Control condition design and implementation features in controlled trials: a meta-analysis of trials evaluating psychotherapy for depression

David C Mohr, Ph.D,¹ Joyce Ho, Ph.D,¹ Tae L Hart, Ph.D,² Kelly G Baron, Ph.D,¹ Mark Berendsen, MLIS,¹ Victoria Beckner, Ph.D,³ Xuan Cai, M.S,¹ Pim Cuijpers, Ph.D,⁴ Bonnie Spring, Ph.D,¹ Sarah W Kinsinger, Ph.D,¹ Kerstin E Schroder, Ph.D,⁵ Jenna Duffecy, Ph.D

¹Northwestern University, Chicago, IL, USA

²Reyerson University, Toronto, ON, Canada

³University of California, San Francisco, CA, USA

⁴Vrije Universiteit Amsterdam, Amsterdam, Netherlands

⁵University of Alabama, Birmingham, AL, USA

Correspondence to: D Mohr
d-mohr@northwestern.edu

Cite this as: *TBM* 2014;4:407–423
doi: 10.1007/s13142-014-0262-3

Abstract

Control conditions are the primary methodology used to reduce threats to internal validity in randomized controlled trials (RCTs). This meta-analysis examined the effects of control arm design and implementation on outcomes in RCTs examining psychological treatments for depression. A search of MEDLINE, PsycINFO, and EMBASE identified all RCTs evaluating psychological treatments for depression published through June 2009. Data were analyzed using mixed-effects models. One hundred twenty-five trials were identified yielding 188 comparisons. Outcomes varied significantly depending control condition design ($p < 0.0001$). Significantly smaller effect sizes were seen when control arms used manualization ($p = 0.006$), therapist training ($p = 0.002$), therapist supervision ($p = 0.009$), and treatment fidelity monitoring ($p = 0.003$). There were no significant effects for differences in therapist experience, level of expertise in the treatment delivered, or nesting vs. crossing therapists in treatment arms. These findings demonstrate the substantial effect that decisions regarding control arm definition and implementation can have on RCT outcomes.

Keywords

Meta-analysis, Depression, Control conditions, Randomized controlled trial design, Methodology

Over the past half century, evidence has accumulated to support a number of psychological and behavioral interventions for mental health and medical conditions [6]. The backbone of treatment outcome research is the randomized controlled trial (RCT), a planned experiment designed to test the efficacy or effectiveness of an intervention. Although many aspects of RCT methodology have received considerable attention [2], until recently, surprisingly little attention has been paid to how to select and implement control conditions. The aim of this paper is to examine the effects of the design and implementation of control conditions on RCT outcomes for the treatment of depression using meta-analysis. These results will be interpreted in

Implications:

Research: The design, selection, and implementation of control arms can have a considerable impact on the outcome of randomized controlled trials. Differences in control arm design and treatment fidelity procedures across treatment arms can substantially inflate effect sizes, while the experience and expertise of clinicians and nesting or crossing clinicians with treatment arms may have relatively little effect on outcomes.

Practice: While evidence derived from randomized controlled trials should be an important part of clinical decision making, practitioners should be aware that the design of clinical trials can influence their outcomes.

Policy: When examining the randomized controlled trial literature for the purposes of systematic review, guideline development, and policy, the impact of trial design and treatment implementation procedures should be considered.

light of recent efforts to formulate a framework to support decisions regarding the selection, design, and implementation of control conditions [20].

RCTs can vary in their aim, from explanatory trials evaluating efficacy or effectiveness under ideal conditions, to more pragmatic trials that evaluate the intervention under conditions found in clinical settings [13, 28]. In either case, the experimental treatment is always determined *relative* to a control condition. Consequently, what an RCT reveals about the effectiveness of the experimental treatment inherently depends as much on the control condition as on the experimental treatment. One of the principal reasons for using a control condition is to eliminate alternative causal explanations. In statistical terms, the purpose of a control condition is to filter out the variance due to factors that are not specific to the experimental intervention, leaving only the variance due specifically

to this treatment. A well-designed control condition should maximize our confidence that any positive results are due to the treatment and not to other factors. Control arms are usually expected to remove traditional threats to validity such as changes in the treatment target (e.g., natural history of the disorder), statistical regression, attrition, and effects repeated testing [5]. In addition, some control conditions, such as treatment as usual (TAU) or active comparators may control for the effects of standard practices or treatments.

While control arms can control for these threats to internal validity, it has been posited that much of the effect size of RCTs depends upon the design and implementation of control conditions [20]. To date a few meta-analyses have included analyses examining the effects of control conditions and have not reported large control condition effects [10, 18]. However, this is the first meta-analysis to focus exclusively on the effects that control conditions and their implementation have on the outcomes of RCTs of psychotherapy for depression using a granularly defined categorization of control conditions and implementation factors. We tested the following hypotheses:

1. Control condition design would have a significant effect on outcomes. The literature would suggest some control conditions such as no-treatment, wait-list, or minimal treatment controls would result in larger effect sizes, while others, such as active comparators, would result in smaller effect sizes. However, we did not make specific a priori hypotheses regarding the order of effect sizes, or differences between specific control condition designs.
2. Control treatments that do not employ treatment fidelity and implementation procedures recommended by Bellg et al. [2], including treatment manuals, therapist training, supervision, and fidelity monitoring would produce larger effect sizes, compared to trials where these are implemented.
3. Larger effect sizes would be observed in trials in which control arms used less experienced clinicians, or where there is less congruence between clinician expertise and control treatment modality relative to experimental treatment clinicians.
4. An exploratory aim evaluated the effect of crossing vs. nesting therapists in treatment arms.

Understanding the effects of control arm design and implementation considerations is important for investigators in supporting decision making during RCT design, methodologists involved in the development of decision making frameworks for RCT design, and consumers of RCT data, such as those conducting systematic reviews, meta-analyses, or developing clinical guidelines.

METHODS

Inclusion criteria

Studies were included based on the following eligibility criteria: (1) adult (age 18+) participants;

(2) inclusion criteria of a depressive disorder diagnosis or use of a cutoff identifying elevated symptom severity; (3) a psychological treatment was tested in at least one arm. Based on findings that substantial effects can be measured in the first 4 weeks of treatment [16], psychological treatments had to include at least four sessions, delivered by a clinician; (4) presence of a control comparison arm; (5) random treatment assignment; (6) a validated measure of depression was used as an outcome; (7) paper was published in English in a peer-reviewed journal.

We selected one clinical condition to limit the potential confounding influence of variability in effect sizes by clinical disorder and varying prevalence of control arm design and implementation factors across clinical fields. The volume of RCT data on depression and the variety of control arms employed was believed to be sufficient to provide adequate numbers of studies to evaluate the proposed analyses.

Information sources and systematic searches

A health sciences librarian searched MEDLINE, PsycINFO, EMBASE, and the Cochrane Library. Appropriate controlled-vocabulary terms specific to each database and keyword searching within title and abstract fields were used to retrieve studies indicating depression in conjunction with behavioral interventions or psychotherapy. The searches were restricted to studies of adult subjects published in English from the mid 1960s through June 2009 and filters were employed to retrieve specific study types. The MEDLINE search strategy is attached as Appendix 1. Additionally, the reviewers examined the reference lists from previous meta-analyses and systematic reviews identified through the searches.

Data collection process

Eight PhD clinical psychologists served as reviewers and performed data extraction. All abstracts were reviewed for exclusion. If the abstract was not clearly excluded, full articles were reviewed. Data were extracted for all included papers. All review and extraction procedures were conducted by teams of two reviewers who were required to achieve consensus on each decision or extraction.

A preliminary coding sheet along with a data extraction definition document were developed and piloted by all eight reviewers on two articles. The reviewers met by conference call to review coding and identify problems with the items. The coding form and definitions document were revised and tested again with another two articles. This process continued iteratively until all reviewers were satisfied that the item definitions were clear and representative of the data, and that reviewers were obtaining similar codes. The coding sheet was then translated into a web-based format. Pairs of re-

viewers coded each paper on the website. The coding website automatically detected discrepancies and alerted reviewers. Reviewers resolved these discrepancies through discussion and consensus. Conference calls were held after reviewer groups completed three to five papers to resolve questions and problems that arose during coding. When sufficient data were not available to allow an item to be coded, the study authors were contacted by e-mail with a request for the information. For each study author contacted, up to three requests were made over a 3-week period.

Data items

Depression outcomes were based on self-report data, as they have been shown to be more conservative than clinician-rated scores [7]. If more than one measure was available, the measure most commonly represented in this meta-analysis was used.

Control condition design

Control conditions were defined based on primarily on definitions described in Mohr et al. [20], which was refined by the coding team based on inspection of the control treatments used in the included papers. Control conditions were defined a priori as (1) no-treatment control, which contained no study treatment and was not conducted in setting where treatment would be available; (2) wait-list control (WLC), which provided no treatment during the period of the experimental treatment, but offered the experimental treatment or some equivalent after post-treatment assessment; (3) TAU, which required that trial was conducted in a clinic where patients had access to some form of treatment; (4) non-specific factors component control, which provided time with therapist equivalent to experimental treatment but only provided non-specific factors provided; (5) specific factors component control, which therapist time equivalent to the experimental condition, but a different or reduced number of specific factors in addition to the non-specific factors; (6) active comparator, an evidence-based treatment that would not be expected to differ from the experimental treatment; and (7) pill placebo. Minimal treatment control, defined as treatments that entailed less than four sessions, was added by the coding team to account for control conditions that did not meet the a priori definitions.

Control condition implementation variables

These variables reflected procedures used to ensure the reliability of clinician-administered control treatments and were derived from the NIH Behavior Change Consortium [2] and procedures identified as likely contributors to outcome variance in our previous work [20]. These variables were only

extracted for control arms that employed therapists. Studies in which these implementation procedures were not employed in the experimental treatment were excluded to focus the comparison on the effects of control arm implementation. Comparisons in which no study administered treatments were employed (no-treatment and WLC) were excluded from these analyses.

Treatment manualization—Treatment manualization for the control condition included three levels: (1) manualized, indicating that a treatment manual, existed, was used, and was cited; (2) treatment definition, indicating that a general treatment approach was described, but no treatment manual was cited or described; (3) no definition, indicating that there was no evidence of a treatment manual or clearly articulated treatment approach.

Therapist training—This refers to training in a treatment model prior to providing the study treatment and included two levels: (1) therapist training evident in both the experimental and control conditions and (2) therapist training evident in the experimental condition but not the control condition.

Therapist supervision—Therapist supervision was defined as any description of supervision of therapists providing study treatments and included two levels: (1) therapist supervision evident in both the experimental and control conditions; and (2) therapist supervision evident in the experimental condition but not the control condition.

Therapist fidelity monitoring—This was coded, based on descriptions in the paper, for both experimental and control treatments and included two levels: (1) therapist fidelity monitoring evident in both the experimental and control conditions, and (2) therapist fidelity monitoring evident in the experimental condition but not the control condition.

Clinician selection biases

These variables reflected potential biases that may arise from differences in clinician selection, which have been noted as potential threats to internal validity.

Therapist experience—This variable included two levels: (1) therapists in the experimental treatment arm were more experienced, based on degree or years in practice, than those in the control arm; and (2) no difference in experience level.

Clinician expertise—This variable reflects an indication that there was a bias in the expertise or treatment orientation of therapists, such as engaging therapists with prior expertise only in the experimental treatment modality for both treatment arms. Three levels were coded: (1) unbiased, indicating no evidence of expertise bias; (2) biased towards experimental condition; and (3) biased towards control condition.

Nesting vs. crossing of therapists—Nesting referred to having therapists provide treatment in only one

treatment arm, while crossing was coded when therapists provided both treatments. Articles in which this could not be discerned were excluded.

Study quality and other variables—Studies were assessed for quality using a version of the PEDro scale [19] modified to include those criteria not already part of the inclusion criteria (e.g., randomization) or primary analyses (e.g., fidelity ratings). Additional items that were coded included mean age, percent female, whether MDD status was used as an inclusion criterion, length of experimental and control intervention (number of weeks), and year of publication.

Data analysis

We combined outcome data on depression symptom severity scores across trials with standard meta-analytic methods. The primary measure of effect size was Hedges' g , a standardized mean difference, which is an unbiased estimate that enables inclusion of different outcome measures in the same synthesis. Hedges' g and 95 % confidence interval were calculated with Comprehensive Meta-Analysis (CMA) version 2.2 [4] using pre- and post-treatment means and standard deviations. CMA calculates $g = d \times J$. The standardized mean difference is calculated $d = \frac{\bar{X}_1 - \bar{X}_2}{\Delta\sigma_{\text{pooled}}}$, where the pooled standard deviation is $\Delta\sigma_{\text{pooled}} = \sqrt{((n_1 - 1) \times \Delta\sigma_1^2 + (n_2 - 1) \times \Delta\sigma_2^2) / (n_1 + n_2 - 2)}$. d is then multiplied by a correction factor (J) to compute g , where $J = 1 - \frac{3}{4(df - 1)}$, where $df = n - 2$.

The Cochran's Q statistic and the Higgins' I^2 statistics were used to determine heterogeneity between studies. An I^2 of 0 indicates no heterogeneity, with 25 % considered low, 50 % considered moderate, and 75 % considered high [15]. A low p value (≤ 0.05) for the Q statistics was considered evidence of significant heterogeneity. As we expected significant heterogeneity, mean effect sizes were calculated using random effects models. Subgroup analyses were conducted using mixed effect models, in which subgroups are pooled with the random effects model, while tests for significant differences between subgroups are conducted with the fixed effects model. Given the multiple analyses, alpha-adjustment is advisable. A Bonferroni adjustment would set alpha at 0.0063, which is a level we are concerned increases the likelihood of type II error. We have therefore set the significance at $p < 0.01$ for omnibus tests, to balance the threats of type I vs. type II error.

Publication bias was evaluated with a funnel plot of pooled effect size versus its standard error [26]. Duval and Tweedie's trim and fill procedure [11] was applied, which yields an effect size estimate after publication bias has been taken into account. These were conducted only on the overall effect size to provide a general indication of publication bias.

These analyses were not run for each subgroup analysis, as non-publication is likely based on the implications of non-significant results for experimental treatments rather than on control treatments, which were the subject of subgroup analyses.

More than two conditions were tested in many RCTs, resulting in multiple comparisons. These multiple comparisons are not independent of each other, which could result in an artificial reduction of heterogeneity. Therefore, we conducted sensitivity analyses to examine the influence of specific studies, in which only one comparison was included from trials with multiple comparisons [27]. One sensitivity analysis included only the comparison with the largest effect size because this was considered the most conservative approach in estimating heterogeneity in the meta-analyses, and one sensitivity analysis included only the smallest effect size. Data analyses were conducted using SAS9.2 [24] and Comprehensive Meta-Analysis (CMA) version 2.2 [4].

RESULTS

Description of included studies

The flow of information is displayed in Fig. 1. A total of 3,251 abstracts were examined and 667 full articles were reviewed for inclusion. This resulted in the inclusion of 125 trials, with 10,077 participants. Of the 52 authors contacted for additional information, 22 provided the requested data.

A single comparison of one experimental treatment and one control/comparison arm was used in 78 studies; 20 studies examined two experimental arms and one control arm; 15 studies examined one experimental arm and two control arms; 7 studies examined two experimental arms and two control arms; 3 studies examined three experimental arms and one control arm; 1 study examined one experimental arm and three control arms; 1 study examined three experimental arms and two control arms.

In these 125 trials, there were 160 experimental treatments and 150 control arms for a total of 188 comparisons. The experimental treatments included cognitive behavior therapy (58 trials; 100 comparisons), behavioral therapy (6 trials; 7 comparisons); interpersonal psychotherapy (10 trials; 11 comparisons); problem-solving therapy (8 trials; 13 comparisons); psychodynamic therapy (3 trials; 4 comparisons); experiential/client centered therapies (5 trials; 6 comparisons); and 35 trials (47 comparisons) with other types of therapies (e.g., social skills training, dialectical behavioral therapy, positive psychotherapy, etc.) in the experimental group. Consistent with previous meta-analyses, there were no significant differences in effect sizes across experimental treatment orientations ($p = 0.25$) [8].

Effect-sizes for comparisons were based on the Beck Depression Inventory (149, 78.7 %), the

Fig 1 | Flow of information

Center for Epidemiological Studies Depression Scale (9, 4.8 %), the Symptom Checklist-90 depression score (6, 3.2 %), and smaller numbers of other scales (25, 13.3 %). In 75 studies, participants met diagnostic criteria for a depressive disorder, while the other 50 studies required high scores on a measure of depression symptom severity for inclusion. Fifteen studies included women only, 18 were aimed at older adults and 21 at depressed patients with general medical disorders (e.g., HIV, cardiovascular disease, cancer, multiple sclerosis, etc.).

Control condition design

The number of control condition design comparisons is listed in Table 1. As TAU can vary in its definition, we note that TAU comparisons included 20 that used a medical care setting with no study-related care enhancement, 11 that used a medical care setting with some mental health enhancement (e.g., giving providers baseline depression information or providing information to participant), and 3 that used TAU mental healthcare settings such as inpatient units. Active comparators were coded as nine comparisons using CBT, five using psychodynamic therapy, and two using interpersonal therapy. The random effects model comparing each of the eight control conditions to their experimental treatment on change in depression showed an overall effect

favoring experimental over control treatments of $g=0.54$ (95 % CI 0.45–0.64), $p<0.0001$ (see Table 1 for all meta-analytic results). Heterogeneity was very high, with $I^2=81.53$ % and $Q=974.39$, $p<0.0001$. There were significant differences in effect size across control conditions ($p<0.0001$). The findings were substantially similar for the sensitivity analyses ($ps<0.0001$).

Egger's test of the intercept was significant ($p=0.0007$), suggesting a publication bias for studies with larger effect sizes. The Duval and Tweedie trim and fill procedure [11] was used to calculate the number of missing studies and provide an estimate of what the effect size would have been without bias. Using this method, 35 studies were imputed, which reduced the overall effect size to 0.26. However, the 95 % CI of 0.15 to 0.38 did not include 0.

Minimal treatment controls, non-specific factor controls, and active comparators did not produce significant effect sizes ($ps>0.11$), while all other control condition designs did produce significant effect sizes ($ps<0.01$). It should be noted that there were only three studies with five comparisons that used minimal treatment controls, which limits their reliability.

Effect sizes for trials using no-treatment controls were significantly larger than those using non-specific component controls ($p=0.03$), active comparators ($p=0.004$), and minimal treatment controls ($p=0.01$). The lower effect sizes of no

Table 1 | Results of meta-analyses examining control condition design and implementation features

Variable	N _{comp}	g	Variance	95% CI	Z	I ² (%)	Q	p value (within*)	p value (between§)
Control Condition Design Overall	188	0.54	0.002	0.45~0.64	11.18	81.53	974.39***	<0.0001	<0.0001
No treatment	13	0.56	0.03	0.22~0.90	3.20	70.14	40.19***	0.001	
Wait-list	42	0.95	0.01	0.74~1.16	8.89	71.76	145.20***	<0.0001	
Minimal treatment	5	-0.13	0.10	-0.75~0.49	-0.40	90.60	351.10***	0.69	
Treatment as usual	34	0.93	0.01	0.72~1.15	8.48	70.84	123.44***	<0.0001	
Non-specific factor	37	0.18	0.01	-0.04~0.40	1.61	83.56	182.47***	0.11	
Specific factor	31	0.35	0.01	0.11~0.58	2.85	90.68	96.41***	0.004	
Pill-placebo	10	0.48	0.04	0.10~0.86	2.48	54.34	32.85**	0.01	
Active comparator	16	0.09	0.03	-0.23~0.40	0.54	0.00	2.74	0.59	
Control Treatment Implementation									
Manualization overall	94	0.25	0.004	0.12~0.37	3.80	76.77	391.68***	0.0001	0.009
Manualized	67	0.17	0.006	0.02~0.32	2.28	71.17	228.96***	0.02	
Approach only	12	0.11	0.04	-0.27~0.48	0.55	72.44	39.91***	0.58	
No definition	15	0.72	0.03	0.39~1.05	4.28	88.60	122.81***	<0.0001	
Therapist Training overall	67	0.26	0.006	0.12~0.41	3.50	77.88	293.79***	0.0005	0.002
No Training	10	0.79	0.04	0.42~1.16	4.22	93.91	147.80***	<0.0001	
Training	57	0.16	0.007	0.001~0.32	1.98	61.64	146.00***	0.05	
Therapist Supervision overall	71	0.21	0.004	0.08~0.34	3.19	70.42	233.27***	0.001	0.009
No Supervision	9	0.69	0.04	0.31~1.08	3.52	89.79	78.33***	0.0004	
Supervision	62	0.15	0.005	0.01~0.29	2.13	60.63	154.94***	0.03	
Therapist Fidelity Monitoring - overall	61	0.23	0.004	0.11~0.36	3.59	67.02	178.90***	0.0003	0.003
No Monitoring	7	0.76	0.04	0.39~1.14	4.03	92.44	79.35***	<0.0001	
Monitoring	54	0.16	0.005	0.03~0.30	2.34	46.76	99.55***	0.02	
Clinician Bias									
Clinician experience - overall	80	0.24	0.006	0.08~0.39	3.02	81.48	421.07***	0.003	0.09
Greater in Exper tx	4	-0.35	0.13	-1.05~0.35	-0.98	93.69	47.54***	0.33	
Equivalent	76	0.27	0.007	0.11~0.43	3.31	79.92	373.53***	0.0009	
Clinician expertise - overall	73	0.22	0.006	0.06~0.38	2.77	81.87	391.60***	0.006	0.93
Bias for Exper tx	52	0.22	0.009	0.04~0.41	2.41	85.57	353.35***	0.02	
No bias	21	0.21	0.02	-0.09~0.51	1.37	47.71	38.25**	0.17	
Therapist Crossed vs. Nested Therapists - overall	75	0.17	0.006	0.02~0.32	2.23	77.53	335.76***	0.03	0.36
Crossed	37	0.10	0.01	-0.11~0.31	0.90	52.96	76.53***	0.37	
Nested	38	0.24	0.01	0.03~0.45	2.23	85.10	259.22***	0.03	

*_c0.05; **_c0.01; ***_c0.001+ Within *p*-values refer to effects within factors§ Between *p*-values refer to effects between factors

treatment controls compared to TAU were not significant ($p=0.08$). Effect sizes for trials using no-treatment controls were significantly smaller than those using WLCs ($p=0.03$). There were no significant differences between no treatment controls and specific factor controls or pill placebo controls ($ps>0.34$).

WLCs produced effect-sizes that were significantly larger than minimal treatment controls ($p=0.0005$), non-specific factor component controls ($p<0.0001$), specific factor controls ($p<0.0001$), pill placebos ($p=0.03$), and active comparators ($p<0.0001$). There were no significant differences between WLCs and TAU ($p=0.94$).

TAU produced significantly larger effect sizes than minimal treatment controls ($p=0.003$), non-specific component controls ($p<0.0001$), specific component controls ($p=0.0008$), and active comparator ($p<0.0001$). The larger effect sizes produced by TAU compared to pill placebo controls reached only trend significance ($p=0.06$).

There were no significant differences for non-specific component controls and specific component controls, pill placebo, active comparator or minimal treatment controls ($ps>0.16$). There were no significant differences for specific component controls compared to minimal treatment controls, pill placebo, or active comparator ($ps>0.23$). Pill placebo control were not significantly larger than active comparators ($p=0.09$) and not significantly different from minimal treatment controls ($p=0.12$). There was no significant difference between active comparator and minimal treatment controls ($p=0.46$).

Sensitivity analyses were substantially similar for all analyses.

Because different control conditions may be used for trials of different length, the number of weeks from baseline to post-treatment assessment was entered as a covariate. Length of trial did not have a significant effect on effect size ($p=0.69$) and control condition design remained significant ($ps<0.0001$). We also verified the findings with an analysis of the pre-post effect sizes within the control group. These findings were substantially similar ($p<0.0001$ for between groups effect).

Because differences in length of treatment between experimental and control conditions could be particularly relevant for specific, non-specific, and active comparator controls, we examined the relationship between the difference in length of treatment between experimental and control conditions, and effect-size in these two conditions. We found no significant difference in length of treatment between experimental and control conditions within non-specific ($p=0.32$), specific, ($p=0.33$), or active comparators ($p=0.72$), nor was there any significant relationship between difference in length of treatment and effect-size within non-specific controls ($p=$

0.25), specific controls ($p=0.25$), and active comparators ($p=0.97$).

Control condition implementation

Comparisons in which no study administered treatments were employed (e.g., no-treatment and WLC) were excluded from these analyses. The difference in the number of comparisons between the number of possible comparisons (133) and the reported number of comparisons reflects the number of comparisons in which the control implementation procedure was not used in the experimental arm.

Manualization of control arm

Level of definition of the control arm was significantly related to treatment outcomes ($p=0.009$). Pairwise analyses revealed that trials in which there was no description of the control treatment approach produced significantly larger effect sizes than trials in which the control condition was manualized ($p=0.003$) but not significantly larger than trials that described theoretical approaches for control treatments ($p=0.09$). There was no significant difference in depression outcomes between trials that used treatment manuals and trials that described theoretical approaches ($p=0.69$).

Therapist training in control arm

In all trials, therapists in the experimental arm received training. Trials in which therapist training was provided to control therapists produced significantly smaller effect sizes, compared to those in which no training was provided ($p=0.002$).

Therapist supervision in control arm

Only one trial provided no therapist supervision in the experimental condition; accordingly, that trial was dropped from this analysis. Trials in which therapist supervision was provided to control therapists produced significantly smaller effect sizes, compared to those in which no supervision was provided ($p=0.009$).

Therapist fidelity monitoring in control arm

Only one study reported no therapist fidelity monitoring in the experimental condition; accordingly, that study was dropped from this analysis. Fidelity monitoring in the control condition was associated with significantly lower effect sizes compared to those that did not ($p=0.003$).

Results from sensitivity analyses were similar for all control implementation analyses.

RCTs that did not use a control treatment manual might be less likely to monitor treatment adherence. We therefore examined the frequencies with which manualization status coincided with other control implementation variables (manualization, training, supervision, and monitoring) among comparisons where control treatments used clinician-delivered treatments. Among the 67 comparisons that used a control treatment manual, 37 included all implementation procedures, 14 included two other procedures, 4 included only one other procedure, and 12 included no other implementation procedures. Of the 12 comparisons that were coded as including a description of control treatment approach, but no manual, 1 included all implementation procedures, 4 included two other procedures, 4 included only one other procedure, and 3 included no other implementation procedures. Of the 15 comparisons that used no control treatment manual or reference to approach, 3 inattention variables in one analysis.

Clinician selection biases

Comparisons using control arms without study therapists (e.g., no treatment, WLC, and TAU) were excluded from analyses of clinician selection bias, as well as trials for which information regarding clinician experience and expertise were unavailable.

Clinician experience

No studies showed a bias in favor of control treatments. There was no significant difference in effect sizes between comparisons in which clinician experience was similar across treatment arms and comparisons in which clinician experience was greater in the experimental arm than in the control arm ($p=0.09$).

Clinician expertise

No studies showed a bias in favor of control treatments. Differences in the expertise of clinicians and their familiarity with the treatments delivered across study arms had no significant effect on outcomes ($p=0.93$).

Therapist nesting vs. crossing

There was no significant difference in effect size for RCTs that nested therapists within treatment condition, compared to those that crossed therapists ($p=0.36$).

Results were similar in sensitivity analyses for all analyses of clinician bias. Metaregression analyses showed no significant for length of trial (all $ps>0.69$) or use of MDD inclusion criterion (all $ps>0.92$) on clinician selection bias analyses,

and all selection bias variables remained non-significant ($ps>0.48$).

Quality indicators and other analyses

There were no significant effects for any quality indicators, including allocation concealment ($p=0.80$), treatment groups similar at baseline ($p=0.67$), outcome obtained in $>85\%$ of sample ($p=0.55$), and intent-to-treat analyses performed ($p=0.50$). Outcomes were also unrelated to age ($p=0.79$) or gender ($p=0.83$) of the participants. Effect sizes were unrelated to the use of MDD as an outcome variable ($p=0.86$) or year of publication ($p=0.92$).

DISCUSSION

The meta-analysis found significant differences in effect sizes generated across control arms, supporting the hypothesis that the choice of control condition can have a very large impact on the outcome of an RCT for psychotherapy to treat depression. In general, WLCs and TAU controls produced the largest trial effect sizes in the 0.93–0.95 range; no treatment, specific factor and placebo controls produced similar results in the moderate 0.35–0.56 range; active comparators, pill placebo, and minimal treatment controls produced small, non-significant results. This range of effect sizes is far greater than the range of effect sizes seen across psychotherapies for depression [8], underscoring the importance of decisions regarding control conditions on the data resulting from RCTs.

Because the care that is actually provided under TAU is often not monitored or adequately reported, it is difficult to interpret these effect sizes. While some TAU settings may provide considerable treatment, other TAU settings provide very little or suboptimal mental health care, which may account for the large effect sizes resulting from the use of TAU controls. We note that the minimal treatment control was a classification developed by the coding team after the initiation of the review to accommodate control conditions that did not meet the a priori categories. The five comparisons included under minimal treatment control come from three small studies [1, 29, 30]. Two of these studies had seemingly anomalous findings, one with a minimal contact bibliotherapy that produced greater improvements than a full treatment [30] and one with a one-session treatment that produced a larger effect than a six-session treatment [1]. Thus, findings related to minimal treatment controls may be related the unreliability of small numbers of small studies, rather than evidence of the lack of efficacy of minimal treatment controls. The remaining analyses of control groups used a priori classification categories and included 10 to 42 comparisons,

increasing confidence in our general finding that different control conditions produced significantly large differences in outcomes.

No-treatment controls produced an effect size of 0.56, which is significantly smaller than those produced by WLCs. While not widely discussed, others have observed no-treatment arms can produce greater improvements than wait-list or minimal treatment controls [14, 25]. Participants in no-treatment controls have effectively been told that they will receive no active treatment and are more likely to engage other forms of help seeking behaviors, which may provide some benefit. In contrast, participants assigned to WLCs, expecting treatment at a future date, may be less likely to engage in help seeking. While these findings do not confirm these specific interpretations, they suggest that different control condition designs may have unexpected effects on patient motivations, behaviors, and expectations, which in turn can affect outcomes.

Effect sizes for pill placebo control arms did not differ significantly from non-specific and specific factor controls. While this suggests that non-specific controls, sometimes considered “psychological placebos,” are equivalent to pill placebo, it should be noted that most pill placebo control conditions also contain case management that includes supportive counseling, often every 1–2 weeks for 20–45 min.

It is possible, and indeed likely, that control conditions may have nonstudy effects outside of the trial that influence outcomes. TAU conditions can increase nonstudy care by promoting nonstudy physician visits or more aggressive nonstudy medical treatment, even when the intervention is not designed or intended to do so [12]. These effects can occur through a variety of mechanisms, such as assessment procedures that identify problems that result in increased care, or participant expectations related to being in a trial that affect requests and receipt of care, resulting in increased care in TAU relative to normal practice or to the experimental treatment itself [12, 31]. Control conditions may also be associated with decreases in nonstudy care. For example, in a three-arm RCT comparing a behavioral intervention for peripheral artery disease to TAU and a non-specific attention control, the non-specific control showed significantly lower use of medications relative to TAU, suggesting that providing attention may decrease help-seeking [21].

Recommendations that treatment fidelity procedures such as manualization, training, supervision, and fidelity monitoring be implemented in trials of behavioral and psychological interventions have been widely adopted [2], at least for the experimental treatments. However, we found

that many trials do not implement these procedures for control arms. Significantly larger effect sizes were observed for trials with control arms that had no treatment manuals, no therapist training or supervision, or no treatment fidelity monitoring. We note that the data for therapist supervision and fidelity monitoring is overlapping as many, but not all, trials use supervision ratings for fidelity monitoring. Given that many studies used subsets of these implementation procedures, we cannot say with certainty that each of these implementation procedures exerts a unique influence. Metaregression to tease apart the individual effects was not feasible, given the small number of comparisons in some of the cells. Nevertheless, taken together, these findings indicate that providing clinicians with support in maintaining fidelity enhances the effect of treatment on outcomes, and that failure to keep treatment implementation procedures constant across treatment arms can inflate RCT effect sizes.

In contrast to treatment implementation procedures, there was no evidence that discrepancies in the clinician experience or expertise across experimental and control arms impacted depression outcomes. We note that there are only four comparisons that contained clinicians of differing levels of experience, which limits our confidence in this analysis. However, it is consistent with a large literature showing therapist experience level has little effect on psychotherapy outcomes [3]. Overall, these findings suggest that how control treatments are implemented may be more important contributors to effect size than who implements them.

Over the past decades, methodological recommendations have tended to favor nesting therapists in treatment arms due to concerns that crossing therapists with treatment arms could lead to systematic biases resulting from therapist knowledge of the hypotheses of incremental biases resulting from therapists’ experience of one treatment being superior to the other [17]. Such biases, it was worried, would lead to larger effect sizes. This meta-analysis found no support for this idea.

These findings should not be interpreted to suggest that any of these control condition types or implementation methods are better or worse. RCT design usually involves many decisions and tradeoffs; these findings can provide information that can be considered in designing RCTs. A framework for making decisions on control arm design and implementation has previously been proposed, which identifies three factors that should be considered together when designing control conditions and their implementation: (1) considerations of statistical power and threats to internal validity, (2) RCT phase, and (3) the

interests of stakeholders (e.g., patients, patient families, clinicians, payers, and researchers) [20].

These findings provide some support for the hypothesis that greater control over threats to internal validity, by having stronger control arms that provide study treatment and implementation procedures that are consistent across trial arms, generally produce smaller effect sizes. Thus, under most circumstances, there is a trade-off between power and control, such that greater control over threats to internal validity decreases the expected effect size, thereby decreasing power.

Traditional validation criteria for a psychological intervention requires a series of trials, commonly moving from early phase I developmental phases of treatment to phase II (early phase efficacy under highly controlled conditions), phase III (late-phase efficacy which may include patient samples with more comorbidities and sometimes multiple sites), and phase IV (effectiveness or pragmatic) trials ([13]; [23]). Earlier phase trials are focused on efficacy, establishing safety and impact on primary outcome(s), minimizing the effects for many other possible determinants such as therapist variability or patient comorbidities. Later phases shift the emphasis towards generalizability. Pragmatic trials (phase IV) address questions of external validity, policy implications, and the usefulness of the intervention under real-world circumstances [13]. As such, these trials are usually employ community or standard alternative interventions, such as TAU or active comparators, and are conducted under real world conditions that can include a wider range of patients and comorbidities, a wider range of real-world clinical settings, and clinicians with varying level of expertise, instruction or training that is less intensive and more flexible, and little or no supervision [28].

As an intervention moves through these phases of evaluation, the sources of potential harm to stakeholders shift. During phase I developmental RCTs, the threat and potential harm to stakeholders from type II error (i.e., when an RCT fails to find a significant effect for an intervention that in fact does have an effect) usually exceeds the threat from type I error (i.e., when a study or RCT finds a significant effect for an intervention that in fact does not have an effect). This is because the potential long-term harm in excluding a possibly useful treatment early in the developmental process could be substantial, as it could eliminate a treatment that might have been beneficial to a population. The potential harm during these early phases of investigation resulting from finding a significant effect when in fact none exists is comparatively small, as

subsequent validation trials would provide additional opportunities to weed out the ineffective treatments. Requiring refined controls, including control therapists, manualization, therapist training, and supervision can result in costs that threaten the feasibility of early phase trials. It may be both prudent and acceptable to sacrifice some level of control over threats to internal validity for the benefit of feasibility and having adequate power to ensure any potential effects are detected. On the other hand, in later clinical trial phases the threat or potential harm shifts from type II to type I error, as data are used for guidelines and policy decisions and stakeholders must be protected from ineffective or harmful treatments. The importance of control for threats to internal validity increases in later stage trials, while concerns regarding power generally decrease.

Thus, information on the impact of control arm design must be considered judiciously. Understanding the impact of control arm design and implementation features on RCT outcomes in the context of a framework for making RCT design decisions will improve our ability to design trials that meet the aims of the study, promote innovation, and establish a knowledge base that can more effectively inform public policy, and protect the interests of stakeholders.

There are a number of limitations and caveats to the present study, which should be considered in interpreting the data. (1) The comparisons of control arm definitions and implementation methods were across trials; few trials randomized patients to multiple control arms or had different methods of implementing the same control treatments. Thus, these findings should be interpreted cautiously as support for our hypotheses, but not confirmation. (2) Consistent with most large meta-analyses of psychotherapy trials, there was an indication of substantial publication bias [9]. Generally, trials are not published due to lack of an effect for the experimental treatment relative to the control condition. This would mean that trials in which control arms produced effect sizes similar to experimental treatments may be underrepresented. (3) As noted above, there is colinearity among the variables investigated (e.g., some but not all of the trials that did not have a control treatment manual may not have trained, supervised or monitored their control arm therapists). Because a metaregression was not possible, we cannot determine the specific effects of one variable independent of the others. (4) Because some control conditions were not used with some forms of psychotherapy, it was not possible to conduct analyses that controlled for the type of psychotherapy. While we saw no overall effect for type of psychotherapy, we

cannot rule out the possibility that some of the observed effects were the result of uneven matching of control condition type and form of psychotherapy. (5) We conducted 8 analyses, which could result in alpha slippage. We used a criterion of $p < 0.01$ to balance the threats of type I and type II error. However, depending on tolerance for type I vs. type II error, investigators may choose to interpret statistical significance differently. (6) Due to a recordkeeping error in the web-based data system, we were unable to calculate reliabilities among the data extractors. (7) We did not attempt to rate the quality of the control implementation procedures. Thus, these findings cannot be interpreted as indicating that differences in the quality of control treatment implementation procedures affects outcomes. (8) The requirement that control treatment implementation procedures be implemented in the experimental procedure resulted in small reductions in the number comparisons in those analyses. While we believe this decision was warranted, as it focused the analysis on the control implementation procedures relative to the recommended implementation methodologies for experimental treatments [2], the changing sample across analyses limits the validity of comparisons of effect sizes between analyses. (9) We only examined RCTs for psychological treatments of depression. Depression has a high spontaneous remission rate and is highly responsive to placebos [22]. Thus, these findings may not be generalizable to RCTs in other disorders or behavioral targets.

In summary, decisions on the design and implementation of control arms have a statistically significant and large effect on the outcomes of trials. Indeed, the effect size of control condition design and implementation is much larger across trials than the effect of the experimental treatments themselves. While some aspects of the implementation of treatment procedures such as clinician experience and nesting vs. crossing clinicians with treatment arms may not have substantial effects on outcomes, differences in the implementation of treatment fidelity procedures across treatment arms may significantly inflate effect sizes and should be avoided. With the exception of maintaining consistency in treatment fidelity procedures across treatment arms, we do not believe these data should be used rigidly to always recommend implementation of the strongest level of control, as this is neither feasible nor desirable. Rather, we argue that this information should be used in the context of a decision making framework that considers the needs of promoting innovation in early phase trials, while protecting stakeholders in later phase trials [20]. Understanding the effects of control arm design and implementation considerations in the context of a

decision-making framework will improve the quality of the evidence base for psychological and behavioral treatments.

Acknowledgements: This meta-analysis was conducted in part through the Society of Behavioral Medicine's Evidence Based Behavioral Medicine Committee. We would like to thank Alfred Rademaker, Ph.D. and Mary Kwasny, Sc.D. for their guidance in the statistical analyses. This work was funded in part by research grant R01-MH095753 and R01-MH059708 from the National Institute of Mental Health.

APPENDIX 1

Search strategy

MEDLINE search strategy

1. meta analysis.pt.
2. systematic review.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
3. meta-analysis.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
4. metaanalysis.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
5. meta analysis.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
6. systematic literature review.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
7. quantitative review.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
8. randomized controlled trial.mp.
9. randomized controlled trial.pt.
10. controlled clinical trial.pt.
11. clinical trial.pt.
12. ((efficacy or effectiveness or intervention) adj10 (study or studies or trial\$)).mp.
13. (random\$ adj10 (trial\$ or assign\$ or allocat\$)).mp.
14. or/1-13
15. autogenic training/ or behavior therapy/ or cognitive therapy/ or "biofeedback (psychology)" or gestalt therapy/ or nondirective therapy/ or psychoanalytic therapy/ or psychotherapy, brief/ or psychotherapy, multiple/ or psychotherapy, rational-emotive/ or reality therapy/
16. ((psychologic\$ or psychodynamic) adj2 (interven\$ or treat\$ or therap\$)).ti,ab,tw.
17. ((cognit\$ or behavior\$) adj2 (treat\$ or interven\$ or therap\$)).ti,ab,tw.
18. 15 or 16 or 17
19. mood disorders/ or exp affective disorders, psychotic/ or exp depressive disorder/
20. Depression/
21. (depress\$ adj2 major).ti,ab,tw.
22. 19 or 20 or 21

23. 14 and 18 and 22
24. limit 23 to english language
25. limit 24 to "all adult (19 plus years)"

APPENDIX 2

Citations for included trials

Alladin A, Alibhai A. Cognitive hypnotherapy for depression: an empirical investigation. *International Journal of Clinical & Experimental Hypnosis*. 2007, 55:147-166.

Allart-van Dam E, Hosman CM, Hoogduin CA, Schaap CP. The Coping With Depression course: Short-term outcomes and mediating effects of a randomized controlled trial in the treatment of subclinical depression. *Behavior Therapy*. 2003, 34:381-396.

Appleby L, Warner R, Whitton A, Faragher B. A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. *BMJ (Clinical research ed.)*. 1997, 314:932-936.

Arean PA, Perri MG, Nezu AM, et al. Comparative effectiveness of social problem solving therapy and reminiscence therapy as treatments for depression in older adults. *Journal of Consulting and Clinical Psychology*. 1993, 61:1003-1010.

Barnhofer T, Crane C, Hargus E, et al. Mindfulness-based cognitive therapy as a treatment for chronic depression: A preliminary study. *Behaviour Research and Therapy*. 2009, 47:366-373.

Barrera M. An evaluation of a brief group therapy for depression. *Journal of Consulting and Clinical Psychology*. 1979, 47:413-415.

Barrett JE, Williams Jr JW, Oxman TE, et al. Treatment of dysthymia and minor depression in primary care: A randomized trial in patients aged 18 to 5 years. *Journal of Family Practice*. 2001, 50:405-412.

Barth J, Paul J, Harter M, Bengel J. Inpatient psychotherapeutic treatment for cardiac patients with depression in Germany: Short-term results. *GMS Psycho-Social-Medicine Vol 2* 2005, 1-8. 2005.

Beach SR, O'Leary KD. Treating depression in the context of marital discord: Outcome and predictors of response of marital therapy versus cognitive therapy. *Behavior Therapy*. 1992, 23:507-528.

Beutler LE, Engle D, Mohr D, et al. Predictors of differential response to cognitive, experiential, and self-directed psychotherapeutic procedures. *Journal of Consulting & Clinical Psychology*. 1991, 59:333-340.

Bodenmann G, Plancherel B, Beach SRH, et al. Effects of coping-oriented couples therapy on depression: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*. 2008, 76:944-954.

Bolton P, Bass J, Neugebauer R, et al. Group interpersonal psychotherapy for depression in rural Uganda: A randomized controlled trial. *Journal of the American Medical Association*. 2003, 289:3117-3124.

Bright JI, Baker KD, Neimeyer RA. Professional and paraprofessional group treatments for depression: a comparison of cognitive-behavioral and mutual support interventions. *Journal of Consulting & Clinical Psychology*. 1999, 67:491-501.

Brown RA, Lewinsohn PM. A psychoeducational approach to the treatment of depression: comparison of group, individual, and minimal contact procedures. *Journal of Consulting & Clinical Psychology*. 1984, 52:774-783.

Burns A, Banerjee S, Morris J, et al. Treatment and prevention of depression after surgery for hip fracture in older people: randomized, controlled trials.[see comment]. *Journal of the American Geriatrics Society*. 2007, 55:75-80.

Butler LD, Waelde LC, Hastings TA, et al. Meditation with yoga, group therapy with hypnosis, and psychoeducation for long-term depressed mood: A randomized pilot trial. *Journal of Clinical Psychology*. 2008, 64:806-820.

Chabrol H, Teissedre F, Saint-Jean M, et al. Prevention and treatment of post-partum depression: a controlled randomized study on women at risk. *Psychological Medicine*. 2002, 32:1039-1047.

Chen CH, Tseng YF, Chou FH, Wang SY. Effects of support group intervention in postnatally distressed women. A controlled study in Taiwan. *Journal of Psychosomatic Research*. 2000, 49:395-399.

Chesney MA, Chambers DB, Taylor JM, Johnson LM, Folkman S. Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. *Psychosomatic Medicine*. 2003, 65:1038-1046.

Cho HJ, Kwon JH, Lee JJ. Antenatal cognitive-behavioral therapy for prevention of postpartum depression: a pilot study. *Yonsei Medical Journal*. 2008, 49:553-562.

Ciechanowski P, Wagner E, Schmalting K, et al. Community-integrated home-based depression treatment in older adults: A randomized controlled trial. *JAMA*. 2004, 291:1569-1577.

Clark R, Tluczek A, Brown R. A mother-infant therapy group model for postpartum depression. *Infant Mental Health Journal*. 2008, 29:514-536.

Constantino MJ, Marnell ME, Haile AJ, et al. Integrative cognitive therapy for depression: A randomized pilot comparison. *Psychotherapy: Theory, Research, Practice, Training*. 2008, 45:122-134.

Czajkowski SM. Effects of treating depression and low perceived social support on clinical events after myocardial infarction: The Enhancing Recovery in Coronary Heart Disease Patients (ENRICH) randomized trial. *Journal of the American Medical Association*. 2003, 289:3106-3116.

David D, Szentagotai A, Lupu V, Cosman D. 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. *Journal of Clinical Psychology*. 2008, 64:728-746.

- Dimidjian S, Hollon SD, Dobson KS, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting & Clinical Psychology*. 2006, 74:658–670.
- Dowrick C, Dunn G, Ayuso-Mateos JL, et al. Problem solving treatment and group psychoeducation for depression: Multicentre randomised controlled trial. *BMJ: British Medical Journal*. 2000, 321:No Pagination Specified.
- Dunn NJ, Rehm LP, Schillaci J, et al. A randomized trial of self-management and psychoeducational group therapies for comorbid chronic posttraumatic stress disorder and depressive disorder. *Journal of traumatic stress*. 2007, 20:221–237.
- Elkin I, Shea MT, Watkins JT, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry*. 1989, 46:971–982; discussion 983.
- Emanuel-Zuurveen L, Emmelkamp PM. Individual behavioural-cognitive therapy v. marital therapy for depression in maritally distressed couples. *British Journal of Psychiatry*. 1996, 169:181–188.
- Emanuel-Zuurveen L, Emmelkamp PM. Spouse-aided therapy with depressed patients. *Behavior Modification*. 1997, 21:62–77.
- Evans RL CR. Comparison of brief group therapies for depressed cancer patients receiving radiation treatment. *Public Health Reports*. 1995, 110(3):298–300.
- Faramarzi M, Alipor A, Esmaelzadeh S, et al. Treatment of depression and anxiety in infertile women: Cognitive behavioral therapy versus fluoxetine. *Journal of Affective Disorders*. 2008, 108:159–164.
- Floyd MS, F. Mckendree-Smith, N. L. Floyd, D. L. Rokke, P. D. Cognitive therapy for depression: A comparison of individual psychotherapy and bibliotherapy for depressed older adults. *Behavior Modification*. 2004, 28:297–318.
- Freedland KE, Skala JA, Carney RM, et al. Treatment of depression after coronary artery bypass surgery: a randomized controlled trial. *Archives of General Psychiatry*. 2009, 66:387–396.
- Fremont J, Craighead LW. Aerobic exercise and cognitive therapy in the treatment of dysphoric moods. *Cognitive Therapy and Research*. 1987, 11:241–251.
- Fry PS. Structured and unstructured reminiscence training and depression among the elderly. *Clinical Gerontologist*. 1983, Vol 1:15–37.
- Gallagher DET, Larry W. Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy: Theory, Research & Practice*. 1982, Vol 19:482–490.
- Gellis ZD, McGinty J, Horowitz A, Bruce ML, Misener E. Problem-solving therapy for late-life depression in home care: a randomized field trial. *The American Journal of Geriatric Psychiatry*. 2007, 15:968–978.
- Goldman RN, Greenberg LS, Angus L. The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. *Psychotherapy Research*. 2006, 16:536–546.
- González, S. G., Rodríguez, C. F., Rodríguez, J. P., & Amigo, I. Secondary prevention of depression in primary care. *Psychology in Spain*. 2007, 11:24–32.
- Gordon VC, Matwychuk AC, Sachs EG, Canedy BH. A 3-yr follow-up of a cognitive-behavioral therapy intervention. *Archives of Psychiatric Nursing*. 1988, 2:218–226.
- Greenberg LS, Watson J. Experiential therapy of depression: Differential effects of client-centered relationship conditions and process experiential interventions. *Psychotherapy Research*. 1998, 8:210–224.
- Harley R, Sprich S, Safren S, Jacobo M, Fava M. Adaptation of dialectical behavior therapy skills training group for treatment-resistant depression. *Journal of Nervous and Mental Disease*. 2008, 196:136–143.
- Hopko DR, Lejuez C, LePage JP, Hopko SD, McNeil DW. A brief behavioral activation treatment for depression: A randomized pilot trial within an inpatient psychiatric hospital. *Behavior Modification*. 2003, 27:458–469.
- Hyer L, Yeager CA, Hilton N, Sacks A. Group, individual, and staff therapy: An efficient and effective cognitive behavioral therapy in long-term care. *American Journal of Alzheimer's Disease & Other Dementias*. 2008, 23:528–539.
- Jacobson NS, Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. A component analysis of cognitive-behavioral treatment for depression. *A Journal of Consulting and Clinical Psychology*. 1996, 64:295–304.
- Jacobson NS, Dobson K, Fruzzetti AE, Schmalings KB, Salusky S. Marital therapy as a treatment for depression. *Journal of Consulting & Clinical Psychology*. 1991, 59:547–557.
- Jarrett RB, Schaffer M, McIntire D, et al. Treatment of atypical depression With cognitive therapy or phenelzine: A double-blind, placebo-controlled trial. *Archives of General Psychiatry*. 1999, 57:1084.
- Johnson W, Ridley CR. Brief Christian and non-Christian rational-emotive therapy with depressed Christian clients: An exploratory study. *Counseling and Values*. 1992, 36:220–229.
- Kelly JA, Murphy DA, Bahr G, Kalichman SC, et al. Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. *American Journal of Psychiatry*. 1993, 150:1679–1686.
- Kingston T, Dooley B, Bates A, Lawlor E, Malone K. Mindfulness-based cognitive therapy for residual depressive symptoms. *Psychology & Psychotherapy: Theory, Research & Practice*. 2007, 80:193–203.
- Kornblith SJ, Rehm LP, O'Hara MW, Lamparski DM. The contribution of self-reinforcement training and behavioral assignments to the efficacy of self-

control therapy for depression. *Cognitive Therapy and Research*. 1983, 7:499–527.

Laidlaw K, Davidson K, Toner H, et al. A randomised controlled trial of cognitive behaviour therapy vs treatment as usual in the treatment of mild to moderate late life depression. *International Journal of Geriatric Psychiatry*. 2008, 23:843–850.

Lang AJ. Brief intervention for co-occurring anxiety and depression in primary care: a pilot study. *International Journal of Psychiatry in Medicine*. 2003, 33:141–154.

Larcombe NA, Wilson PH. An evaluation of cognitive-behaviour therapy for depression in patients with multiple sclerosis. *British Journal of Psychiatry*. 1984, 145:366–371.

Latour D, Cappeliez P. Pretherapy training for group cognitive therapy with depressed older adults. *Canadian Journal on Aging*. 1994, 13:221–235.

Lesperance F, Frasur-Smith N, Koszycki D, et al. Effects of citalopram and interpersonal psychotherapy on depression in patients with coronary artery disease: The Canadian Cardiac Randomized Evaluation of Antidepressant and Psychotherapy Efficacy (CREATE) trial. *Journal of the American Medical Association*. 2007, 297:367–379.

Lustman PJ, Griffith LS, Freedland KE, Kissel SS, Clouse RE. Cognitive behavior therapy for depression in type 2 diabetes mellitus. A randomized, controlled trial. *Annals of internal medicine*. 1998, 129:613–621.

Luty SE, Carter JD, McKenzie JM, et al. Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression. *British Journal of Psychiatry Vol 190 Jun 2007*, 496–502. 2007.

Markowitz JC, Kocsis JH, Bleiberg KL, Christos PJ, Sacks M. A comparative trial of psychotherapy and pharmacotherapy for “pure” dysthymic patients. *Journal of Affective Disorders*. 2005, 89:167–175.

Markowitz JC, Kocsis JH, Fishman B, et al. Treatment of depressive symptoms in human immunodeficiency virus-positive patients. *Archives of General Psychiatry*. 1998, 55:452–457.

Meeks S, Looney SW, Van Haitsma K, Teri L. BE-ACTIV: A staff-assisted behavioral intervention for depression in nursing homes. *Gerontologist*. 2008, 48:105–114.

Milgrom J, Negri LM, Gemmill AW, McNeil M, Martin PR. A randomized controlled trial of psychological interventions for postnatal depression. *British Journal of Clinical Psychology*. 2005, 44:529–542.

Miller IW, Norman WH, Keitner GI. Cognitive-behavioral treatment of depressed inpatients: Six- and twelve-month follow-up. *American Journal of Psychiatry*. 1989, 146:1274–1279.

Misri S, Kostaras X, Fox D, Kostaras D. The impact of partner support in the treatment of postpartum depression. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*. 2000, 45:554–558.

Mohr DC, Boudewyn AC, Goodkin DE, Bostrom A, Epstein L. 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*. 2001, 69:942–949.

Mohr DC, Hart SL, Julian L, et al. Telephone-administered psychotherapy for depression. *Archives of General Psychiatry*. 2005, 62:1007–1014.

Mohr DC, Likosky W, Bertagnolli A, et al. Telephone-administered cognitive-behavioral therapy for the treatment of depressive symptoms in multiple sclerosis. *Journal of Consulting & Clinical Psychology*. 2000, 68:356–361.

Mynors-Wallis LM, Gath DH, Lloyd-Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *British Medical Journal*. 1995, 310:441–445.

Nezu AM. Efficacy of a social problem-solving therapy approach for unipolar depression. *Journal of Consulting and Clinical Psychology*. 1986, 54:196–202.

Nezu AM, Perri MG. Social problem-solving therapy for unipolar depression: an initial dismantling investigation. *Journal of Consulting & Clinical Psychology*. 1989, 57:408–413.

O'Hara MW, Stuart S, Gorman LL, Wenzel A. Efficacy of interpersonal psychotherapy for postpartum depression. *Archives of General Psychiatry*. 2000, 57:1039–1045.

Oxman TE, Hegel MT, Hull JG, Dietrich AJ. Problem-solving treatment and coping styles in primary care for minor depression. *Journal of Consulting and Clinical Psychology*. 2008, 76:933–943.

Pace TM, Dixon DN. Changes in depressive self-schemata and depressive symptoms following cognitive therapy. *Journal of Counseling Psychology*. 1993, 40:288–294.

Parker JC, Smarr KL, Slaughter JR, et al. Management of depression in rheumatoid arthritis: a combined pharmacologic and cognitive-behavioral approach. *Arthritis & Rheumatism*. 2003, 49:766–777.

Pecher DR, Edwards KJ. A comparison of secular and religious versions of cognitive therapy with depressed Christian college students. *Journal of Psychology and Theology*. 1984, 12:45–54.

Prendergast J, Austin M-P. Early childhood nurse-delivered cognitive behavioural counselling for postnatal depression. *Australasian Psychiatry*. 2001, 9:255–259.

Propst LR, Ostrom R, Watkins P, Dean T, Mashburn D. Comparative efficacy of religious and nonreligious cognitive-behavioral therapy for the treatment of clinical depression in religious individuals. *Journal of Consulting & Clinical Psychology*. 1992, 60:94–103.

Ransom D, Heckman TG, Anderson T, et al. Telephone-delivered, interpersonal psychotherapy for HIV-infected rural persons with depres-

- sion: A pilot trial. *Psychiatric Services*. 2008, 59:871–877.
- Ravindran AV, Anisman H, Merali Z, et al. Treatment of primary dysthymia with group cognitive therapy and pharmacotherapy: clinical symptoms and functional impairments. *American Journal of Psychiatry*. 1999, 156:1608–1617.
- Rehm LP, Kornblith SJ, O'Hara MW, et al. An evaluation of major components in a self-control therapy program for depression. *Behavior Modification*. 1981, 5:459–489.
- Richards DA, Lovell K, Gilbody S, et al. Collaborative care for depression in UK primary care: A randomized controlled trial. *Psychological Medicine*. 2008, 38:279–287.
- Rohan KJ, Roecklein KA, Lindsey KT, et al. A randomized controlled trial of cognitive-behavioral therapy, light therapy, and their combination for Seasonal Affective Disorder. *Journal of Consulting and Clinical Psychology*. 2007, 75:489–500.
- Rude SS. Relative benefits of assertion or cognitive self-control treatment for depression as a function of proficiency in each domain. *Journal of Consulting & Clinical Psychology*. 1986, 54:390–394.
- Safren SA, O'Cleirigh C, Tan JY, et al. A randomized controlled trial of cognitive behavioral therapy for adherence and depression (CBT-AD) in HIV-infected individuals. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*. 2009, 28:1–10.
- Sallis JF, et al. Anxiety and depression management for the elderly. *International Journal of Behavioral Geriatrics*. 1983, 1:3–12.
- Sanders MR, McFarland M. Treatment of depressed mothers with disruptive children: A controlled evaluation of cognitive behavioral family intervention. *Behavior Therapy*. 2000, 31:89–112.
- Savard J, Simard S, Giguere I, et al. Randomized clinical trial on cognitive therapy for depression in women with metastatic breast cancer: psychological and immunological effects. *Palliative & Supportive Care*. 2006, 4:219–237.
- Schmidt MM, Miller WR. Amount of therapist contact and outcome in a multidimensional depression treatment program. *Acta Psychiatrica Scandinavica*. 1983, 67:319–332.
- Scott C, Tacchi J, Jones R, Scott J. Acute and one-year outcome of a randomised controlled trial of brief cognitive therapy for major depressive disorder in primary care. *British Journal of Psychiatry*. 1997, 171:131–134.
- Seligman ME, Rashid T, Parks AC. Positive psychotherapy. *American Psychologist*. 2006, 61:774–788.
- Selmi PM, Klein MH, Greist JH, Sorrell SP, Erdman HP. Computer-administered cognitive-behavioral therapy for depression. *American Journal of Psychiatry*. 1990, 147:51–56.
- Serrano JP, Latorre JM, Gatz M, Montanes J. Life review therapy using autobiographical retrieval practice for older adults with depressive symptomatology. *Psychology and Aging*. 2004, 19:272–277.
- Shapiro DA, Barkham M, Rees A, et al. Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology*. 1994, 62:522–534.
- Simon GE, Ludman EJ, Tutty S, Operskalski B, Von Korff M. Telephone psychotherapy and telephone care management for primary care patients starting antidepressant treatment: A randomized controlled trial. *JAMA: Journal of the American Medical Association*. 2004, 292:935–942.
- Simpson S, Corney R, Fitzgerald P, Beecham J. A randomized controlled trial to evaluate the effectiveness and cost-effectiveness of psychodynamic counselling for general practice patients with chronic depression. *Psychological Medicine*. 2003, 33:229–239.
- Simson U, Nawarotzky U, Frieze G, et al. Psychotherapy intervention to reduce depressive symptoms in patients with diabetic foot syndrome. *Diabetic medicine : a journal of the British Diabetic Association*. 2008, 25:206–212.
- Spek V, Nykek I, Smits N, et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. *Psychological Medicine*. 2007, 37:1797–1806.
- Spinelli MG, Endicott J. Controlled clinical trial of interpersonal psychotherapy versus parenting education program for depressed pregnant women. *The American journal of psychiatry*. 2003, 160:555–562.
- Strauman TJ, Vieth AZ, Merrill KA, et al. Self-system therapy as an intervention for self-regulatory dysfunction in depression: a randomized comparison with cognitive therapy. *Journal of Consulting & Clinical Psychology*. 2006, 74:367–376.
- Swartz HA, Frank E, Zuckoff A, et al. Brief interpersonal psychotherapy for depressed mothers whose children are receiving psychiatric treatment. *American Journal of Psychiatry*. 2008, 165:1155–1162.
- Taylor FG, Marshall WL. Experimental analysis of a cognitive-behavioral therapy for depression. *Cognitive Therapy and Research*. 1977, 1:59–72.
- Teichman Y, Bar-el Z, Shor H, Sirota P, Elizur A. A comparison of two modalities of cognitive therapy (individual and marital) in treating depression. *Psychiatry*. 1995, 58:136–148.
- Thomas J, Petry RA, Goldman JR. Comparison of cognitive and behavioral self-control treatments of depression. *Psychological Reports*. 1987, 60:975–982.
- Thompson LWG, Dolores. Efficacy of psychotherapy in the treatment of late-life depression. *Advances in Behaviour Research & Therapy*. 1984, Vol 6:127–139.
- Thyme KE, Sundin EC, Stahlberg G, et al. The outcome of short-term psychodynamic art therapy compared to short-term psychodynamic verbal therapy for depressed women. *Psychoanalytic Psychotherapy*. 2007, 21:250–264.

Tsai YF, Wong TK, Tsai HH, Ku YC. Self-worth therapy for depressive symptoms in older nursing home residents. *Journal of advanced nursing*. 2008, 64:488–494.

Usaf SO, Kavanagh DJ. Mechanisms of improvement in treatment for depression: Test of a self-efficacy and performance model. *Journal of Cognitive Psychotherapy*. 1990, 4:51–70.

van den Hout JH, Arntz A, Kunkels FH. Efficacy of a self-control therapy program in a psychiatric day-treatment center. *Acta Psychiatrica Scandinavica*. 1995, 92:25–29.

van Schaik AvM, H.; Ader, H.; van Dyck, R.; Haan, M.; Penninx, B.; Kooij, K.; von Hout, H.; Beekman, A. Interpersonal psychotherapy for elderly patients in primary care. *American Journal of Geriatric Psychiatry*. 2006, 14:777–786.

Verduyn C, Barrowclough C, Roberts J, Tarrier T, Harrington R. Maternal depression and child behaviour problems. Randomised placebo-controlled trial of a cognitive-behavioural group intervention. *British Journal of Psychiatry*. 2003, 183:342–348.

Ward E, King M, Lloyd M, et al. Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: clinical effectiveness.[see comment]. *BMJ*. 2000, 321:1383–1388.

Watson JC, Gordon LB, Stermac L, Kalogerakos F, Steckley P. Comparing the effectiveness of process-experiential with cognitive-behavioral psychotherapy in the treatment of depression. *Journal of Consulting & Clinical Psychology*. 2003, 71:773–781.

Watt LMC, P. Integrative and instrumental reminiscence therapies for depression in older adults: Intervention strategies and treatment effectiveness. *Aging & Mental Health*. May 2000, Vol 4:166–177.

Wickberg B, Hwang CP. Counselling of postnatal depression: A controlled study on a population based Swedish sample. *Journal of Affective Disorders*. 1996, 39:209–216.

Williams JW, Jr., Barrett J, Oxman T, et al. Treatment of dysthymia and minor depression in primary care: A randomized controlled trial in older adults. *JAMA: Journal of the American Medical Association*. 2000, 284:1519–1526.

Wilson GL. Psychotherapy with depressed incarcerated felons: a comparative evaluation of treatments. *Psychological Reports*. 1990, 67:1027–1041.

Wilson PH. Combined pharmacological and behavioural treatment of depression. *Behaviour Research and Therapy*. 1982., 20:173–184.

Wollersheim JP, Wilson GL. Group treatment of unipolar depression: A comparison of coping, supportive, bibliotherapy, and delayed treatment groups. *Professional Psychology: Research and Practice*. 1991, 22:496–502.

Wong DFK. Cognitive and health-related outcomes of group cognitive behavioural treatment for people with depressive symptoms in Hong Kong: randomized wait-list control study. *Austra-*

lian & New Zealand Journal of Psychiatry. 2008, 42:702–711.

Zettle RD, Rains JC. Group cognitive and contextual therapies in treatment of depression. *Journal of Clinical Psychology*. 1989, 45:436–445.

APPENDIX 3

Reviews that were examined in the course of the search

Ackermann RT, Williams JW, Jr. Rational Treatment Choices for Non-major Depressions in Primary Care: An Evidence-based Review. *Journal of General Internal Medicine*. 2002, 17:293–301.

Adamek ME, Slater GY. Depression and anxiety. *Journal of Gerontological Social Work*. 2008, 50 Suppl 1:153–189.

Akechi T, Okuyama T, Onishi J, Morita T, Furukawa TA: Psychotherapy for depression among incurable cancer patients. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd, 2008.

Barbato A, D'Avanzo B: Marital therapy for depression. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd, 2006.

Baskin TW, Tierney, S.C., Minami, T., Wampold, B.E., . Establishing specificity in psychotherapy: ameta-analysis of structural equivalence of placebo controls. *J. Consult. Clin Psychology*. 2003, 71:973–979.

Coelho HF, Canter PH, Ernst E. Mindfulness-based cognitive therapy: evaluating current evidence and informing future research. *Journal of Consulting & Clinical Psychology*. 2007, 75:1000–1005.

Conte HR, Plutchik R, Wild KV, Karasu TB. Combined psychotherapy and pharmacotherapy for depression. A systematic analysis of the evidence. *Archives of General Psychiatry*. 1986, 43:471–479.

Cuijpers P, van Straten A, Andersson G, van Oppen. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *J. Consult. Clin Psychology*. 2008, 76:909–922.

Cuijpers P, van Straten A, Smit F. Psychological treatment of late-life depression: a meta-analysis of randomized controlled trials. *International Journal of Geriatric Psychiatry*. 2006, 21:1139–1149.

Gildengers AG, Houck PR, Mulsant BH, et al. Course and rate of antidepressant response in the very old. *Journal of Affective Disorders*. 2002, 69:177–184.

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

Henken HT, Huibers MJH, Churchill R, Restifo K, Roelofs J: Family therapy for depression. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd, 2007.

Iardi S, Craighead W. The role of non specific factors in cognitive-behavioral therapy for depres-

sion. *Clinical Psychology: Science and Practice*. 1994; 1:138–156.

Pampallona S, Bollini P, Tibaldi G, Kupelnick B, Munizza C. Combined pharmacotherapy and psychological treatment for depression: A systematic review. *Archives of General Psychiatry*. 2004; 61:714–719.

Steinbrueck SM, Maxwell SE, Howard GS. A meta-analysis of psychotherapy and drug therapy in the treatment of unipolar depression with adults. *Journal of Consulting and Clinical Psychology*. 1983; 51:856–863.

Wampold B, Minami T, Baskin T, Callen Tierney S. A meta-(re)analysis of the effects of cognitive therapy versus 'other therapies' for depression. *Journal of Affective Disorders*. 2002; 68:159–165.

Wilson KCM, Mottram PG, Vassilas CA: Psychotherapeutic treatments for older depressed people. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd, 2008.

1. Appleby L, Warner R, Whittom A, Faragher B. A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. *BMJ*. 1997;314(7085):932–936.
2. Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*. 2004;23(5):443–451.
3. Beutler LE, Malik M, Alimohamed A, Harwood TM. Therapist variables. In: Lambert MJ, ed. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 5th ed. New York: John Wiley & Sons; 2004:227–306.
4. Borenstein M, Hedges L, Higgins J, Rothstein H. *Comprehensive Meta-analysis Version 2*. Englewood, NJ: Biostat; 2005.
5. Campbell DT, Stanley JC. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally; 1966.
6. Chambless DL, Hollon SD. Defining empirically supported therapies. *J Consult Clin Psychol*. 1998;66(1):7–18.
7. Cuijpers P, Li J, Hofmann SG, Andersson G. Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clin Psychol Rev*. 2010;30(6):768–778. doi: 10.1016/j.cpr.2010.06.001.
8. Cuijpers P, van Straten A, Andersson G, van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *J Consult Clin Psychol*. 2008;76(6):909–922.
9. Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychol Med*. 2010;40(2):211–223. doi: S0033291709006114 [pii]10.1017/S0033291709006114.
10. Cuijpers P, Van Straten A, Warmerdam L, Smits N. Characteristics of effective psychological treatments of depression: a meta-regression analysis. *Psychother Res*. 2008;18(2):225–236. doi: 10.1080/10503300701442027.
11. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455–463.
12. Freedland KE, Mohr DC, Davidson KW, Schwartz JE. Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosom Med*. 2011;73(4):323–335. doi: PSY.0b013e318218e1fb [pii]10.1097/PSY.0b013e318218e1fb.
13. Glasgow RE, Davidson KW, Dobkin PL, Ockene J, Spring B. Practical behavioral trials to advance evidence-based behavioral medicine. *Ann Behav Med*. 2006;31(1):5–13.
14. Harris KB, Miller WR. Behavioral self-control training for problem drinkers: components of efficacy. *Psychol Addict Behav*. 1990;4:82–90.
15. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
16. Ilardi SS, Craighead WE. The role of nonspecific factors in cognitive-behavior therapy for depression. *Clin Psychol Sci Pract*. 1994;1:138–156.
17. Kazdin AE. *Methodological issues and strategies in clinical research*. Washington DC: American Psychological Association; 1992.
18. Kazdin AE, Bass D, Ayers WA, Rodgers A. Empirical and clinical focus of child and adolescent psychotherapy research. *J Consult Clin Psychol*. 1990;58(6):729–740.
19. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther*. 2003;83(8):713–721.
20. Mohr DC, Spring B, Freedland KE, et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychother Psychosom*. 2009;78(5):275–284. doi: 000228248 [pii]10.1159/000228248.
21. Pagoto SL, McDermott MM, Reed G, et al. Can attention control conditions have detrimental effects on behavioral medicine randomized trials? *Psychosom Med*. 2013;75(2):137–143. doi: 10.1097/PSY.0b013e3182765dd2.
22. Rief W, Nestorciuc Y, Weiss S, Welzel E, Barsky AJ, Hofmann SG. Meta-analysis of the placebo response in antidepressant trials. *J Affect Disord*. 2009;118(1–3):1–8. doi: S0165-0327(09)00038-X [pii]10.1016/j.jad.2009.01.029.
23. Rounsaville BJ, Carroll KM, Onken LS. A stage model of behavioral therapies research: getting started and moving on from stage I. *Clin Psychol Sci Pract*. 2001;8(2):133–142.
24. SAS Institute Inc. *SAS version 9.02*. Cary, N.C.: SAS Institute, Inc.; 2007.
25. Schmidt MM, Miller WR. Amount of therapist contact and outcome in a multidimensional depression treatment program. *Acta Psychiatr Scand*. 1983;67(5):319–332.
26. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol*. 2001;54(10):1046–1055. doi: S0895-4356(01)00377-8 [pii].
27. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA*. 2000;283(15):2008–2012. doi: jst00003 [pii].
28. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62(5):464–475. doi: 10.1016/j.jclinepi.2008.12.011.
29. Wilson PH. Combined pharmacological and behavioral treatment of depression. *Behav Res Ther*. 1982;20(2):173–184. doi: 10.1016/0005-7967(82)90116-4.
30. Wollersheim JP, Wilson GL. Group treatment of unipolar depression - a comparison of coping, supportive, bibliotherapy, and delayed treatment groups. *Professional Psychology-Research and Practice*. 1991;22(6):496–502. doi: Doi 10.1037/0735-7028.22.6.496.
31. Woody GE, Poole SA, Subramaniam G, et al. Extended vs short-term buprenorphine-naloxone for treatment of opioid-addicted youth: a randomized trial. *JAMA*. 2008;300(17):2003–2011. doi: 10.1001/jama.2008.574.