion of funded grants not resulting in publication, and 2) the proportion of funded grants resulting in publication but not reporting the data requisite for meta-analysis. These proportions provide a direct assessment of the extent of study publication bias. In case of non-publication, we contacted the investigator to inquire why the data were not published and to request the unpublished data.

One author (ED) extracted data from published articles. We calculated post-treatment effect sizes for four different types of comparisons: 1) psychological treatment versus control conditions, 2) psychological treatment versus another psychological treatment, 3) psychological treatment versus antidepressant medication, and 4) psychological treatment combined with antidepressant medication versus medication mono-treatment. We distinguished no-treatment controls from controls in which some kind of treatment was provided and subdivided the latter into a) treatment-as-usual, b) pill-placebo, and c) psychological-placebo conditions intended to control for time and contact (nonspecific controls).

Within each study, comparative effect sizes were calculated by subtracting mean post-treatment score for one condition from the other and dividing the result by the pooled standard deviation, so that positive signs indicated superiority of psychological treatment. If means and standard deviations were not reported, we used other data (e.g., reported effect sizes, t-value, or event rates; in that order) to compute the effect size. When no such data were available, the study was considered as not reporting the requisite data for meta-analysis. We calculated effect sizes as Hedges' g, because it provides a less biased estimate than Cohen's d when sample sizes are small. All meta-analyses were conducted using random effects models in Comprehensive Meta-analysis (version 2.2.046; Borenstein, Hedges, Higgins, & Rothstein, 2005).

We used depressive symptoms as the sole outcome measure for this meta-analysis, most commonly measured by the Hamilton Depression Rating Scale (Hamilton, 1960) and the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). When results were reported separately for multiple instruments, different subgroups, or multiple comparisons within a given study, we computed combined effects according to the procedures implemented in Comprehensive Meta-analysis. Heterogeneity was examined using the Q-value and I^2-statistic, the latter including a 95% confidence interval as recommended by Ioannidis Patsopoulos, and Evengelou (2007). I^2's 95% confidence interval was calculated by means of the heterogi module (Orsini, Higgins, Bottai, & Buchan, 2005) within Stata (Stata Statistical Software: Release 11), using the non-central chi-squared-based approach (StataCorp, 2009).

We calculated separate pooled mean effect sizes for published and unpublished studies for each comparison. We then compared the two by means of subgroup analyses as implemented in Comprehensive Meta-analysis, using a fully random effects analysis and pooling study-to-study variance across subgroups, as

recommended when subgroups include small numbers of studies (Borenstein et al., 2009). We next calculated the overall pooled mean effect size by combining published and unpublished studies and finally calculated the change between this overall pooled mean effect size point estimate and the pooled mean effect size point estimate for published studies.

Comparisons between different psychological treatments have no natural order and the strategies used in earlier quantitative reviews either overestimate pooled mean effect size (making all signs positive) or artificially create a pooled mean effect size of zero (assigning signs on a random basis; Wampold et al., 1997). For studies including multiple psychological treatment conditions, we therefore ranked the different psychological treatment conditions in terms of the investigators' presumed expected efficacy based on explicit statements to that effect when indicated and, when not indicated, their order of presentation in the title or method section (in this order). In the case of dismantling studies that isolated separate components of a larger psychological treatment package, conditions including all components were ranked highest in terms of expected efficacy, with lower ranks for conditions with fewer components included. Two authors judged presumed expected efficacy independently and disagreements were resolved by group consensus. When no consensus could be reached, sensitivity analyses were conducted for each of the different rankings. Comparative effect sizes for different psychological treatments were then calculated in such a way that positive signs indicated superiority of the psychological treatment that we presumed the investigators expected to be more efficacious, while negative signs indicated superiority of the psychological treatment that we presumed the investigators expected to be less efficacious.

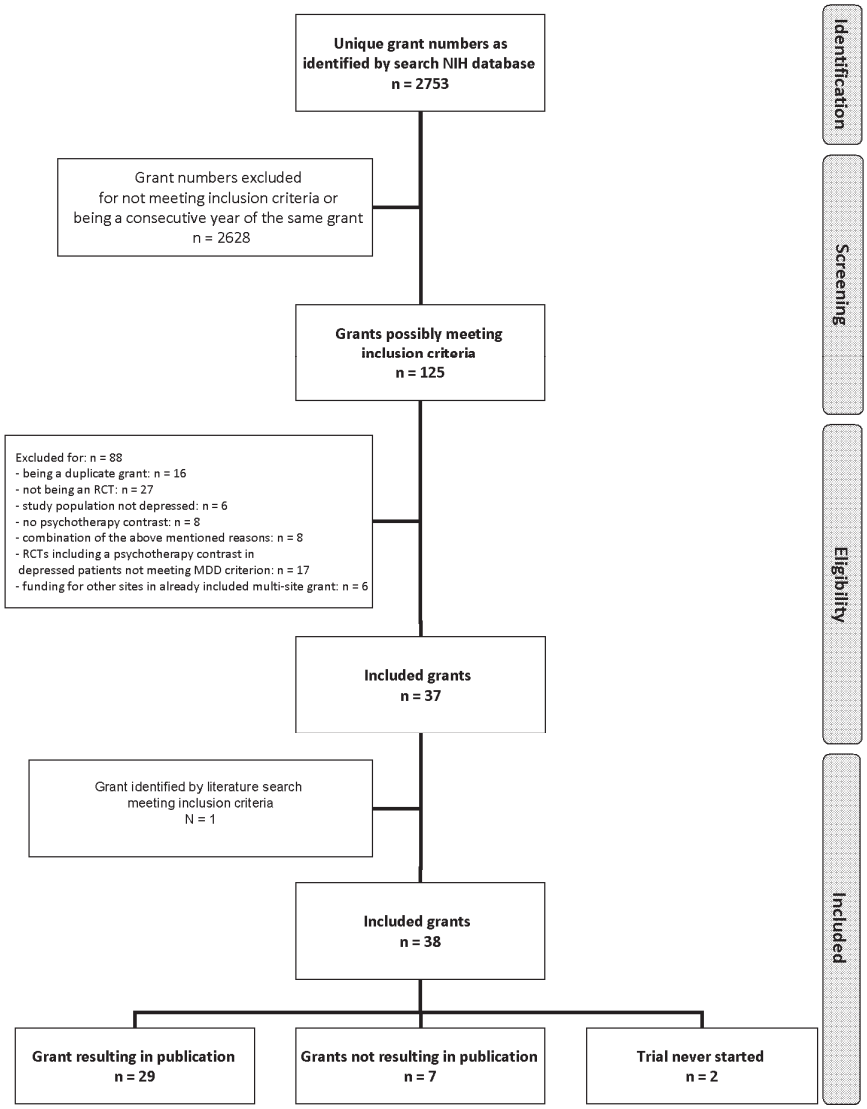
Results

As shown in Figure 1, 2753 grants were identified, of which 2628 were excluded based on titles and abstracts as not meeting inclusion criteria or as consecutive years of the same grant. Of the remaining 125 grants, 88 were excluded after a review of the published article or contacting the investigator, most often because the study was not an randomized clinical trial (*n*=27). This left 37 grants that met our criteria.

We examined the completeness of our grant search strategy by checking all the published studies in the above-mentioned database for US National Institute of Health grant acknowledgements, resulting in another 26 articles that acknowledged support from US National Institute of Health grants not discovered in our search. Only one met our inclusion criteria. In twelve of these articles the study participants were not selected on the basis of major depressive disorder diagnosis. In another five studies all patients received psychological treatment. Three of the publications acknowledged US National Institute of Health support that was not directly used to fund the study. One study was funded by a grant from the military and four studies were funded by US National Institute of Health grants awarded after our cut-off year of 2000. Finally, one article reported the results of an acute treatment trial (Reynolds, Miller et al., 1999)

not described in the abstract of an US National Institute of Health grant intended to fund a maintenance trial. Given that the acute trial met our inclusion criteria, we included the grant in our dataset, bringing the total number to 38.

Figure 1: PRISMA flow chart



Note: RCT = randomized controlled trial

Extent of study publication bias

The 38 included grants are described in Table 1. Multiple grants that funded the same study are counted as a single grant, as are two grants that generated multiple published randomized clinical trials. Two grant-funded studies were never started, one because of difficulty recruiting suitable patients (S. Stockard, personal communication, March 3, 2011) and the other because of difficulty finding psychodynamic therapists willing to participate in clinical research with Hispanic elders (J. Szapocznik, personal communication, August 30, 2010). These last two grants were excluded from further consideration.

Of the 36 grants that started studies, we were able to locate published articles corresponding to 29 (81%), but not for 7 (19%). These seven studies met our definition of unpublished studies. We were able to obtain the original data from six of these studies. The investigator for the seventh reported that the sample was small (no more than a dozen patients per condition) and the differences negligible (Gottlieb, personal communication, June 10, 2012), so we estimated the effect size as *g*=0.00 and *n*=10 per condition and conducted sensitivity analyses excluding this study from the comparisons in which it was included (psychological treatment versus antidepressant medication and combined treatment versus antidepressant monotreatment). All published studies reported sufficient outcome data to calculate effect sizes so that none were lost to analysis.

Adjusting for study publication bias

The results of the meta-analyses are presented in Table 2.

Psychological treatment versus control conditions

The pooled mean effect size of psychological treatment relative to control conditions was significantly lower (*p*=0.03) in the unpublished studies (*g*=0.16; -0.12-0.43, *n*=4) than in the published studies (*g*=0.50; 0.35-0.64; *n*=14). Adding unpublished studies to published studies resulted in a 14% decrease in effect size point estimate to *g*=0.43 (0.28-0.57). When psychological treatment was compared to a no-treatment control condition, lower pooled mean effect sizes also were found in the one unpublished study (*g*=0.50; -0.57-1.56) than in the three published studies (*g*=0.86; 0.37-1.36), but these differences were not significant (*p*=0.54). Comparing psychological treatment to treatment control conditions also resulted in significantly lower (*p*=0.04) pooled mean effect sizes for the unpublished studies (*g*=0.12; -0.12-0.37) than for the published studies (*g*=0.41; 0.28-0.54). Adding unpublished to published studies resulted in a 15% decrease in effect size point estimate to *g*=0.35 (0.23-0.46). Differentiating among the treatment controls resulted in a nonsignificant trend for comparisons to pill-placebo and nonsignificant differences for treatment-as-usual and psychological controls.

Psychological treatment versus antidepressant medication

When psychological treatment was compared to antidepressants, a lower pooled mean effect size was found in the unpublished studies (*g*=-0.27; -0.70-0.15; *n*=3) than in the published studies (*g*=0.05; -0.16-0.17; *n*=13), but this difference was not

Table 1: Characteristics of the included grants

	PI	Grant number	Reference	Study comparison(s)	PT type
Published					
1	Beck	2R01MH019989	Rush et al., 1977	PT vs ADM	CT
	Rush	1R03MH027759			
2	Beutler	2R01MH039859	Beutler et al., 1991	PT vs CTRL-NS; PT vs PT	CT; FEP
3	Covi	2R01MH033585	Covi & Lipman, 1987	PT vs PT	CBT; IPP
4	DeRubeis	1R10MH055877[a]	DeRubeis et al., 2005	PT vs ADM; PT vs CTRL-PLAC	CT
	Hollon	1R10MH055875			
	Hollon	1K02MH001697			
5	DiMascio	2R01MH026467	DiMascio et al., 1979	PT vs ADM; PT vs CTRL-NS; PT+ADM vs ADM	IPT
	Weissman	1R01MH026466	Weissman et al., 1979		
6	Freedland	1R21MH052629	Lustman et al., 1998	PT+CTRL-NS vs CTRL-NS	CBT
7	Glick	1R01MH034466	Clarkin et al., 1990; Glick et al., 1993	PT+CTRL-TAU+ADM vs CTRL-TAU+ADM	IFI
8	Greenberg	1R01MH045040	Goldman et al., 2006	PT vs PT	CCT; EFT
9	Hersen	5R01MH028279	Hersen et al., 1984	PT+ADM vs ADM; PT+CTRL-PLAC vs PT+CTRL-PLAC	SST; STPP
10	Hollon	1R01MH033209	Hollon et al., 1992	PT vs ADM; PT+ADM vs ADM	CT
11	Imber	2U01MH033753	Elkin et al., 1989	PT vs ADM; PT vs CTRL-PLAC; PT vs PT	CBT; IPT
	Watkins	1U01MH033760			
	Sotsky	2U01MH033762			
12	Jacobson	2R01MH033838	Jacobson et al., 1991	PT vs PT	CBT; BMT; CO
13	Jacobson	1R01MH055502	Dimidjian et al., 2006	PT vs ADM; PT vs CTRL-PLAC; PT vs PT	CT; BA
14	Jacobson	1R37MH044063	Jacobson et al., 1996	PT+PT+PT vs PT+PT; PT+PT+PT vs PT; PT+PT vs PT	CT; P-CT; BA
15	Jarrett	5R01MH045043	Jarrett et al., 1999	PT vs ADM; PT vs CTRL-PLAC	CT
16	Lustman	2R01DK036452	Hayden et al., 2012	PT vs CTRL-NS	
17	Lynch	1R03MH057799	Lynch et al., 2003	PT+ADM vs ADM	DBT
18	Miller	1R01MH035945	Miller et al., 1989	CTRL-TAU+ADM+PT vs CTRL-TAU+ADM; CTRL-TAU+ADM+PT vs CTRL-TAU+ADM+PT	CT; SST
19	Miranda	1R01MH056864	Miranda et al., 2003	PT vs ADM; PT vs CTRL-TAU	CBT
20	Mohr	1R01MH059708	Mohr et al., 2001	PT vs ADM; PT vs PT	CBT
21	Murphy	1R01MH032756	Murphy et al., 1984	PT vs ADM; PT+ADM vs ADM	CT
			Murphy et al., 1995	PT vs ADM; PT vs CTRL-NS	CBT

22	O'Hara	1R01MH050524	O'Hara et al., 2000	PT vs CTRL-NT	IPT
23	Reynolds	2/5R01MH37869	Reynolds, Miller et al., 1999	PT+ADM vs ADM; PT+CTRL-PLAC vs CTRL-PLAC	IPT
24	Schulberg	1R01MH045815[b]	Schulberg et al., 1996	PT vs ADM; PT vs CTRL-TAU	IPT
25	Spinelli	1K20MH001276	Spinelli & Endicott, 2003	PT vs CTRL-NS	IPT
26	Thompson	1R01MH032157	Gallagher & Thompson, 1982	PT vs PT	BT; CT; STPP
27	Thompson	1R01MH037196	Thompson et al., 1987	PT vs CTRL-NT; PT vs PT	BT; CT
			Thompson et al., 2001	PT vs ADM; PT+ADM vs ADM	STPP; CBT
28	Weissman	1R01MH034501	Foley et al., 1989	PT vs PT	IPT; IPT-CM
29	Wright	1R21MH057470	Wright et al., 2005	PT vs CTRL-NT; PT vs PT	CT; C-CT
Unpublished					
30	Blum	1R01MH025258		PT vs CTRL-NT; PT vs PT	STPP; TOP; CP
31	Chisholm	1F32MH012228	*		
32	Delgado	2R01MH048977		PT vs ADM	CT
33	Gottlieb	1K07MH000597		PT vs ADM; PT+ADM vs ADM	CBT
34	Hauenstein	1R18MH049101		PT vs CTRL-NS	CBT
35	Miller	1R01MH058866		PT+ADM vs ADM; PT+PT+ADM vs PT+ADM	CT; FT
36	Stuart	1R01MH059103		PT vs CTRL-TAU	IPT
37	Szapocznik	5R01MH037379	*		
38	Thase	2R01MH041884		PT vs ADM; PT vs CTRL-PLAC	CBT

* Trial was never started.

[a] Grant number reported incorrectly in the published article, but confirmed with authors. [b] Grant number omitted in published article but confirmed with authors

ADM = Antidepressant medication; BA = behavioral activation; BMT = behavioral marital therapy; BT = behavior therapy; CBT = cognitive behavioral therapy; CCT = client-centered therapy; C-CT = computer-assisted cognitive therapy; CO = combined treatment of BMT and CT; CP = counseling psychotherapy; CT = cognitive therapy; CTRL = Control conditions in general; CTRL-NS = non-specific (psychological placebo) control condition; CTRL-PLAC = placebo control condition; CTRL-NT = no-treatment control condition; CTRL-TAU = treatment-as-usual control condition; DBT = dialectical behavior therapy; EFT = emotion-focused therapy; FEP = focused expressive psychotherapy; FT = family therapy; IFI = inpatient family intervention; IPP = interpersonal psychodynamic psychotherapy; IPT = interpersonal psychotherapy; IPT-CM = conjoint marital interpersonal psychotherapy; P-CT = partial cognitive therapy; PI = principle investigator; PT = psychological treatment; Ref. = reference number SST = social skills training; STPP = short-term psychodynamic/psychoanalytic psychotherapy, TOP = theme-oriented psychotherapy.

Table 2: Meta-analyses of studies examining the effect of psychological treatment for depression

Comparison	*N*	*g*	95% CI	*Z*	*Q (df)*	I^2 *(95% CI)*	Δg^a (%)	Qbetween (df)	*p*
PT vs. CTRL								4.59 (1)	.03
Unpublished	4	0.16	-0.12-0.43	1.11	2.25 (3)	0 (0-85)			
Published	14	0.50	0.35-0.64	6.72**	19.13 (13)	32 (0-64)			
Published + unpublished	18	0.43	0.28-0.57	5.87**	*27.48* (17)	38 (0-65)	-0.07 (-14%)		
PT vs CTRL-NT								0.37 (1)	.54
Unpublished	1	0.50	-0.57-1.56	0.92	0.00 (0)	0 (-)			
Published	3	0.86	0.37-1.36	3.41**	*5.02* (2)	60 (0-89)			
Published + unpublished	4	0.80	0.39-1.21	3.83**	5.72 (3)	48 (0-83)	-0.06 (-7%)		
PT vs CTRL-T								4.11 (1)	.04
Unpublished	3	0.12	-0.12-0.37	0.99	1.54 (2)	0 (0-90)			
Published	11	0.41	0.28-0.54	6.18**	6.12 (10)	0 (0-60)			
Published + unpublished	14	0.35	0.23-0.46	5.93**	11.86 (13)	0 (0-55)	-0.06 (-15%)		
- PT vs CTRL-TAU								0.26 (1)	.61
Unpublished	1	0.31	-0.10-0.71	1.48	0.00 (0)	0 (-)			
Published	2	0.42	0.22-0.63	4.01**	0.02 (1)	0 (-)			
Published + unpublished	3	0.40	0.21-0.58	4.25**	0.28 (2)	0 (0-90)	-0.02 (-6%)		
- PT vs CTRL-PLAC								3.00 (1)	.08
Unpublished	1	-0.09	-0.59-0.40	-0.36	0.00 (0)	0 (-)			
Published	4	0.38	0.18-0.58	3.76**	1.36 (3)	0 (0-85)			
Published + unpublished	5	0.31	0.12-0.51	3.20**	4.35 (4)	8 (0-81)	0.07 (-17%)		

- PT vs CTRL-NS								1.50 (1)	.22
Unpublished	1	0.09	-0.39-0.57	0.36	0.00 (0)	0 (-)			
Published	5	0.45	0.12-0.79	2.65**	4.66 (4)	14 (0-82)			
Published + unpublished	6	0.34	0.05-0.64	2.31*	6.72 (5)	26 (0-69)	-0.11 (-24%)		
PT vs. ADM								1.94 (1)	.16
Unpublished	3	-0.27	-0.70-0.15	-1.26	1.46 (2)	0 (0-90)			
Published	13	0.05	-0.12-0.23	0.58	24.83* (12)	52 (9-74)			
Published + unpublished	16	0.01	-0.16-0.17	0.07	28.62* (15)	48 (6-71)	-0.05 (-89%)		
PT vs. PT								0.35 (1)	.55
Unpublished	1	0.44	-0.38-1.26	1.06	0.00 (0)	0 (-)			
Published	11	0.19	-0.02-0.40	*1.73*	13.31 (10)	25 (0-63)			
Published + unpublished	12	0.20	-0.00-0.40	*1.95*	13.85 (11)	21 (0-59)	+0.01 (+6%)		
PT + ADM vs. ADM								0.21 (1)	.65
Unpublished	2	0.43	0.02-0.83	2.07*	1.30 (1)	23.19 (-)			
Published	7	0.32	0.09-0.55	2.74**	4.27 (6)	0 (0-71)			
Published + unpublished	9	0.34	0.15-0.54	3.40**	5.78 (8)	0 (0-65)	+0.03 (+8%)		

* $p < .05$; ** $p < .01$; Italic numbers indicate a non-significant trend ($p < .10$)

[a] difference between effect size point estimates published studies only and published + unpublished studies, calculated using 5 decimals.

ADM = Antidepressant medication, CTRL = Control conditions in general, CTRL-NS = non-specific (psychological placebo) control condition, CTRL-NT = No-treatment control condition, CTRL-PLAC = placebo control condition; CTRL-T = treatment control condition, CTRL-TAU = treatment-as-usual control condition, PT = psychological treatment.

significant (p=0.16). Adding unpublished to published studies resulted in a pooled mean effect size of g=0.01 (-0.16-0.17), indicating no difference between psychological treatments and antidepressants. Excluding the unpublished study for which original data were not available did not alter this result pattern.

Psychological treatment versus another psychological treatment
One unpublished study was found comparing different types of psychological treatment. The effect size in that study was higher (g=0.44; -0.38-1.26) than in the published studies (g=0.19; -0.02-0.40), but this difference was not significant (p=0.55). Adding this unpublished study to the published studies resulted in an effect size point estimate of g=0.20 (-0.00-0.40, n=11). We also conducted sensitivity analyses for one study intended to minimize allegiance effects (Elkin et al., 1989); changing the psychological treatment ranks did not alter the pattern of results.

Combined psychological treatment and antidepressants versus antidepressant monotherapy
The pooled mean effect size of psychological treatment added to antidepressant medication versus medication mono-therapy did not differ significantly between unpublished (g=0.43; 0.02-0.83; n=2) and published studies (g=0.32; 0.09-0.55; n=7; p=0.65). Adding unpublished to published studies resulted in an effect size point estimate of g=0.34 (0.15-0.54). Excluding the unpublished study for which original data were not available did not alter this result.

Comparing study publication bias across literatures
Turner et al. (2008) identified 74 studies comparing antidepressants to pill-placebo registered with the Food and Drug Administration, of which 51 (69%) were published and 23 (31%) were unpublished. By comparison, we found 14 studies comparing psychological treatment to nonspecific controls, of which 11 (79%) were published and 3 (21%) were unpublished. The proportions were not significantly different (Fisher's Exact Test: p=0.54). The total proportion of unpublished psychological treatment trials (7/36=19%) also did not differ from antidepressant medications (23/74=31%; p=0.26). Using data from Turner et al. (2008), adding unpublished (g=0.15; 0.08-0.22) to published studies (g=0.41; 0.37-0.45) reduced the overall effect size of antidepressants to g=0.34 (0.30-0.39). By comparison, for psychological treatments comparisons to treatment controls, adding unpublished (g=0.12; -0.12-0.37) to published studies (g=0.41; 0.28-0.54) resulted in an effect size of g=0.35; 0.23-0.46). Among comparisons to pill-placebo, adding unpublished (g=-0.09; -0.59-0.40) to published studies (g=0.38; 0.18-0.58) resulted in an effect size of g=0.31 (0.12-0.51).

Discussion

We found clear indications of study publication bias in US National Institute of Health funded randomized clinical trials examining the efficacy of psychological treatment for

major depressive disorder; 19% of the grants that started studies did not result in publication. Although somewhat lower, this proportion did not differ significantly from the 31% non-publication rate reported by Turner et al. (2008) for the industry-funded antidepressant trials. Study publication bias effects were most apparent in comparisons of psychological treatments to controls, resulting in a 14% effect size point estimate decrease when unpublished findings were added to the published ones. Both the current study on psychological treatment and the earlier study by Turner et al. (2008) on antidepressants found that the effect size should be adjusted downward by 0.07 standard deviations. We can have less confidence in the estimate for psychotherapy since it was based on a much smaller sample, but the similarity between the two literatures is striking. Comparisons between different types of psychological treatment were largely unaffected, as were comparisons between psychological treatment alone or combined with antidepressants versus antidepressant monotherapy. Psychological treatment is efficacious and specific, but as for antidepressants, it is less efficacious than the published literature makes it out to be.

The major strength of this study is that it provides a direct assessment of the extent of study publication bias in the psychological treatment literature, instead of relying on an inferred estimate based on statistical procedures. By restricting our attention to an inception cohort of US National Institute of Health funded grants we were able to track the actual number of unpublished trials in that particular subset of the psychological treatment literature. It is reassuring that all of the published studies in our sample reported data in a manner suitable for meta-analysis and that all but one of the unpublished studies could make their data available for our analyses.

The major limitation of the study is that we were not able to ascertain the extent to which the published results might have been inflated by outcome reporting bias. Unfortunately, there is as yet no equivalent to the Food and Drug Administration database for psychological treatment studies that would have allowed us to determine whether published reports accurately reflected the data actually obtained as Turner et al. (2008) were able to do with the antidepressant studies. We tried to obtain the original grants from the US National Institute of Health but only seven of the more recent protocols (18%) had been retained. This contrasts with the success of Chan, Krleža-Jerić, Schmid, and Altman (2004) in obtaining protocols from the Canadian Institutes of Health Research. If we had similar access to the US National Institute of Health protocols, we could have compared the methods proposed in the grants to those described in the publications to determine whether a priori plans had been followed. We believe that the US National Institute of Health should archive protocols and make them publicly available as does its Canadian counterpart (Chan et al., 2004).

However, even with full access to the protocols, we still would not know whether the results reported were obtained using the methods specified a priori in the protocols. Therefore, we cannot say to what extent the effect sizes adjusted for study publication bias alone represents the ‘true’ effect of psychological treatment. If outcome reporting bias did occur and was detected, the adjusted effect would have been lower than what we report after adjusting for study publication bias alone. In

effect, we can never come to terms with the extent of publication bias in the literature unless we can assess both study publication bias and outcome reporting bias. The field will need not just a clinical trials registry but also a data repository for psychological treatment trials similar to the Food and Drug Administration repository for industry-funded pharmacotherapy trials.

Another limitation of this study is the possibility that our search missed grant-funded randomized clinical trials that were begun but never published. We did find one instance in which we had excluded a grant that resulted in a publication that fell within our criteria. We suspect that was a rare occurrence that came about only because the sample studied was available for inclusion in a separate acute treatment trial when they were excluded from the funded maintenance trial (Reynolds, Frank et al., 1999).

A final limitation is that some of the unpublished psychological treatment studies may still be published. Two of the more recent trials obtained effect sizes that might facilitate acceptance and both investigators indicated that they intended to submit their studies. If both were to be accepted that would lower the non-publication rate from 19% to 14% and increase the disparity in effect size between the published and unpublished trials, but it would not alter the “true” effect of psychotherapy based on all initiated trials.

Only two of the other five unpublished studies were submitted for review, largely because the investigators considered their findings to be unimpressive. That is not to say that journals do not play a role in shaping author expectations, but the proximal cause of non-publication in our sample was as likely to be non-submission by the authors as rejection by the journals. This is consistent with an earlier study of US National Institute of Health funded clinical trials that found that findings often remained unsubmitted when the investigators considered their results "not interesting" (Dickersin & Min, 1993) and a recent review found that up to a third of US National Institute of Health funded trials remained unpublished more than four years following study completion (Ross et al., 2012). The bias against null findings is so ingrained in the field that we doubt that simple exhortations to authors or editors will do much to change behavior. Instead, we join with others who recommend that funding agencies or journals should archive both original protocols and raw data from any funded randomized clinical trial (Alsheikh-Ali, Qureshi, Al-Mallah, & Ioannidis, 2011; Chan et al., 2006; Perneger, 2011; Smith & Roberts, 2006; Vickers, 2011).

In conclusion, study publication bias is present in studies funded by US government grants to examine the efficacy of psychological treatment for depression and leads to an overestimation of the effects of psychological treatment for major depressive disorder relative to control conditions. This overestimation is as large in magnitude as that found in antidepressant studies funded by the pharmaceutical industry (Turner et al., 2008). Clinicians, guidelines developers, and decision makers should be aware of overestimated effects of the predominant treatments for major depressive disorder.

References

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

Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H., & Ioannidis, J. P. A. (2011). Public availability of published research data in high-impact journals. *Public Library of Science ONE*, *6(9),* e24357. doi:10.1371/journal.pone.0024357

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

* Beutler, L. E., Engle, D., Mohr, D., Daldrup, R. J., Bergan, J., Meredith, K., & Merry, W. (1991). Predictors of differential response to cognitive, experiential, and self-directed psychotherapeutic procedures. *Journal of Consulting and Clinical Psychology, 59(2),* 333-340.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. UK: Wiley.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis Version 2*. Englewood, NJ: Biostat.

Chan, A-W., Krleža-Jerić, K., Schmid, I., & Altman, D. G. (2004). Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *Canadian Medical Association Journal, 171(7),* 735-740.

Chan, A-W., Upshur, R., Singh, J. A., Ghersi, D., Chapuis, F., & Altman, D. G. (2006). Waiving confidentiality for the greater good. *British Medical Journal*, *332(7549),* 1086-1089.

* Clarkin JF, Glick ID, Haas GL, Spencer, J. H., Lewis, A. B., Peyser, J., ... Lestelle, V. (1990). A randomized clinical trial of inpatient family intervention. V. Results for affective disorders. *Journal of Affective Disorders, 18(1),* 17-28.

* Covi, L., & Lipman, R. S. (1987). Cognitive behavioral group psychotherapy combined with imipramine in major depression. *Psychopharmacology Bulletin, 23(1),* 173-176.

Cuijpers, P., Smit, F., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *The British Journal of Psychiatry, 196,* 173-178.

Cuijpers, P., van Straten, A., Smit, F., & Andersson, G. (2009). Is psychotherapy for depression equally effective in younger and older adults? A meta-regression analysis. *International Psychogeriatrics*, *21(1),* 6-24.

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.

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., . . . Gallop, R. (2005). Cognitive therapy vs. medications in the treatment of moderate to severe depression. *Archives of General Psychiatry, 62,* 409–416.

Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. *Journal of the American Medical Association, 263(10),* 1385–1389.

Dickersin, K., & Min, Y. I. (1993). NIH clinical trials and publication bias. *The Online Journal of Current Clinical Trials*, *2,* Doc No 50.

* DiMascio, A., Weissman, M. M., Prusoff, B. A., Neu, C., Zwilling, M., & Klerman, G. L. (1979). Differential symptom reduction by drugs and psychotherapy in acute depression. *Archives of General Psychiatry, 36(13),* 1450-1456.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., . . . Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the

acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74,* 658–670.

Driessen, E., Cuijpers, P., Hollon, S. D., & Dekker, J. J. M. (2010). Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *Journal of Consulting and Clinical Psychology, 78,* 668-680.

Duval, S., & Tweedie, R. (2000a). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95(449),* 89-98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455-463.

* Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., . . . Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26(1), 57-63.

* Foley, S. H., Rounsaville, B. J., Weissman, M. M., Sholomskas, D., & Chevron E. (1989). Individual versus conjoint interpersonal psychotherapy for depressed patients with marital disputes. *International Journal of Family Psychiatry*, *10(1-2),* 29-42.

Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., & Fawcett, J. (2010). Antidepressant drug effects and depression severity: A patient-level meta-analysis. *Journal of the American Medical Association, 303,* 47–53.

* Gallagher, D. E., & Thompson, L. W. (1982). Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy: Theory, Research and Practice, 19(4),* 482-490.

* Glick, I. D., Clarkin, J. F., Haas, G. L., & Spencer, J. H. (1993). Clinical significance of inpatient family intervention: conclusions for a clinical trial. *Hospital and Community Psychiatry*, *44(9),* 869-873.

* Goldman RN, Greenberg LS, Angus L. (2006). The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. *Psychotherapy Research*, 16(5), 537-549.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23,* 56–62.

* Hayden, T., Perantie, D. C., Nix, B. D., Barnes, L. D., Mostello, D. J., Holcomb, W. L., … Hershey, T. (2012). Treating prepartum depression to improve infant developmental outcomes: a study of diabetes in pregnancy. *Journal of Clinical Psychology in Medical Settings*, 19(3), 285-292. doi: 10.1007/s10880-011-9294-8

* Hersen, M., Bellack, A. S., Himmelhoch, J. M., & Thase, M. E. (1984). Effects of social skills training, amitriptyline and psychotherapy in unipolar depressed women. *Behavior Therapy, 15(1),* 21-40.

* Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1992). Cognitive therapy and pharmacotherapy for depression: singly and in combination. *Archives of General Psychiatry*, *49(10),* 774-781.

Hollon, S. D. (2011). Cognitive and behavior therapy in the treatment and prevention of depression. *Depression and Anxiety*, *28(4),* 263-266. doi:10.1002/da.20797

Ioannidis, J. P. A., Patsopoulos, N. A., & Evengelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *British Medical Journal, 335(7626),* 914-916.

* Jacobson, N. S., Dobson, K., Fruzzetti, A. E., Schmaling, K. B., & Salusky, S. (1991). Marital therapy as a treatment for depression. *Journal of Consulting and Clinical Psychology*, *59(4),* 547-557.

* Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., … Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology*, 64(2), 295-304.

* Jarrett, R. B., Schaffer, M., McIntire, D., Witt-Browder, A., Kraft, D., & Risser, R. C. (1999). Treatment of atypical depression with cognitive therapy or phenelzine: a double-blind, placebo-controlled trial. *Archives of General Psychiatry*, 56(5), 431-437.

Kessler, R. C. (2012). The costs of depression. *Psychiatric Clinics of North America, 35(1),* 1-14.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., … Wang, P. S. (2003). The epidemiology of major depressive disorder: results from the national comorbidity survey replication (NCS-R). *Journal of the American Medical Association, 289(23)*, 3095–3105.

Kupfer, D. J., Frank, E., & Phillips, M. L. (2012). Major depressive disorder: new clinical neurobiological, and treatment perspectives. *Lancet,* 379*(9820),* 1045-1055.

* Lustman, P. J., Griffith, L. S., Freedland, K. E., Kissel, S. S., & Clouse, R. E. (1998). Cognitive behavior therapy for depression in type 2 diabetes mellitus. A randomized, controlled trial. *Annals of Internal Medicine*, *129(8),* 613-621.

* Lynch, T. R., Morse, J. Q., Mendelson, T., & Robins, C. J. (2003). Dialectical behavior therapy for depressed older adults: a randomized pilot study. *American Journal of Geriatric Psychiatry*, *11(1),* 33-45.

Mathers, C. D., & Loncar D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *Public Library of Science Medicine*, *3,* e442. doi:10.1371/journal.pmed.0030442

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.

* Miller, I. W., Norman, W. H., & Keitner, G. I. (1989). Cognitive-behavioural treatment of depressed inpatients. *Behavior Therapy, 20(1),* 25-47.

* Miranda, J., Chung, J. Y., Green, B. L., Krupnick, J., Siddique, J., Revicki, D. A., & Belin, T. (2003). Treating depression in predominantly low-income young minority women: a randomized controlled trial. *Journal of the American Medical Association,* 290(1), 57-65.

* Mohr, D. C., Boudewyn, A. C., Goodkin, D. E., 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. *Journal of Consulting and Clinical Psychology,* 69(6), 942-949.

* Murphy, G. E., Carney, R. M., Knesevich, M. A., Wetsel, R. D., & Whitworth, P. (1995). Cognitive behavior therapy, relaxation training and tricyclic antidepressant medication in the treatment of depression. *Psychological Reports*, 77(2), 403-420.

* Murphy, G. E., Simons, A. D., Wetzel, R. D., & Lustman, P. J. (1984). Cognitive therapy and pharmacotherapy. Singly and together in the treatment of depression. *Archives of General Psychiatry, 41(1),* 33-41.

* O'Hara, M. W., Stuart, S., Gorman, L. L., & Wenzel, A. (2000). Efficacy of interpersonal therapy for postpartum depression. *Archives of General Psychiatry*, *57(11),* 1039-1045.

Orsini, N., Higgins, J., Bottai, M., & Buchan, I. (2005). Heterogi: Stata module to quantify heterogeneity in a meta-analysis.

Perneger, T. V. (2011). Sharing raw data: another of Francis Galton's ideas. *British Medical Journal*, *342(7807),* d3035.

Reynolds, C. F., Frank, E., Perel, J. M., Imber, S. D., Cornes, C., Miller, M. D., … Kupfer, D. J. (1999). Nortriptyline and interpersonal psychotherapy as maintenance therapies for recurrent major depression. A randomized controlled trial in patients older than 59 years. *Journal of the American Medical Association*, *281(1),* 39-45.

* Reynolds, C. F., Miller, M. D., Pasternak, R. E., Frank, E., Perel, J. M., Cornes, C., … Kupfer, D. J. (1999). Treatment of bereavement-related major depressive episodes in later life: a controlled study of acute and continuation treatment with nortriptyline and interpersonal psychotherapy. *American Journal of Psychiatry, 156(2),* 202-208.

Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin, 86,* 638-641.

Ross, J. S., Tse, T., Zarin, D. A., Xu, H., Zhou, L., & Krumholz, H. M. (2012). Publication of NIH funded trials registered in ClinicalTrials.gov: Cross sectional analysis. *British Medical Journal, 344,* d7292. doi:10.1136/bmj.d7292.

* Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. D. (1977). Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy Research, 1(1),* 17-37.

* Schulberg, H. C., Block, M. R., Madonia, M. J., Scott, P., Rodriguez, E., Imber, S. D., … Coulehan, J. L. (1996) Treating major depression in primary care practice. Eight-month clinical outcomes. *Archives of General Psychiatry*, *53(10),* 913-919.

Smith, R., & Roberts, I. (2006). Patient safety requires a new way to publish clinical trials. *Public Library of Science Clinical Trials*, *1(1),* e6. doi:10.1371/journal.pctr.0010006.

Song, F., Parekh-Bhurke, S., Hooper, L., Loke, Y. K., Ryder, J. J., Sutton, A. J., … Harvey, I. (2009). Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. *BMC Medical Research Methodology, 9,* 79. doi:10.1186/1471-2288-9-79

* Spinelli, M. G., & Endicott, J. (2003). Controlled clinical trial of interpersonal psychotherapy versus parenting education program for depressed pregnant women. *American Journal of Psychiatry*, 160(3), 555-562.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35(6),* 773-785.

StataCorp. (2009). *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Sterne, J. A., Egger, M., & Smith, G. D. (2001). Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *British Medical Journal, 323(7304),* 101-105.

* Thompson, L. W., Coon, D. W., Gallagher-Thompson, D., Sommer, B. R., & Koin. D. (2001). Comparison of desipramine and cognitive/behavioral therapy in the treatment of elderly outpatients with mild-to-moderate depression. *American Journal of Geriatric Psychiatry,* 9(3), 225-240.

* Thompson, L. W., Gallagher, D., & Breckenridge, J. S. (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology,* 55(3), 385-390.

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,* 252–260.

Vickers, A. J. (2011). Making raw data more widely available. New initiative by the Wellcome Trust sets out some guiding principles. *British Medical Journal*, *342(7805),* d2323.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: empirically, “all must have prizes”. *Psychological Bulletin, 122(3),* 203-215.

* Weissman, M. M., Prusoff, B. A., DiMascio, A., Neu, C., Goklaney, M., & Klerman, G. L. (1979). The efficacy of drugs and psychotherapy in the treatment of acute depressive episodes. *American Journal of Psychiatry, 136(4B),* 555-558.

* Wright, J. H., Wright, A. S., Albano, A. M., Basco, M. R., Goldsmith, L. J., Raffield, T., & Otto, M. W. (2005). Computer-assisted cognitive therapy for depression: Maintaining efficacy while reducing therapist time. *American Journal of Psychiatry, 162(6),* 1158-1164.