# Dropout from Randomized, Controlled Treatments for Depression

Andrew A. Cooper

Dropout from Randomized, Controlled Treatments for Depression

Premature termination of psychotherapy, often referred to as "dropout", is a commonly occurring phenomenon that can have deleterious effects on both clinical practice and research. Unexpected termination can increase wasted "no-show" hours, and it may demoralize therapists and reduce their effectiveness (Barrett *et al.*., 2008; Klein, Stone, Hicks & Pritchard, 2003; Sledge, Moras, Hartley & Levine, 1990). Dropout may also pose unique challenges in a research context, as patients who fail to complete study protocols can affect analyses and outcomes, and high dropout rates may complicate interpretation of results. However, to date, dropout has not been extensively studied in randomized controlled trials (RCTs) of psychotherapies. Estimates of dropout and data on variables that predict dropout may improve the quality of psychotherapy research by informing study design and providing a more appropriate comparison for observed dropout rates. This paper uses meta-analytic techniques to provide an estimate of dropout in one common kind of psychotherapy RCT – namely, studies that involve individual treatments for depression – and to investigate potential predictors of dropout in this context.

**Psychotherapy Dropout**

Two kinds of studies dominate the existing literature on estimates of psychotherapy dropout and predictors of dropout. The first consists of large-scale epidemiological investigations of dropout across all kinds of mental health treatments, which are typically assessed using survey methods (see Edlund, Wang, Berglund, Katz, Lin & Kessler, 2002; Olfson *et al.*., 2009; Wang, 2007; Wang, Gilman, Guardino, Christiana, Morselli, Mickelson, *et al.*., 2000). The second consists of smaller-scale studies evaluating patient-level predictors, mostly conducted in the context of community mental health clinics or private practice (see Mueller & Pekarik, 2000; Oei & Kazmierczak, 1997; Persons, Burns & Perloff, 1988; Rusch *et al.*, 2008).

Estimates of the dropout rate in community treatment samples have been reported as ranging from 24% to 66% (Bados, Balaguer, & Saldana, 2007). Some patient characteristics have emerged as predictors from this research; however, results have not been consistent, perhaps in part due to considerable methodological variability in studies of this kind (Reis & Brown, 1999).

Defining dropout in a clinical context is a major issue in this domain of research, as variation in definitions can dramatically affect estimates of overall rates (see Barrett *et al.*, 2008; Reis & Brown, 1999; Wierzbicki & Pekarik, 1993). For instance, Wierzbicki and Pekarik (1993) noted a 12% difference in overall dropout rate based on whether dropout was defined as the failure to attend a scheduled session versus basing dropout on therapist judgment. This issue highlights a potentially meaningful difference between clinical practice and psychotherapy research. In clinical practice, defining dropout may be complicated due to ambiguous markers of treatment attendance and expected length; by contrast, researchers using RCTs to investigate psychotherapeutic interventions often impose structure as part of the study protocol that makes identifying dropouts more straightforward. This may make dropout in a research context somewhat different than dropout in community clinical samples, and suggests the possibility that dropout estimates might differ between these two settings.

In the context of psychotherapy RCTs, dropout has the potential to affect more than therapist effectiveness and patient care. Dropouts may unbalance group designs for data analyses and reduce power to detect effects. Conventional methods for dealing with dropout in data analyses (e.g., intent-to-treat analyses) may yield distorted results (Lane, 2008) that may be difficult to interpret, especially when dropout levels are very high in one treatment arm but not others (see, for example, Coffman, Martell, Dimidjian, Gallop & Hollon, 2007). Yet despite the

increasing number of psychotherapy RCTs being conducted, the existing dropout literature is absent an extensive investigation of the phenomenon in the context of psychotherapy RCTs.

Researchers conducting RCTs may have need for information about expected dropout rate, and may also benefit from information about any study characteristics that predict changes in dropout. This information may be critical for procurement of grants, recruitment efforts, resource management, determining study and group size, and for structuring data analyses and estimating power. In addition, such data may be relevant to researchers with projects already underway to know if they are experiencing atypical dropout rates. However, the existing literature on dropout predominantly focuses on community clinical samples and private practice, which may not be well suited to comparison to RCTs. Looking to similar research designs for a comparison may yield a wide range of dropout rates. As such, psychotherapy researchers may be without a clear precedent or estimate by which to gauge the likely rate of dropout in a given research design.

An empirically-derived estimate of dropout from psychotherapy research may be useful to investigators developing RCTs, and information about study or treatment characteristics that predict changes in dropout rate may also help inform research design and give context to observed dropout rates. As psychotherapy research is a heterogeneous field, it may be especially important for such data to incorporate a range of study designs, sizes and psychotherapy formats, so as to provide a comparison for researchers working with a multitude of protocols. A study-level analysis of dropout rates, and of predictors of dropout, may provide researchers with a more appropriate and empirically-informed method of considering how dropout is likely to affect their own RCTs.

In order to facilitate a more straightforward investigation of this issue, I will focus on a single diagnosis – major depressive disorder. This approach preserves the heterogeneity in study design that allows for the evaluation of study-level predictors, while removing much of the potential for complications arising from disorder-specific differences in treatment format and duration. Evidence of differential dropout rates by primary diagnosis (see Persons, Burns & Perloff, 1988; Stark, 1992; Wang, 2007) highlights both the need for further study on this matter, and the potential value of focusing on a single disorder for an initial study such as this. Major depressive disorder is an ideal candidate in this regard, as it has been extensively studied in a large number of clinical trials.

As RCTs vary with respect to the conditions being tested (e.g., psychotherapy versus medication, psychotherapy versus psychotherapy), evaluating predictors of dropout at the study level may prove imprecise with respect to specific features of psychotherapy conditions. As such, data on dropout will be collected at both the treatment-level (that is, a particular arm of a psychotherapy RCT), and at study level (that is, overall dropout rate). Treatment-level analyses will be restricted to active psychotherapy interventions. With this information collected, secondary analyses exploring differences in dropout rate as a function of treatment and study characteristics can be conducted.

**Study Characteristics as Predictors of Dropout**

Although many variables have been investigated as predictors of dropout, the existing literature has failed to identify any robustly consistent predictors. As such, the following variables were selected for this meta-analysis based on both empirical evidence, and their potential to be useful and easily measured variables for psychotherapy researchers: treatment type, therapist experience/credentials, intervention length, comorbid conditions (separated into

Axis I diagnoses, personality disorders, and anxiety disorders), mean age and proportion of minority patients of the study sample, socioeconomic status (SES), and effect of intervention on depressive symptoms.

Psychotherapy treatment type may be an especially compelling treatment variable in light of the long history of comparative studies in treatment of depression. In an analysis involving a subset of the studies utilized herein, Cuijpers and colleagues (Cuijpers, van Straten, Andersson & van Oppen; hereafter, Cuijpers *et al.*, 2008a) reported a slightly higher risk of dropout in studies involving cognitive-behavioral therapy, and a slightly lower risk of dropout in those involving supportive-expressive therapy; however, the authors noted that the number of samples of each therapy included in these analyses were disparate (28 for CBT vs. 5 for SE) and cautioned interpreting the finding without considering differences in definition of dropout. Nonetheless, this finding suggests that there may be some useful variability in dropout rates across different treatment types.

A number of studies have found support for the notion that therapist experience reduces risk of dropout (for a brief review, see Roth & Fonagy, 2005, p.453), though null findings have also been reported (Wampold & Brown, 2005). Therapist training and credentials have been less extensively studied, and the results have also been largely inconsistent to date (for a review, see Beutler *et al.*., 2003; Hamilton, Moore, Crane & Payne, 2010). With respect to treatment length, there is some evidence to suggest that short-term interventions tend to have lower dropout rates when contrasted with longer forms of similar interventions (Sledge *et al.*, 1990). This topic has not been studied extensively as a predictor of dropout, perhaps because of its relationship to dropout itself in treatment-as-usual clinical settings, where duration may be negotiated between client and therapist as opposed to explicitly defined by study parameters.

The preponderance of available evidence suggests that younger patients, those with lower incomes (associated with low SES) and those from minority groups are more likely to drop out (Edlund *et al.*, 2002; Wang *et al.*, 2000) although there are contradictory findings for each of these variables (Arnow *et al.*, 2007; Olfson *et al.*, 2009). There are also empirical reasons to suspect that comorbid diagnoses may increase dropout rate dropout. Olfson *et al.* (2009) reported a trend level prediction of elevated risk of dropout amongst patients with multiple psychiatric conditions. Similarly, Arnow *et al.* (2007) reported that dropouts were more likely to have an anxiety disorder than completers; this may be an especially useful comparison, as anxiety and depression commonly co-occur. In a sample of CT-treated patients, Fournier *et al.* (2008) reported a trend level interaction between treatment assignment and Axis II diagnoses in the prediction of dropout, such that people with personality disorders were more likely to drop out of cognitive therapy.

Finally, the relationship between the effect size of an intervention and the dropout rate is a complex issue that has not been extensively explored in the existing literature. Dropout has considerable potential to influence observed results from an analytic standpoint, having the potential to increase effect sizes in completer analyses and potentially dampen them in intent-to-treat analyses. As such, exploring the relationship between dropout and effect of intervention will require thoughtful consideration in the context of this meta-analysis.

**Criteria for Inclusion in Meta-Analysis**

Capturing some variability in terms of study characteristics is critical to a comprehensive effort to identify study-level predictors of dropout. However, too much variability may yield results that are difficult to interpret, so three restrictions will be added to the criteria for inclusion in this meta-analysis in order to aid the ease of interpretation. Included studies must be

investigations of (1) outpatient psychotherapy, involving (2) face-to-face individual treatment with (3) patients formally diagnosed with major depressive disorder. The rationale for each of these restrictions is addressed in turn below.

The exclusion of studies that involve non-outpatient treatment (e.g., in-patient or emergency psychiatric care) is predicated upon concerns that this population may have meaningful differences in how dropout is defined. It is possible that such patients may be less likely to be lost to follow-up in the same way that a traditional outpatient might be, by virtue of being cared for in an in-patient setting. Studies in which patients are recruited while in-patients for later outpatient treatment will be included (e.g., recruiting pregnant mothers in hospital). Interventions that are not strictly individual psychotherapy will also be excluded (i.e., group, couples, or guided bibliotherapy). There is some reason to believe that rates may differ across these modes of delivery (see, for example, Organista, Munoz & Gonzalez, 1994; Minniti *et al.*, 2010; Rush & Watkins, 1981). Furthermore, as with in-patient care, there may be differences in definition of dropout in these settings (e.g., patients may be allowed to miss more sessions of group than in individual treatments). As an extension of this, only therapies delivered during face-to-face contact will be acceptable. Finally, studies must have as inclusion criteria a formal, primary diagnosis of major depressive disorder or post-partum depression, using DSM or comparable criteria through a formal diagnostic process. As such, analogue populations and other depressive disorders (e.g., dysthymia, minor depressive disorder) will be removed. It is by no means absolutely clear that such diagnostic differences will affect dropout rate; however, one can easily imagine that they might do so in some non-random way (e.g., less severely ill patients may be less motivated to attend). Although variability in symptom severity would be present to

some degree within a formally-diagnosed sample, such effects might occur in a non-random fashion for the studies employing analogue or minor depressive populations.

**Definition of Dropout**

As noted above, the definition of dropout is a critical issue for any investigation of dropout rates. For the purpose of this study, the primary dropout variable, *treatment-level dropout*, will refer to the dropout rate of an individual psychotherapy condition that is part of a randomized controlled trial. A given study may thus contribute multiple conditions to this analysis[1]. These analyses are approached from the standpoint of understanding dropout in psychotherapy research in terms of unexpected patient loss, among individuals who were accepted into the study but who failed to complete it. For this reason, the earliest point at which a patient can be considered a dropout is post-randomization to treatment condition; patients lost prior to that point are not considered dropouts because they have not been enrolled in the study and classified as a study patient (e.g., patients who attend an assessment but never return to be randomized). Patients who are randomized but refuse their assignment, those who never attend a session, and those who stop attending sessions or withdraw consent before the end of the treatment period are considered dropouts. The second dropout variable is based on *overall dropout*, which is the global dropout rate across all conditions in particular study. Although this does not speak to the primary question of interest, it may provide useful data for comparison purposes and also to provide an estimate of overall dropout rate in RCTs involving treatments for depression. Overall dropout rates were only collected from studies that were capable of being included in the treatment-level analysis; that is, studies that did not report dropout by condition were not included in this meta-analysis.

In light of variability in treatment duration, information about dropout was only collected for the initial or acute phase of treatment in multi-stage studies. Additionally, administrative removals of study patients and instances of data loss were not treated as dropouts. The rationale for their exclusion is that these rates are often reported separately, and often result from failures of the inclusion procedures or issues related to concurrent treatment.

## Methods

In lieu of a primary literature search, this study utilizes a comprehensive database of RCTs for depression as a starting point for determining study inclusion. This database is available for public use as part of an ongoing effort to provide a comprehensive collection of articles for researchers interested in conducting meta-analyses; a description of the efforts to develop the database was offered by Cuijpers, van Straten, Warmerdam, and Andersson (2008b), and additional information can be found on the website itself (www.psychotherapyrcts.org).

As described by Cuijpers *et al.*(2008b), studies were identified in the following way: a comprehensive literature search was conducted, identifying articles from 1966 to May 2007 resulting in 6947 abstracts from PubMed, PsycINFO, EMBASE and the Cochrane Central Register of Controlled trials, based on terms indicative of psychological treatment and depression. Studies were also selected from primary meta-analyses of depression and from published dissertations. Treatments were defined as interventions where verbal communication between client and therapist was a core feature (allowing for the inclusion of multiple different modalities, including guided bibliotherapy and telephone-based interventions). Studies involving non-adult samples were removed, as were studies where the psychotherapeutic intervention could not be identified from other aspects of standard care. Studies in which an effect size could not be established were also removed from the database. As of July 2010, the

psychotherapyrcts.org database had identified 243 articles involving a randomized controlled trial of one or more psychotherapy in the treatment of depression. A comprehensive list of these articles is available on their website, alongside a database of study characteristics compiled by the authors of the website (viz., Cuijpers *et al.*, 2008b). Figure 1 summarizes the steps for generating the dataset in flowchart format.

Basic descriptive information provided on the website lead to the removal of 130 articles for failure to include a formal diagnostic test of major depression (76 articles failed this criterion) and/or for not involving an individual, outpatient psychotherapy (92 articles failed this criterion). Some studies failed to meet both criteria. Thus, a total of 113 articles were accepted for a more thorough review. An additional 67 articles were removed from the dataset for a number of reasons (outlined in Appendix 1), most commonly for providing insufficient information on dropout or for depression diagnosis issues (e.g., primary dysthymia diagnosis). One study was removed because it was based on the same sample as another study that had been included. Thus, the total number of separate studies included in the meta-analysis was 46. From this group of studies, a total of 69 individual psychotherapy conditions were acceptable for use in the primary treatment-level dropout analyses.

**Primary Data Collection**

The RCT database provides information about date of publication, target population, type of comparison (e.g., psychotherapy versus pharmacotherapy, etc.) and type of psychotherapy, divided into 7 categories (per Cuijpers *et al.*, 2008b): (a) cognitive-behavioral, (b) problem-solving, (c) interpersonal, (d) non-directive / supportive therapy, (e) behavioral activation, (f) psychodynamic, and (g) other. Target population is also provided, classified as: adults, older adults, women with post-partum depression (PPD), patients with general medical conditions, and

"other". Dropout rates and all other potential predictor variables were collected during the review process. Additional information was recorded concerning the specificity of the variables in question – that is, whether or not estimates were provided by condition, across the entire study, or only for completers. Estimates not provided by condition were only retained if the variables in question were identified as not varying meaningfully between conditions, or between completers and dropouts. Primary analyses were conducted using all available data, unless otherwise discussed.

Table 1 indicates the availability of data on each of the potential predictors in the treatment-level analysis, along with mean and weighted mean values where applicable. Some of the variables were collected in multiple formats. For instance, information about therapist credentials was collected in two ways: a categorical variable reflecting the most common degree of therapists in the study, and as a continuous estimate of mean years of experience. Similarly, information about treatment duration was collected in terms of both intended and observed duration, and in terms of weeks and number of sessions, where available. Notably, socioeconomic status was not included in Table 1 because the information was provided in less than a quarter of cases (16 of 69), and estimates were of sufficient variety and imprecision that creating an overall index proved impossible. It also bears noting that most of the data on comorbid conditions took the form of exclusion criteria; as such, the mean values of these variables may appear low, as the most common value was zero when exclusion criteria were in place. Information about these conditions was also generally less commonly available.

Information about the effect of intervention was collected by recording pre-post scores on the two most commonly utilized measures of depressive symptoms: the Beck Depression Inventory, versions I (Beck, Ward, Mendelson, Mock & Erbaugh, 1961) and II (Beck, Steer &

Brown, 1996), and the Hamilton Rating Scale for Depression (Hamilton, 1960). Cohen's *d* effect sizes were calculated for these variables. Condition-level mean values that were explicitly identified as coming from an intent-to-treat population (that is, from the entire sample including dropouts) were separated from those identified as completers-only, or where no information was provided. In cases where no information was provided, a conservative approach was adopted in performing analyses, such that effect size estimates were calculated assuming these were completer analyses (viz., using pooled standard deviations for calculations based on number of completers).

**Analytic Strategy**

Both treatment- and study-level dropout rates were transformed using the variance stabilizing Freeman-Tukey double arcsine transformation (Freeman & Tukey, 1950), which weights the proportions very slightly towards 50% and thus allows for the inclusion of trials with zero proportions (for an explanation, see Lipsey & Wilson, 2001). Heterogeneity in the sample was evaluated using Cochran's Q-statistic (Cochran, 1954) and the related $I^2$ statistic (Higgins, Thompson, Deeks & Altman, 2003), which describes the percentage of variation across studies that is the result of true heterogeneity as opposed to chance. $I^2$ is easily interpreted, with values of 25%, 50% and 75% corresponding to low, moderate and high heterogeneity. When heterogeneity (between-study variability) is high, analyses should be conducted in a random-effects model (Higgins *et al.*, 2003). A lengthy discussion of this issue is not appropriate, and can be found elsewhere (see Hedges & Vevea, 1998; Schulze, 2004); however, a simple version of the random versus fixed effects issue is that fixed effects models assume that all studies included in the meta-analysis are drawn from the same theoretical population. Conversely, in random effects models, study weights are calculated as the sum of the weight used in a fixed

effects models plus the between-study variability. In practice, these weights are typically very similar to those produced in fixed effects models, but they are better able to reflect between-study variation and provide more equal weighting.

Primary analyses consisted of providing an overall estimate of dropout rate across all studies, along with estimates for each of several relevant, categorical subgroups (where statistically appropriate based on group size). Meta-regression analyses utilized mixed regression models to predict the transformed dropout rate using the variables noted above (akin to models reported by Cuijpers, van Straten, Warmerdam, & Smits, 2008c). All primary and secondary analyses were performed using SAS 9.2, based on guidelines provided by Arthur, Bennett & Huffcutt (2001). Analyses of heterogeneity were first performed in Excel per instructions provided by StatsDirect (2009), and subsequently confirmed in SAS. Finally, a funnel plot was constructed to address the possibility of a "file drawer" problem - the tendency in research literature towards a positive publication bias. While such an effect has not been explicitly suggested with respect to dropout rate, it is plausible that studies with extremely high rates may not be published. Additionally, a previous study using the same initial collection of articles reported some indication of a bias with respect to treatment effect size (Cuijpers, Smit, Bohlmeijer, Hollon & Andersson, 2010). This plot was constructed per guidelines suggested by Light & Pillemer (1984; see also, Egger, Smith, Schneider & Minder, 1997).

## Results

The final collection of articles included 46 separate publications (19% of the psychotherapyrcts.org database), with a total of 2802 patients and 69 treatment conditions represented (see Appendix 2 for a full listing of studies and conditions)[2]. The mean number of

conditions per article represented in these analyses was 1.5, with a range of 1 to 3. Four authors contributed more than one article as the primary author. The final collection included studies published between 1976 and 2009, and 60% of the articles were published in 2000 or later. The majority of articles were American in origin (54%), with the next largest number of contributions originating from the UK (19%). With respect to other basic characteristics of these studies, the majority involved general adult outpatient clientele (65%). The most commonly examined therapy was CBT, which was evaluated in 61% of the studies. Interventions examined ranged from 6 to 30 sessions in duration, with a modal intervention length of 16 sessions. Overall study sizes ranged from 20 to 681 (mean = 112.1, SD = 114). With respect to comparisons being tested in the studies, the most frequently occurring kinds were comparisons of an active therapy versus a control (41%), or anti-depressant medication (ADM) versus an active therapy (40%). Comparisons between two active psychotherapies were only reported in about a quarter of the studies included in this sample. These categories are not mutually exclusive; that is, a study could be classified as involving an active therapy versus control and also a comparison between two psychotherapies if it satisfies both criteria (i.e., two psychotherapy arms versus waitlist). Concerning implementation, slightly over a third of these studies (36%) included some reference to fidelity checking, such as ratings of therapy process and re-training or removal of non-adherent therapists.

The 69 individual treatments conditions ranged in size from 10 to 228 subjects (mean = 40.6, SD = 39.9). Once again, the most common psychotherapy format was CBT (52% of conditions). Twenty percent of the included conditions were combined psychotherapy and pharmacotherapy, and only 3 % involved a pill placebo in conjunction with therapy. Most

treatments were delivered by clinical psychologists; 38% of studies involved only psychologists, and an additional 20% involved a mixed group of psychologists and other professionals.

**Heterogeneity Analyses**

All heterogeneity analyses were conducted using formulae specifically applicable to Freeman-Tukey double arcsine transformed data. (In the interest of interpretability and convenience, all statistical reporting and figures utilize actual proportion estimates, as the transformed data are less clearly interpretable). For the treatment-level analysis, the observed value of the Q-statistic was 242.42. The Q-statistic is compared relative to a critical chi-square value which, for 68 degrees of freedom and $\alpha = .05$, is 88.25. As the observed value is considerably greater than the critical value, this indicates there is significant heterogeneity in this sample. The $I^2$ value for this sample corresponds to 71.9%, with a 95% confidence interval of 64.1 to 78.1, calculated using the method suggested by Higgins and Thompson (2002). Interpreted in line with the anchors suggested by Higgins and colleagues (2003), this is a high level of variability, with approximately 72% of the total variability in dropout estimates being attributable to true heterogeneity between studies as opposed to sampling error. For the overall dropout rate, the observed value of the Q statistic was 227.5, which is greater than the critical value of 61.66 (for 45 degrees of freedom and $\alpha = .05$). Accordingly, the overall dropout estimate is also considered significantly heterogeneous. The $I^2$ value for this sample corresponds to 80.2%, a high level of variability, with a 95% confidence interval of 72.3 to 85.9. These results unambiguously support the use of random effects models in all subsequent analyses.

As the results of these analyses show, the heterogeneity observed in the treatment-level sample is quite high. As therapy type may be the most prominent distinction between treatments in this meta-analysis, heterogeneity analyses were also conducted for each subgroup. Results of

these analyses are reported on the right side of Table 2. Nearly all therapy types evidenced significant heterogeneity with respect to predicted dropout rate; the sole exceptions were behavioral therapies (5 conditions) and those classified as "other" (3 conditions). A $Q_{between}$ score, representing the difference between the overall Q-statistic and the sum of the individual therapy type Q-statistics, tests the hypothesis that the grouping variable (viz., therapy type) accounts for significant variability in dropout estimates. The $Q_{between}$ estimate is 32.8 and is larger than the critical value ($Q_{crit}$ =12.6); thus, there are meaningful differences in the heterogeneity of these subgroups.

**Estimates of Mean Dropout Rate**

In meta-analysis, the mean dropout rate is calculated as a pooled proportion estimate, weighted by study size. In a fixed-effects model, this corresponds to the squared product of the sine of the standardized dropout rate. The random-effects model incorporates between-study variance into this estimate. For the treatment-level analyses, the pooled proportion estimate for the random effects model is .1696. By way of comparison, the fixed-effects estimate is slightly higher at .1879, suggesting a non-trivial difference in estimates based on model specifications. Thus, the best estimate of dropout rate across the entire sample is about 17 percent. Figure 2 shows a forest plot of the treatment-level dropout rates, which are not transformed in this figure in the interest of interpretability. For the overall dropout rate, the pooled proportion effect for the random effects model is .1929, with the comparable fixed-effects estimate slightly higher at .2089. The best estimate of overall dropout rate, across all 46 studies included in the sample was about 19%, slightly higher than the mean dropout rate for psychotherapy conditions. Figure 3 shows a forest plot of the overall dropout rate, with the vertical line representing this average

value. Table 2 provides estimates of mean dropout rate for each therapy type; these estimates range from 14.8% (cognitive-behavioral) to 24.8% ("other" therapies).

**Rater Reliability Estimates**

A subset of the sample (11 studies, including 17 conditions) was randomly selected for the purpose of conducting reliability analyses on variables collected as a part of this study. An advanced undergraduate student was trained in the coding system, provided with a sample study, and then asked to make ratings based on this information. Intra-class correlations (ICCs) were conducted for continuous variables (McGraw & Wong, 1996), and are reported in Table 1. The ICC estimates generally suggested good reliability, with values of .89 or greater for all but one variable. The ICC for comorbid substance abuse was .13, likely attributable to considerable discordance in a single study and a small number of comparisons. Nevertheless, this variable was removed from subsequent analyses due to concerns about reliability. Encouragingly, ICC estimates for condition-level dropout and study-level dropout were .93 and .97 respectively, suggesting that the primary predicted variable was reliably identified in the subset of studies rated here.

The only categorical predictor variable included in the dataset was therapist credentials. Visual inspection of the confirmatory ratings suggested very low reliability, due in part to differences in encoding studies involving therapists from multiple categories (e.g., studies in which most therapists were Ph.D. psychologists, but one therapist was a nurse or social worker). As a consequence, this variable was removed from the dataset, and from further analyses due to concerns about reliability.

**Meta-Regression Analyses**

Table 3 shows the results of the meta-regression analyses. As per the results of the heterogeneity analyses, all meta-regression analyses were conducted with between-studies variance modeled as a random effect. However, because not all variables were available for all conditions, most analyses involve slightly different datasets. There were no statistically significant predictors of treatment-level dropout. Whereas the $Q_{between}$ analyses suggest meaningful differences in the level of heterogeneity observed by subgroup, treatment type was not a significant predictor of dropout, and none of the planned contrasts between the groups were significant. However, intended duration in weeks appeared to predict dropout at the level of a non-significant trend ($t(57) = 1.73$, $p = .09$). A median split was performed on this data, resulting in 33 treatments classified as "short" duration, and 26 as "longer" duration. Conducting estimates of mean dropout rate, studies with shorter duration had an estimated 14% dropout, whereas those with longer duration had an estimated 20% dropout; however, when tested as a contrast, this difference was not statistically significant ($p = .176$). Additional analyses were conducted while controlling for intended duration in weeks; they are not reported here in the interest of space. As in the primary analyses, none of the aforementioned variables significantly predicted dropout in this model. Percentage of comorbid personality disorders also appeared to predict dropout rate at the level of a trend ($t(13) = 1.91$, $p = .08$). However, as reported in Table 1 and noted above, data on this variable was infrequently provided, and often only in the form of exclusion criteria. As such, the result should be interpreted with caution.

In the interest of evaluating potential predictors of overall dropout rate, a second meta-regression was conducted using some of the study-level variables provided in the RCT database (namely, target group, publication year, type of comparison, and type of therapy). In these

analyses, the therapy type variable refers to whether or not a given therapy was used in one or more conditions in the study, unlike the treatment-level analyses, which were separated and classified on the basis of a single therapy type. Results are reported in Table 4. As in the treatment-level analyses, none of these variables predicted overall dropout rate.

**Publication Bias**

A funnel plot was constructed to evaluate the possibility of publication bias related to dropout rate (see Figure 4). The vertical axis measures standard error of the dropout estimate, with the dropout rate on the horizontal axis. In contrast to typical funnel plots, data-points closer to the left side of the horizontal axis likely represent the best outcomes, as they depict lower dropout rates. The funnel plot appears somewhat symmetrical, although three studies separate from the general tendency toward symmetry by reporting lower dropout rates in a larger study. Also, there are conspicuously few small studies that report high dropout rates (that is, studies low on the vertical axis but high on the horizontal axis). A plausible interpretation is that there may be a tendency for small studies to report lower dropout rates, suggesting the possibility that small studies with higher dropout rates may be subject to a "file drawer" effect.

## Discussion

The primary aim of this meta-analysis was to provide an estimate of dropout rates in psychotherapy treatment conditions in RCTs for major depression. Those estimates were 17% for individual conditions, and 19% for overall dropout rates in trials of this kind, with both populations evidencing high levels of heterogeneity in the estimate. Secondary aims of the study were to evaluate potential predictors of dropout from a collection of study- and treatment-level variables. Despite notable heterogeneity in the dropout estimates (implying true variability in the

dropout rate and thus the possibility of detecting meaningful variation), there were no significant predictors of dropout rate in either the treatment- or the study-level analysis. Treatment duration predicted dropout at the rate of a non-significant trend, but a dichotomous split of this variable did not yield significant differences. Heterogeneity estimates were significantly predicted by therapy type in the treatment-level, but this may have been due to disproportionate representation across all 7 categories. Almost all therapy subgroups evidenced significant estimates of heterogeneity, with the two exceptions (behavioral and "other" therapies) being comprised of a small number of conditions, and in one case consisting of multiple treatments from the same study in the group.

Considering the observed mean dropout rates in the context of the existing literature highlights the utility of deriving more precise estimates of dropout relative to particular conditions (e.g., in psychotherapy research; for a particular diagnosis; in a community sample, etc.). For instance, a researcher conducting a small RCT involving cognitive therapy for depression might look to the existing RCT literature and observe that dropout from CT conditions ranges from zero to 37 percent, as in our sample. In searching for an average estimate, the researcher might find dramatically different rates reported in dropout literature in clinical practice using CT (40 - 50%; Persons *et al.*, 1993; Trepka, 1986) versus large-scale community samples with variable diagnoses and treatment approaches (16 – 20%; Edlund *et al.*, 2002; Olfson *et al.*, 2009). It is likely that an argument could be made for using any of these estimates; however, a more precise and applicable estimate would strengthen the researcher's ability to use it in an informative way in designing the study. With additional efforts to refine and expand this process to incorporate a broader range of treatment types and diagnoses, there is potential for a superior method of evaluating one's dropout rate to emerge.

With respect to the process of identifying dropout rates, it bears noting that the majority of studies that were excluded from the sample were removed in the interest of reducing certain sources of variability (e.g., disallowing group treatments, etc.). As such, the true proportion of all depression RCTs that provided dropout rates cannot be easily ascertained from the estimates provided herein. However, from the studies that were reviewed more extensively for dropout information, a surprising number were removed on the grounds of failure to report dropout rate by condition in an unambiguous fashion (see Appendix 1). Although directives such as the CONSORT statement (Moher *et al.*, 2010) may be expected to improve the quality of reporting in journals that implement these standards, there is a pressing need to increase transparency about dropout rates in RCTs, regardless of the journal of publication. The estimates derived in this study only represent a fraction of the available literature, and otherwise strong studies that fail to report on dropout in a straightforward fashion do a disservice to the field.

There are multiple ways of interpreting the failure to detect significant predictors. It is plausible that they reflect a true non-relation between the selected variables and risk of dropout. If true, this would suggest either a problem identifying relevant predictor variables or an issue with the type of analysis being performed. The former outcome is certainly possible, as the variables were selected largely on the basis of their role as predictors in large-scale epidemiological investigations of dropout, a literature that is not without contradictory findings. Similarly, it may be that dropout rates in psychotherapy treatments of RCTs for depression are better predicted by intermediate variables that were not included, or are not amenable to condition- or study-level investigations. Higgins and Thompson (2004) commented that in meta-analyses with very high heterogeneity with numerous predictor variables, spurious predictive findings are likely to be reported due to the nature of the analyses. Although no significant

predictors were identified in this study, the potential for misleading results in future analyses should be addressed by thoughtful, *a priori* selection of predictor variables.

The analyses related to therapy type warrant additional consideration. Although there was no evidence of differences in dropout rates between therapy types – a finding that may be in part due to small numbers of treatments for most therapies – the differences in heterogeneity give rise to the potential for differential predictors (or predictive abilities) by therapy. The lack of differences between therapy types contrasts with findings reported by Cuijpers and colleagues (2008a). Using overall dropout rates in a study involving only psychotherapy versus psychotherapy comparisons (with a database that included many studies in common with the present study), they reported an elevated risk of dropout in studies involving cognitive therapy, and a reduced risk in problem-solving therapies. The authors noted that the disproportionate representation of these therapies might be a reason for caution in interpreting their findings. As the same imbalance is true of the present study, with cognitive therapy being better represented in the sample than other therapies, this non-finding should also be interpreted with caution.

An additional possibility is suggested by the finding that therapy types differ with respect to heterogeneity of dropout estimates. Although there was no evidence of differences in dropout rates between therapy types – a finding that may be in part due to small numbers of treatments for most therapies – the differences in heterogeneity give rise to the potential for differential predictors (or predictive abilities) by therapy. That is, heterogeneity in "true" estimate variability across therapy types suggests at least the possibility that different predictors may be at play. Unfortunately, the small numbers of examples of most therapy types do not allow for an adequate evaluation of this question, as moderation analyses using therapy type as a moderator would likely be under-powered, especially since almost none of the predictor variables were

available for every condition (Hedges & Pigott, 2003). Conducting moderator analyses in low power conditions may yield results that are difficult to interpret, especially in the absence of significant findings. Future efforts to extend this research area to additional RCT variants may open up the possibility of moderator analyses that can adequately test hypotheses.

Considering the results of the funnel plot and the issue of publication bias, there may be a relationship between study size and dropout rate in this collection of studies. This was an exploratory analyses conducted without a clear precedent to suspect such an effect; however, the observed result may make some intuitive sense in light of some known properties of published trials. The rationale for investigating publication bias is that studies with significant findings may be more likely to be published, and those without significant findings more likely to be relegated to file drawers. Additionally, RCTs conducted with larger samples may be more likely to be published than those conducted in smaller samples (perhaps because they are grant-supported or considered more precise due to better power to detect effects). When dropout rates are high, they may complicate interpretation of results, perhaps especially so in last-observation-carried-forward analyses, where they may reduce the magnitude of observed effects (Lane, 2008). Small studies may be particularly susceptible to these effects, and especially sensitive to aberrant dropout rates. As such, high dropout rates in small studies may be conspicuous in their own right, and may interfere with findings of interest, compounding the likelihood that the study will not be submitted or accepted for publication.

**Limitations**

The estimates derived in this study may prove informative to those interested in dropout in the context of psychotherapy research, and, more specifically, in those types of studies to which these analyses were restricted. In the interest of constructing a straightforward

comparison, several types of studies were removed from consideration, and as such these results may not be easily generalized to studies involving those variables. Additional efforts to construct larger databases will be necessary to improve generalizability.

A critical limitation of this study is that, while one can informally compare estimates of dropout derived from epidemiological designs or clinical contexts to the results derived in this paper, a formal comparison has not been conducted. Additionally, as the estimates of overall dropout rate were only obtained from studies with at least one viable psychotherapy condition, it is possible that the observed study-level dropout estimate in this sample differs from one constructed using all available studies. However, there is no strong reason to suspect such a difference, at least with respect to studies that would meet other inclusion criteria for this study.

Finally, efforts to evaluate reliability of ratings suggest that while most variables were reliably identified, there are still domains in which inter-rater agreement could be improved. Future efforts in this dataset and others will incorporate additional ratings for more precise estimates, also opening up the potential for consensus ratings, which may improve reliability in difficult categorical variables such as therapist experience.

**Conclusions**

The phenomenon of dropout is of considerable interest to both clinicians and researchers alike, but the existing literature on the topic may be suboptimal for researchers hoping to anticipate dropout rates when designing psychotherapy research studies, or interpret those observed in their own RCTs. This meta-analysis provides an estimate of dropout for a small piece of the RCT literature: individual outpatient psychotherapies for major depression. The estimates derived in this study may help researchers conducting RCTs in this field make informed decisions with respect to their anticipated dropout rates.

**References**


*Alexopoulos, G.S., Raue, P., & Arean, P. (2003). Problem-solving therapy versus supportive therapy in geriatric major depression with executive dysfunction. *American Journal of Geriatric Psychology, 11,* 46-52.

Arnow, B. A., Blasey, C., Manber, R., Constantino, M. J., Markowitz, J. C., Klein, D. N., ... Rush, A.J. (2007). Dropouts versus completers among chronically depressed outpatients. *Journal of Affective Disorders, 97,* 197-202.

Arthur, W. Jr., Bennett, W. Jr., & Huffcutt, A.I. (2001). *Conducting meta-analysis using SAS.* Mahwah, N.J.: Lawrence Erlbaum Associates.

Bados, A., Balaguer, G., & Saldana, C. (2007). The efficacy of cognitive-behavioral therapy and the problem of drop-out. *Journal of Clinical Psychology, 63,* 585-592.

Barrett, M.S., Chua, W-J., Crits-Christoph, P., Gibbons, M.B., Casiano, D. & Thompson, D. (2008). Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy, 45,* 247-267.

*Beck A.T., Hollon, S.D., Young, J.E., Bedrosian, R.C., & Budenz, D. (1985). Treatment of depression with cognitive therapy and amitriptyline. *Archives of General Psychiatry, 42,* 142-148.

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.

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Manual for Beck Depression Inventory II (BDI-II).* San Antonio, TX: Psychology Corporation.

Beutler, L.E., Malik, M., Alimohamed, S., Harwood, T.M., Talebi, H., Noble, S., & Wong, E. (2003). Therapist variables. In M. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (5th ed, pp.227-307). New York: John Wiley & Sons.

*Blackburn IM & Moore RG. (1997). Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with current depression. *British Journal of Psychiatry, 171,* 328-334.

*Burnand, Y., Andreoli, A., Kolatte, E., Venturini, A., & Rosset, N. (2002). Psychodynamic psychotherapy and clomipramine in the treatment of major depression. *Psychiatric Services, 53,* 585-589.

*Castonguay, L.G., Schutt, A.J., Aikins, D.E., Constantino, M.J., Laurenceau, J.P., Bologh, L., & Burns, D.D. (2004). Integrative cognitive therapy for depression: A preliminary investigation. *Journal of Psychotherapy Integration, 14,* 4–20.

Cochran, W. (1954). Some methods of strengthening the common $\chi 2$ test. *Biometrics, 10,* 417-451.

Coffman, S.J., Martell, C.R., Dimidjian, S., Gallop, R. & Hollon, S.D. (2007) Extreme nonresponse in cognitive therapy: can behavioral activation succeed where cognitive therapy fails *Journal of Consulting and Clinical Psychology, 75,* 531-541.

*Constantino, M.J., Marnell, M.E., Haile, A.J., Kanther-Sista, S.N., Wolman, K., Zappert, L., & Arnow, B.A. (2008). Integrative cognitive therapy for depression: A randomized pilot comparison, *Psychotherapy, 45,* 122 – 134.

*Cooper, P., Murray, L., Wilson, A. & Romaniuk, H. (2003). Controlled trial of the short- and long-term effect of psychological treatment of past-partum depression. I. Impact on maternal mood. *British Journal of Psychiatry, 182,* 412-419.

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. *British Journal of Psychiatry, 196,* 173-178.

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008a). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology, 76,* 909-922.

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

Cuijpers, P., van Straten, A., Warmerdam, L. & Smits, N. (2008c). Characteristics of effective psychological treatments of depression: a metaregression analysis. *Psychotherapy Research, 18,* 225-236.

*David, D., Szentagotai, A., Lupu, V., & Cosman, D. (2008). Rational emotive behavior therapy, cognitive therapy, and medication in the treatment of major depressive disorder: a randomized clinical trial, post-treatment outcomes, and six-month follow-up. *Journal of Clinical Psychology, 64,* 728-746.

*de Jonghe, F., Hendricksen, M, van Aalst, G., Kool, S., Peen, V., Van, R., van den Eijnden, E., & Dekker, J. (2004). Psychotherapy alone and combined with pharmacotherapy in the treatment of depression. *British Journal of Psychiatry, 185,* 37-45.

*de Jonghe, F., Kool, S., van Aalst, G., Dekker, J. & Peen, J. (2001).Combining psychotherapy and antidepressants in the treatment of depression.. *Journal of Affective Disorders, 64,* 217-229.

*Dekker , J.J.M., Koelen, J.A., Van, H.L., Schoevers, R.A., Peen, J., Hendriksen, M., ... de Jonghe F. (2008). Speed of action: The relative efficacy of short psychodynamic supportive psychotherapy and pharmacotherapy in the first 8 weeks of a treatment algorithm for depression. *Journal of Affective Disorder, 109,* 183 - 188.

*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.

*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 antidepre ssant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74,* 658-670.

*Dozois, D.J.A., Bieling, P.J., Patelis-Siotis, I., Hoar, L., Chudzik, S., McCabe, K., & Westra, H.A. (2009). Changes in self-schema structure in cognitive therapy for major depressive disorder: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 77,* 1078-1088.

Edlund, M.J, Wang, P.S. Berglund, P.A., Katz, S.J., Lin, E. & Kessler, R.C. (2002). Dropping out of mental health treatment: patterns and predictors among epidemiological survey respondents in the United States and Ontario. *American Journal of Psychiatry, 159,* 845-851.

Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629-634.

Fournier, J.C., DeRubeis, R.J., Shelton, R.C., Gallop, R., Amsterdam, J.D., & Hollon, S.D. (2008). Anti-depressant medications v. cognitive therapy in people with depression with or without personality disorder. *British Journal of Psychiatry, 192,* 124-129.

Freeman, M.F. & Tukey, J.W. (1950). Transformations related to the angular and the square root. *The Annals of Mathematical Statistics, 21,* 607 – 611.

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

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

Hamilton, S., Moore, A.M., Crane, D.R. & Payne, S.H. (2010). Psychotherapy dropouts: differences by modality, license and DSM-IV diagnosis. *Journal of Marital and Family Therapy*, doi: 10.1111/j.1752-0606.2010.00204.x

Hedges, L.V. & Pigott, T.D. (2003). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426-445.

Hedges, L.V., & Vevea, J.L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3,* 486-504.

Higgins, J.P., &Thompson, S.G. (2002). Quantifying heterogeneity in meta-analysis. *Statistics in Medicine, 21,* 1539-1558.

Higgins, J.P. & Thompson, S.G. (2004). Controlling the risk of spurious findings from meta-analysis. *Statistics in Medicine, 23,* 1663 – 1682.

Higgins, J.P., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557-560.

*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,* 774-781.

*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,* 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,* 431-437.

*Keller, M.B., McCullough, J.P., Klein, D.N., Arnow, B., Dunner, D.L., Gelenberg, A.J., ... Zajecka, J. (2000). A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *The New England Journal of Medicine, 342,* 1462-1470.

Klein, E.B., Stone, W.N., Jicks, M.W. & Pritchard, I.L. (2003). Understanding dropouts. *Journal of Mental Health Counseling, 25,* 89-100.

*Laidlaw, K., Davidson, K., Toner, H., Jackson, G., Clark, S., Law, J., ... Cross, S. (2008). A randomised controlled trial of cognitive behaviour therapy vs treatment as usual in the treatment of mild to moderate late life depression. *International Journal Geriatric Psychiatry, 23,* 843 - 850.

Lane, P. (2008). Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharmaceutical Statistics, 7,* 93-106.

*Lespérance, F., Frasure-Smith, N., Koszycki, D., Laliberté, M-A., van Zyl, L.T., Baker, B., ... Guertin M-C. (2007). Effects of citalopram and interpersonal psychotherapy in coronary artery disease patients. The Canadian Cardiac Randomized Evaluation of Antidepressant and Psychotherapy Efficacy Trial (CREATE). *Journal of the American Medical Association, 297,* 367-412.

Light, R.J., & Pillemer, D.B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage Publications.

*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,* 613-21.

Macaskill, N.D., & Macaskill, A. (1996). Rational-emotive therapy plus pharmacotherapy vs. pharmacotherapy alone in the treatment of high cognitive dysfunction depression. *Cognitive Therapy and Research, 20,* 575-592.

McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-46.

*McKnight, D.L., Nelson-Gray, R.O. & Barnhill, J. (1992). Dexamethasone suppression test and response to cognitive therapy and antidepressant medication. *Behavior Therapy, 23*, 99-111.

Minniti, A., Bissoli, L., DiFrancesco, V., Fantin, F., Mandragona, R., Olivieri, M., … Zamboni, M. (2007) Individual versus group therapy for obesity: comparison of dropout rate and treatment outcome. *Eating and Weight Disorders, 12,* 161-167.

*Misri, S., Reebye, P., Corral, M., & Milis, L. (2004). The use of paroxetine and cognitive-behavioral therapy in postpartum depression and anxiety: a randomized controlled trial. *Journal of Clinical Psychiatry, 65,* 1236-1241.

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gotzsche, P.C., Devereaux, P.J. ... for the CONSORT group. (2010) Explanation and elaboration: Updated guidelines for reporting parallel group randomized trial. *British Medical Journal, 340,* c869.

*Mohr, D.C., Boudewyn,A.C., Goodkin, D.E., Bsotrom, A., & Epstein, L. (2001). Comparative outcomes for individual cognitive-behavior therapy, supportive-expressive, group psychotherapy, and sertraline for treatment of depression in multiple-sclerosis. *Journal of Consulting and Clinical Psychology, 69,* 942-949.

Mueller, M. & Pekarik, G. (2000). Treatment duration prediction: client accuracy and its relationship to dropout, outcome and satisfaction. *Psychotherapy,* 37, 117 – 123.

*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,* 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,* 33-41.

*Mynors-Wallis, L.M., Gath, D.H., Day, A. & Baker, F. (2000). Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *British Medical Journal, 320,* 26-30.

*Mynors-Wallis, L.M., Gath, D.H., Lloyd-Thomas, A.R., & Tomlinson, D. (1995). Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *British Medical Journal, 310,* 441-445.

Oei, T.P.S., & Kazmierczak, T. (1997) Factors associated with dropout in a group cognitive behavior therapy for mood disorders. *Behavior Research and Therapy, 36,* 1025-1030.

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

Olfson, M., Mojtabai, R., Sampson, N.A., Hwang, I., Druss, B., Wang, P.S., ... Kessler, R.C. (2009). Dropout from outpatient mental health care in the United States. *Psychiatric Services, 60,* 898-907.

Organista, K.C, Munoz, R.F. & Gonzalez, G. (1994). Cognitive-behavioral therapy for depression in low-income and minority medical outpatients: description of a program and exploratory analysis. *Cognitive Therapy and Research, 18,* 241-259.

*Padfield, M. (1976). The comparative effects of two counseling approaches on the intensity of depression among rural women of low socioeconomic status. *Journal-of-Counseling Psychology, 23,* 209-214.

Persons, J.B., Burns, D.D., & Perloff, J.M. (1988). Predictors of dropout and outcome in cognitive therapy for depression in a private practice setting. *Cognitive Therapy and Research, 12,* 557-575.

Reis, B.F. & Brown, L.G. (1999). Reducing psychotherapy dropouts: maximizing perspective convergence in the psychotherapy dyad. *Psychotherapy, 36,* 123-136.

Roth, A. & Fonagy, P. (2005). *What works for whom?: A critical review of psychotherapy research.* New York: Guilford Press.

Rusch, N., Schiel, S., Corrigan, P.W., Leihener, F., Jacob, G.A., Olschewski, M. *et al.*. (2008). Predictors of dropout from inpatient dialectical behavior therapy among women with borderline personality disorder. *Journal of Behavior Therapy & Experimental Psychiatry,* 39, 497 – 503.

*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 and Research, 1,* 17-37.

Rush, A.J. & Watkins, J.T. (1981). Group versus individual cognitive therapy: a pilot study. *Cognitive Therapy and Research, 5,* 95-103.

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

*Schulberg, H.C., Block, M.R., Madonia, M.J., Scott, C.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,* 913-919.

Schulze, R. (2004). *Meta-analysis: A comparison of approaches.* Toronto: Hogrefe & Huber.

*Scott, A.I., & Freeman, C.P. (1992). Edinburgh primary care depression study: treatment outcome, patient satisfaction, and cost after 16 weeks. *British Medical Journal: Clinical Research Edition, 304,* 883-887.

Sledge, W.H., Moras, K., Hartley, D., & Levine, M. (1990). Effect of time-limited psychotherapy on patient dropout rates. *American Journal of Psychiatry, 147,* 1341 -1347.

*Sloane, R.B., Staples, F.R., & Schneider, L.S. (1985). Interpersonal therapy versus nor-triptyline for depression in the elderly. In: G.D. Burrows & L. Dennerstein (Eds.) *Clinical and Pharmacological Studies in Psychiatric Disorders*. London, UK: John Libby.

*Smit, A., Kluiter, H., **Conradi, H.J.,** Van der meer, K., Tiemens, B.G., Jenner, J.A., ... Ormel, J. (2006). Short-term effects of enhanced treatment for depression in primary care: results from a randomized controlled trial. *Psychological Medicine, 36,* 15–26.

Stark, M.J. (1992). Dropping out of substance abuse treatment: a clinically oriented review. *Clinical Psychology Review, 12,* 93-116.

Statsdirect.com (2009). Proportion meta-analysis from Statsdirect help. Retrieved September 22, 2010 from http://www.statsdirect.com/help/meta_analysis/proportion_meta_analysis.htm

Sterne, J.A.C., & Harbord, R.M. (2004). Funnel plots in meta-analysis. *The Stata Journal, 2,* 127 – 141.

*Swartz, H.A., Frank, E., Zuckoff, A., Cyranowski, J.M., Houck, P.R., Cheng, Y., ... Shear, M.K.(2008). Brief interpersonal psychotherapy for depressed mothers whose children are receiving psychiatric treatment. *American Journal of Psychiatry, 165,* 1155 - 1162.

*Teasdale, J.D., Fennal, M.J., Hibbert, G.A., & Amies, P.L. (1984). Cognitive therapy for major depressive disorder in primary care. *British Journal of Psychiatry,144,* 400-406.

*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. *The American Journal of Geriatric Psychiatry, 9,* 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,* 385-390.

Trepka, C. (1986). Attrition from an out-patient psychology clinic. *British Journal of Medical Psychology, 59,* 181-186.

*van Schaik, A., van Marwijk, H., Ader, H., van Dyck, R., de Haan, M., Penninx, B., ... Beekman, A. (2006). Interpersonal psychotherapy for elderly patients in primary care. *American Journal of Geriatric Psychiatry, 14,* 777-786.

*Vitriol, V.G., Ballesteros, S.T., Florenzano, R.U., Weil, K.P., & Benadof, D.F. (2009). Evaluation of an outpatient intervention for women with severe depression and a history of childhood trauma. *Psychiatric Services,* 60, 936-942.

Wampold, B.E., & Brown, G.S. (2005). Estimating variability in outcome attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914-923.

Wang, J. (2007). Mental health treatment dropout and its correlates in a general population sample. *Medical Care, 45,* 224-229.

Wang, P.S., Gilman, S.E., Guardino, M., Christiana, J.M., Morselli, P.L., Mickelson, K., & Kessler, R.C. (2000). Initiation of and adherence to treatment for mental disorders: Examination of patient advocate group members in 11 countries. *Medical Care, 38,* 926-936.

Wierzbicki, M. & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice, 24,* 190-195.

*Wiles, N.J., Hollinghurst, S., Mason, V., Musa, M., Burt, V., Hyde, J., ... Kessler, D. (2008). A randomized controlled trial of cognitive behavioural therapy as an adjunct to pharmacotherapy in primary care based patients with treatment resistant depression: A pilot study. *Behavioural Cognitive Psychotherapy, 36,* 21 – 33.

*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,* 1158-1164.

## Footnotes

1. There is some controversy as to whether or not it is appropriate to include estimates from samples within the same study in a meta-analysis and treat them as independent of one another; however, Schulze (2004) has argued that inclusion of groups from the same study is likely no worse than including multiple studies from the same research group, which are also likely to share common sources of variance. As such, for the purposes of this project, multiple conditions within a given study were included in the analyses, so long as participants were ineligible from participating in multiple conditions (i.e., participants from one condition cannot have participated in the other condition).

2. The 46 articles used in the final meta-analysis are marked with an asterisk in the References for ease of identification. Note that articles were labeled and identified by first author based on the version of the database available at time of downloading. Subsequent revisions to that list (due to updated information on later publications) resulted in a name change of one primary author: the paper cited in the meta-analysis as "Conradi, 2006" appears in the reference section as "Smit et al, 2006". In order to maintain connections with the primary database, the original name has been retained. Original database citations are also given in Appendix 1 and 2.

Table 1. Mean values and availability of treatment- and study-level predictor variables, with intra-class correlations (ICCs).

| | $N_{cond}$ (%) | Mean | Weighted Mean | ICC |
|---|---|---|---|---|
| Mean Age | 65 (94) | 44.4 | 42.5 | 1.00 |
| % Female Patients | 69 (100) | 73.0 | 70.7 | 1.00 |
| Therapist Experience (years) | 25 (36) | 8.5 | 9.2 | 1.00 |
| Duration | | | | |
| Sessions Intended | 55 (80) | 16.4 | 16.7 | .89 |
| Weeks Intended | 59 (86) | 13.9 | 15.2 | 1.00 |
| Sessions Observed | 21 (30) | 11.6 | 12.6 | 1.00 |
| % Minority Patients | 38 (55) | 14.3 | 13.2 | 1.00 |
| Comorbid Substance Use | 37 (52) | 3.8 | 3.5 | .13 |
| Comorbid Axis II | 15 (22) | 31.7 | 45.4 | 1.00 |
| Comorbid Axis I | 11 (16) | 34.9 | 43.1 | 1.00 |
| Comorbid anxiety disorder | 10 (14) | 42.5 | 34.4 | 1.00 |
| Pre-post BDI *d* effect size | | | | |
| Any version | 37 (54) | .28 | .27 | 1.00 |
| Confirmed ITT | 23 (33) | .31 | .29 | |
| Pre-post HRSD *d* effect size | | | | |
| Any version | 45 (65) | .46 | .38 | .99 |
| Confirmed ITT | 30 (43) | .38 | .34 | |

$N_{cond}$ refers to the number of conditions for which an estimate the corresponding variable was available. Mean values not given for categorical variables. ICC values for BDI and HRSD reflect average agreement across pre- and post- scores. Comorbid substance use was removed from further analyses due to poor reliability.

Table 2. Heterogeneity statistics and dropout estimates for treatment-level analyses.

| **Condition** | **$N_{cond}$** | **$N_{subj}$** | **Estimate** | **Q** | **$I^2$** |
|---|---|---|---|---|---|
| All treatments | 69 | 8506 | .1879 (F)<br>.1696(R) | 242.4* | 71.9 |
| Cognitive | 36 | 3251 | .1474 (F)<br>.1481 (R) | 77.2* | 54.7 |
| Interpersonal | 8 | 1296 | .1862(F)<br>.1674(R) | 56.7* | 87.7 |
| Problem-solving | 6 | 593 | .1811(F)<br>.1606(R) | 18.1* | 72.3 |
| Supportive | 3 | 242 | .1325(F)<br>.1348(R) | 6.6* | 69.5 |
| Behavioral | 4 | 532 | .1658(F)<br>.1909(R) | 7.3 | 59.1 |
| Dynamic | 9 | 1202 | .2396(F)<br>.2288(R) | 40.2* | 80.1 |
| Other | 3 | 1390 | .2479(F)<br>.2446(R) | 3.5 | 43.6 |

$N_{cond}$ = total number of individual treatment conditions represented in the primary analyses
$N_{subj}$ = total number of subjects in the relevant treatment conditions
Estimate = estimated dropout rate, with (F) and (R) denoting fixed and random effects calculations used in deriving estimates. When heterogeneity is significant, random effects models are considered the most appropriate estimates.
* $p < .05$

Table 3. Meta-regression results predicting treatment-level dropout.

| **Predictor** | **Estimate** | **SE** | **df** | ***t* / *F*** | ***p*** |
|---|---|---|---|---|---|
| Therapy Type* | - | - | 62 | .59 | .73 |
| Population* | - | - | 64 | .68 | .61 |
| Mean Age | -.002 | .004 | 63 | -.48 | .63 |
| % Female Patients | .343 | .311 | 67 | 1.11 | .27 |
| Therapist Experience | | | | | |
| Years | -.007 | .015 | 23 | -.47 | .64 |
| Duration | . | | | | |
| Sessions Intended | .004 | .013 | 53 | .27 | .79 |
| Weeks Intended | .020 | .012 | 57 | 1.73 | .09 |
| Sessions Observed | .5975 | 1.03 | 13 | .58 | .58 |
| % Minority Patients | -.006 | .594 | 36 | .09 | .93 |
| Comorbid Axis II | -.798 | .418 | 13 | 1.91 | .08 |
| Comorbid Axis I | .515 | .524 | 9 | .98 | .35 |
| Comorbid anxiety disorder | -.002 | .005 | 8 | -.31 | .77 |
| Pre-post BDI *d* effect size | | | | | |
| Any version | -.283 | .393 | 35 | -.72 | .48 |
| Confirmed ITT | -.249 | .418 | 21 | -.59 | .56 |
| Pre-post HRSD *d* effect size | | | | | |
| Any version | -.279 | .216 | 43 | -.83 | .41 |
| Confirmed ITT | -211 | .276 | 28 | -.77 | .45 |

Estimates are of unstandardized regression coefficients, with corresponding standard error (SE) values. All tests involved random effects models. Results involving categorical variables (identified with a *) are given as F statistics, and estimates of regression coefficients are not reported.

Table 4. Meta-regression results predicting overall dropout rate.

| **Predictor** | **df** | ***F*** | ***p*** |
|---|---|---|---|
| Target Group | 41 | .48 | .75 |
| Country of Origin | 41 | .87 | .49 |
| Publication Year | 24 | .44 | .97 |
| Psychotherapy vs. Control | 44 | .51 | .48 |
| Psychotherapy vs Medication | 44 | 1.22 | .28 |
| Psychotherapy vs. Combined treatment | 44 | 2.57 | .12 |
| Medication vs. Combined treatment | 44 | .27 | .61 |
| Combined vs. Psychotherapy + Placebo | 44 | .06 | .81 |
| Psychotherapy vs Psychotherapy | 44 | .40 | .53 |
| Overall study size | 5 | .49 | .91 |
| Studies involving Cognitive therapy | 44 | 1.75 | .19 |
| Studies involving Interpersonal therapy | 44 | 1.14 | .29 |
| Studies involving Problem-Solving therapy | 44 | .45 | .51 |
| Studies involving Supportive-Expressive therapy | 44 | 1.43 | .24 |
| Studies involving Behavioral therapy | 44 | 0 | .99 |
| Studies involving Psychodynamic therapy | 44 | .28 | .60 |
| Studies involving “other” therapies | 44 | .04 | .84 |

Variables reported above were available for all 46 studies included in the meta-analysis. Individual studies could be categorized into multiple treatment type and comparison type variables.

Figure 1. Flow diagram of study collection process.

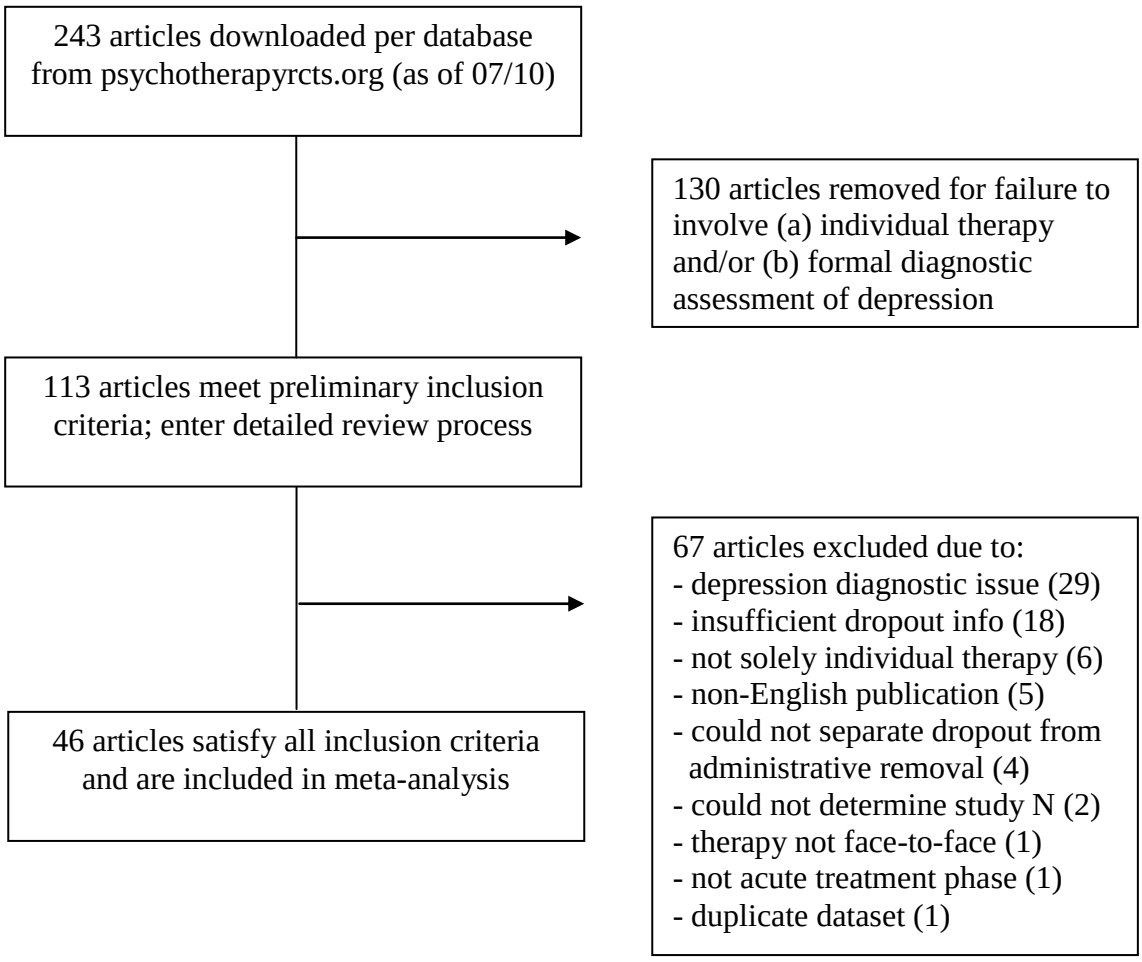
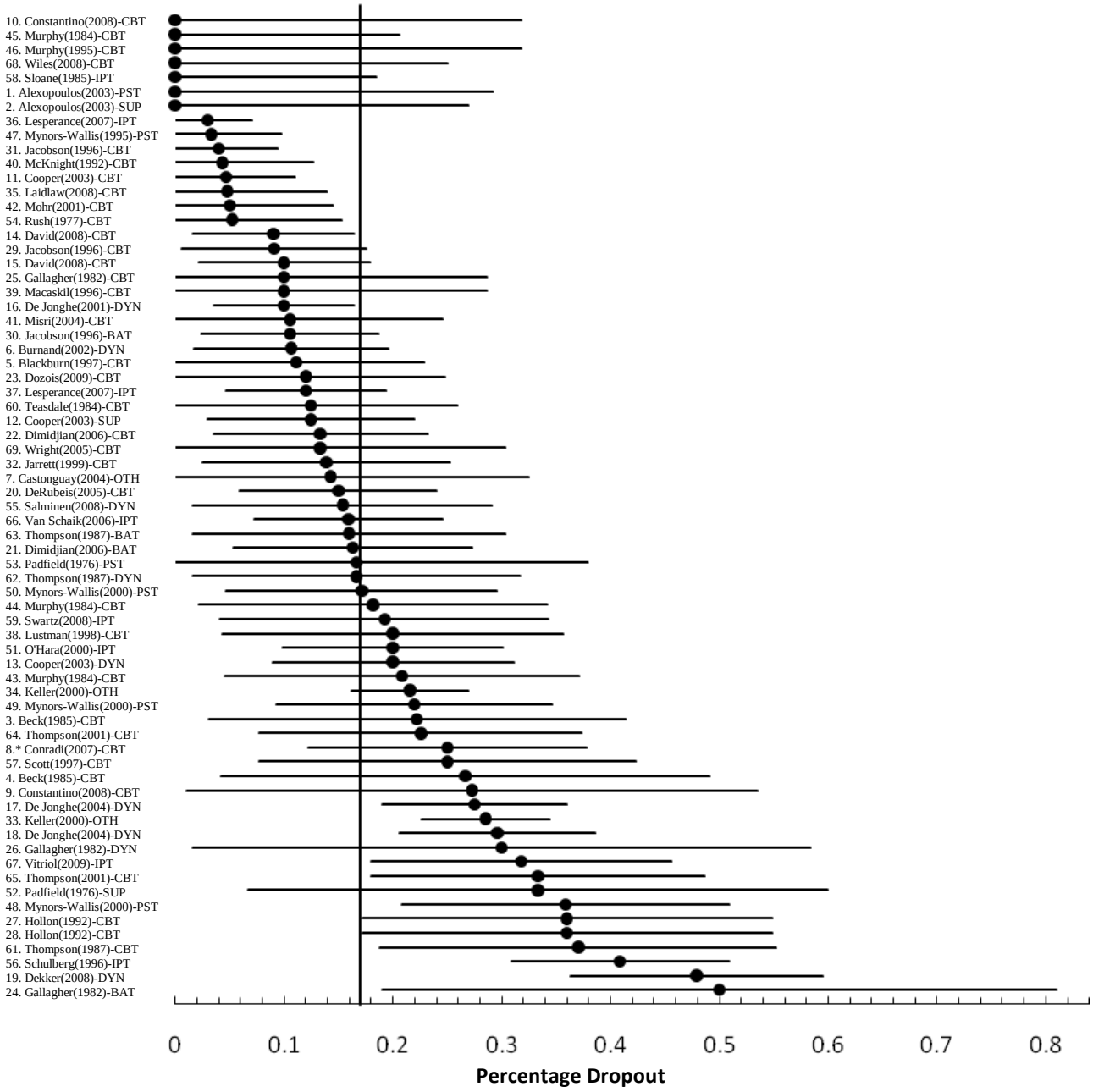
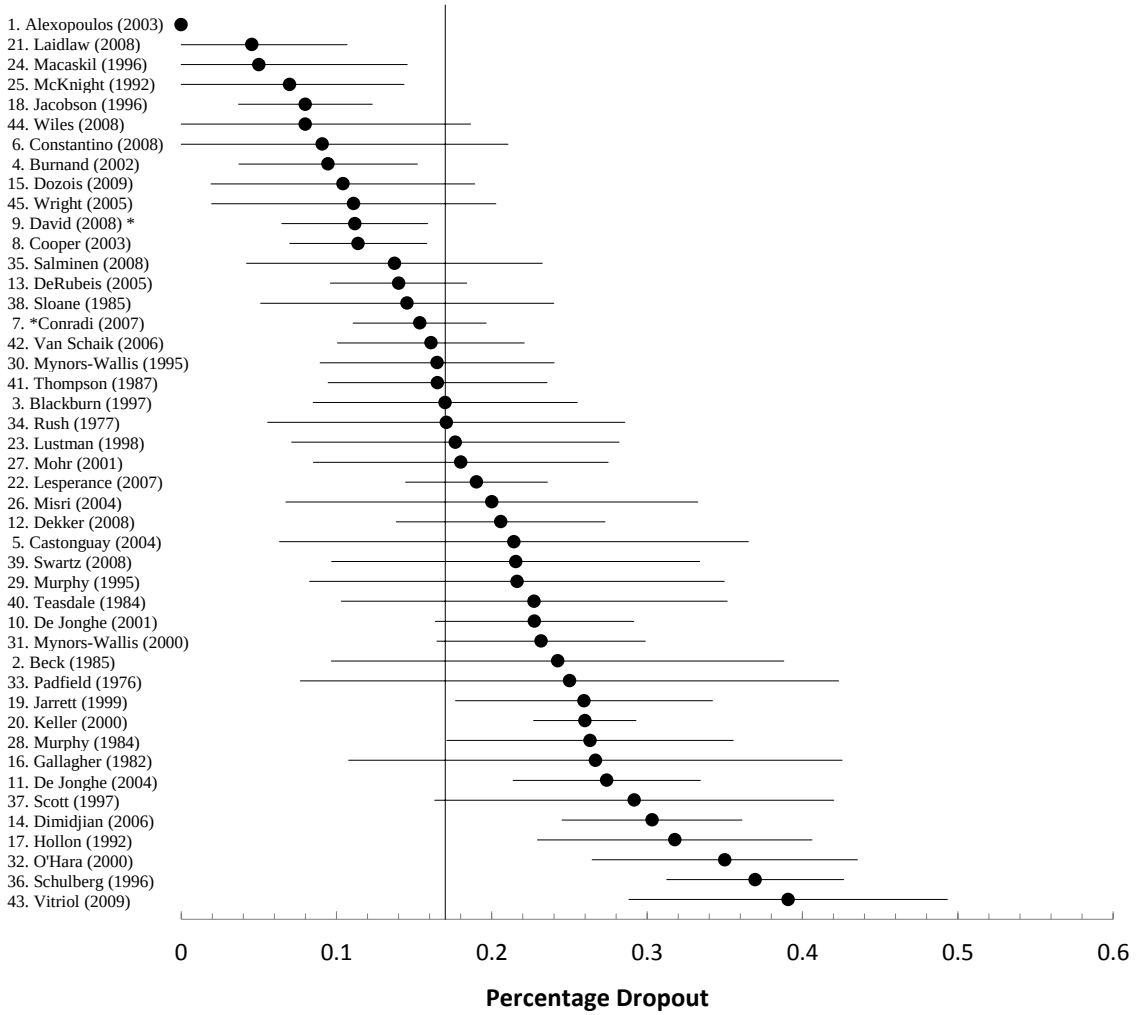


Figure 2. Forest plot of treatment-level dropout rate.



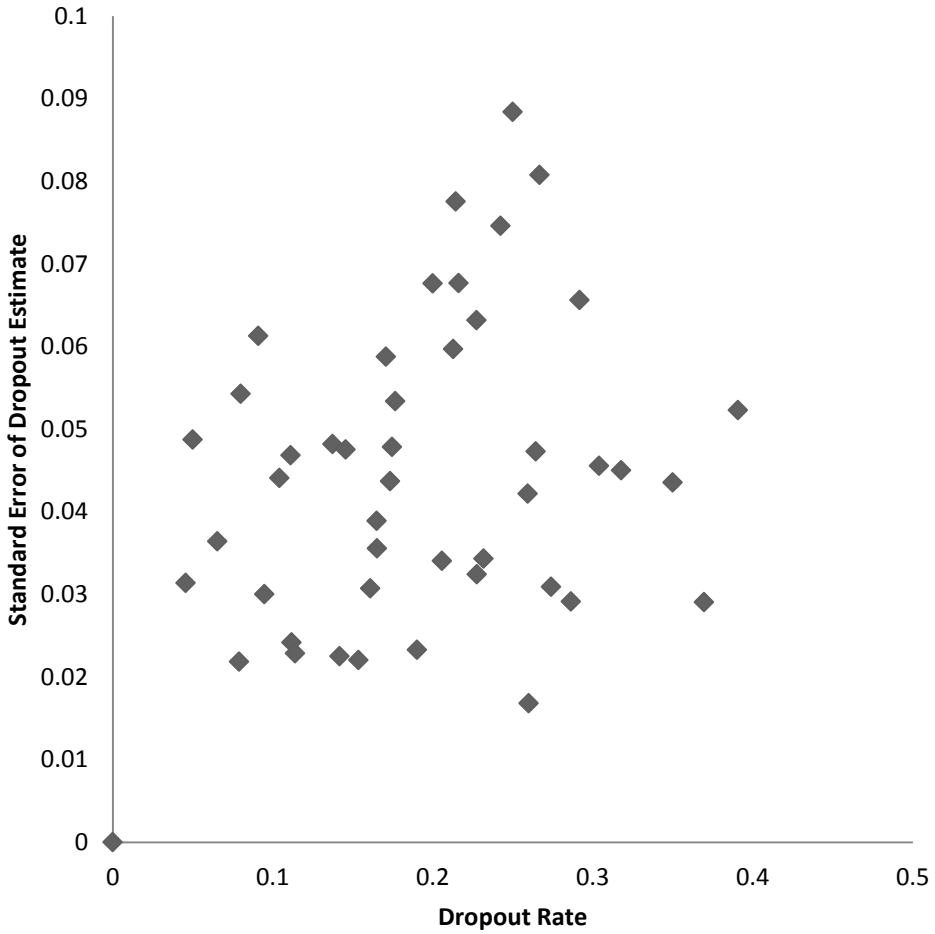

The figure above depicts dropout rates, with 95% confidence intervals, for each condition included in the primary analysis. The vertical line denotes the group mean dropout rate (approximately 17%). Labels on the vertical axis include treatment number (as per Appendix 2), name of the first author, year of publication, and treatment type of the condition. Note that because all values represent probabilities, low dropout rates with wide confidence intervals are in some cases necessarily bounded on one side by zero.

Figure 3. Forest plot of study-level dropout rates.




The figure above depicts untransformed dropout rates, with 95% confidence intervals, for each study included in the secondary analysis. The vertical line denotes the group mean dropout rate (approximately 19%). Labels on the vertical axis include study number, name of the first author and year of publication. Note that because all values represent probabilities, low dropout rates with wide confidence intervals are in some cases necessarily bounded on one side by zero.

Figure 4. Funnel plot of standard error of dropout estimate by dropout rate.




The figure above features the standard error of the dropout estimate, plotted against the dropout rate, for all 46 studies included in this meta-analysis (per Sterne & Harbord, 2004).

## Appendix 1

*Note: Studies are identified by first author and date of publication. Complete citations can be found in at psychotherapyrcts.org.

*Removed due to inability to distinguish dropout from withdrawal by study personnel*

| | | | |
|---|---|---|---|
| Bellack, 1981 | Bellino, 2006 | Elkin, 1990 | Weissman, 1979 |

*Removed due to depression diagnostic issue (e.g., no formal assessment; primary dysthymia)*

| | | | |
|---|---|---|---|
| Appleby, 1997 | Dowrick, 2000 | Markowitz, 2005 | Teichmann, 1995 |
| Barrett, 2001 | Dunner, 1996 | Markowitz, 2008 | Teri, 1986 |
| Beutler, 2003 | Floyd, 2004 | McGrath, 1996 | Teri, 1997 |
| Brown, 1984 | Freedland, 2009 | Prendergast, 2001 | Wierzbicki, 1987 |
| Browne, 2002 | Gallagher, 1994 | Safran, 2009 | Williams, 2000 |
| Carpenter, 2008 | Hernandez, 2004 | Selmi, 1990 | |
| Cho, 2008 | Holden, 1989 | Serfaty, 2009 | |
| De Mello, 2001 | Maina, 2005 | Strauman, 2006 | |

*Removed because information on dropout / treatment-level dropout could not be identified*

| | | | |
|---|---|---|---|
| Alladin, 2007 | Marshall, 2008 | Schiffer, 1990 | Spinelliy, 2003 |
| Bedi, 2000 | McBride, 2007 | Schmitz, 2001 | Thompson, 1984 |
| Blom, 2007 | Pecheur, 1980 | Scoot, 1992 | Wickberg, 1996 |
| Carrington, 1979 | Reynolds, 1999 | Shamsei, 2008 | |
| Martin, 2001 | Ross, 1985 | Sirey, 2005 | |

*Removed because paper was not published in English*

| | | | |
|---|---|---|---|
| Hautzinger, 1996 | Lopez, 2004 | Maldonado- | Maldonado- |
| Krampen, 1997 | | Lopez, 1982 | Lopez, 1984 |

*Removed because treatment was not exclusively individual / one-to-one (e.g. couples)*

| | | |
|---|---|---|
| Beach, 1992 | Leff, 2000 | Milgrom, 2005 |
| Bodenmann, 2008 | McLean, 1979 | Scott, 1990 |

*Removed for other reasons*

Chaput, 2008 – not acute treatment phase
Ransom, 2008 – intervention not always delivered face-to-face
Banken, 1993 – could not identify critical study data (i.e., study N)
Cullen, 2002 – could not identify critical study data (i.e., study N)

## Appendix 2

| Study | ID # | Condition Descriptor | Study | ID # | Condition Descriptor |
|---|---|---|---|---|---|
| Alexopolous, 2003 | 1 | Problem-solving | Lesperance, 2007 | 36 | Interpersonal + citalopram |
| Alexopolous, 2003 | 2 | Supportive | Lesperance, 2007 | 37 | Interpersonal + placebo |
| Beck, 1985 | 3 | CT | Lustman, 1998 | 38 | CT |
| Beck, 1985 | 4 | CT + amitryptiline | Macaskil, 1996 | 39 | RET + lofepramine |
| Blackburn, 1997 | 5 | CT | McKnight, 1992 | 40 | CT |
| Burnand, 2002 | 6 | Dynamic + clomipramine | Misri, 2004 | 41 | CT + paroxetine |
| Castonguay, 2004 | 7 | Integrative CT | Mohr, 2001 | 42 | CT |
| Conradi, 2007* | 8 | Brief CT + DRPP | Murphy, 1984 | 43 | CT |
| Constantino, 2008 | 9 | CT | Murphy, 1984 | 44 | CT + ADM |
| Constantino, 2008 | 10 | Integrative CT | Murphy, 1984 | 45 | CT + placebo |
| Cooper, 2003 | 11 | CT | Murphy, 1995 | 46 | CT |
| Cooper, 2003 | 12 | Non-directive counseling | Mynors-Wallis, 1995 | 47 | Problem-solving |
| Cooper, 2003 | 13 | Dynamic | Mynors-Wallis, 2000 | 48 | Problem-solving (GP) |
| David, 2008 | 14 | REBT | Mynors-Wallis, 2000 | 49 | Problem-solving (RN) |
| David, 2008 | 15 | CT | Mynors-Wallis, 2000 | 50 | Problem-solving + ADM |
| De Jonghe, 2001 | 16 | Analytic + ADM | O'Hara, 2000 | 51 | Interpersonal |
| De Jonghe, 2004 | 17 | Supportive Dynamic | Padfield, 1976 | 52 | Positive reinforcement |
| De Jonghe, 2004 | 18 | Supportive Dynamic + ADM | Padfield, 1976 | 53 | Behavioral |
| Dekker, 2008 | 19 | Brief Dynamic | Rush, 1977 | 54 | CT |
| DeRubeis, 2005 | 20 | CT | Salminen, 2008 | 55 | Interpersonal |
| Dimidjian, 2006 | 21 | Behavioral | Schulberg, 1996 | 56 | Interpersonal |
| Dimidjian, 2006 | 22 | CT | Scott, 1997 | 57 | Brief CT + TAU |
| Dozois, 2009 | 23 | CT + ADM | Sloane, 1985 | 58 | Interpersonal |
| Gallagher, 1982 | 24 | Behavioral | Swartz, 2008 | 59 | IPT-MOMS |
| Gallagher, 1982 | 25 | CT | Teasdale, 1984 | 60 | CT + TAU |
| Gallagher, 1982 | 26 | Insight-oriented | Thompson, 1987 | 61 | CT |
| Hollon, 1992 | 27 | CT | Thompson, 1987 | 62 | Brief dynamic |
| Hollon, 1992 | 28 | CT + imipramine | Thompson, 1987 | 63 | Behavioral |
| Jacobson, 1996 | 29 | Behavioral + AT | Thompson, 2001 | 64 | CT |
| Jacobson, 1996 | 30 | Behavioral | Thompson, 2001 | 65 | CT + desipramine |
| Jacobson, 1996 | 31 | CT | Van Schaik, 2006 | 66 | Interpersonal |
| Jarrett, 1999 | 32 | CT | Vitriol, 2009 | 67 | Brief IPT |
| Keller, 2000 | 33 | CBASP | Wiles, 2008 | 68 | CT |
| Keller, 2000 | 34 | CBASP + nefazodone | Wright, 2005 | 69 | CT |
| Laidlaw, 2008 | 35 | CT | | | |