Clinical Psychology Review 31 (2011) 1304–1312

Contents lists available at SciVerse ScienceDirect

Clinical Psychology Review

Evidence-based treatments for depression and anxiety versus treatment-as-usual: A meta-analysis of direct comparisons ☆

Bruce E. Wampold [a,b,*], Stephanie L. Budge [c], Kevin M. Laska [a], A.C. Del Re [a,d], Timothy P. Baardseth [a], Christoph Flűckiger [a,e], Takuya Minami [a], D. Martin Kivlighan II [a], Wade Gunn [a]

[a] University of Wisconsin, Madison, USA
[b] Forskningsinstituttet, Modum Bad Psychiatric Center, Norway
[c] University of Louisville, USA
[d] VA Palo Alto Health Care System & Stanford University School of Medicine, USA
[e] University of Bern, Switzerland

ARTICLE INFO

Article history:
Received 11 February 2011
Received in revised form 7 July 2011
Accepted 8 July 2011
Available online 17 September 2011

Keywords:
Evidence-based treatments
Meta-analysis
Treatment-as-usual
Anxiety
Depression

ABSTRACT

Objective: The aim of this study was to examine the relative efficacy of evidence-based treatments (EBTs) versus treatment-as-usual (TAU) in routine care for anxiety and depression in adults.
Method: A computerized search of studies that directly compared an EBT with a TAU was conducted. Meta-analytic methods were used to estimate effectiveness of EBTs relative to TAU and to model how various confounding variables impacted the results of this comparative research.
Results: A total of 14 studies were included in the final meta-analysis. There was significant heterogeneity in the TAU conditions, which ranged from unknown and/or minimal mental health treatment to psychotherapeutic interventions provided by trained professionals. Although the effect for EBT vs. TAU was significantly greater than zero, the effect for EBT vs. TAUs that were psychotherapeutic interventions was not statistically different from zero.
Conclusions: Heterogeneity of TAU conditions in this meta-analysis highlight the importance of clarifying the research questions being asked when investigating and drawing conclusions from EBT–TAU comparisons. Researchers need to clarify if they are comparing an EBT to psychotherapeutic services in routine care or to minimal mental health services. Extant research on EBT versus TAU reveals that there is insufficient evidence to recommend the transportation of EBTs for anxiety and depression to routine care, particularly when the routine care involves psychotherapeutic services.

Contents

☆ The views expressed here are those of the authors and do not necessarily represent those of the Department of Veterans Affairs.

* Corresponding author at: Department of Counseling Psychology, University of Wisconsin, 335 Educational Building, 1000 Bascom Mall, Madison, WI 53706, USA. Tel.: +1 608 262 2878; fax: +1 608 265 3347.
E-mail address: wampold@education.wisc.edu (B.E. Wampold).

0272-7358/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cpr.2011.07.012

1. Introduction

Anxiety and depressive disorders are among the most common psychiatric diagnoses in the United States. Lifetime prevalence for Major Depressive Disorder (MDD) is 16.6%, resulting in an economic cost to society of roughly $83.1 billion (Greenberg et al., 2003; Kessler, Berglund, Demler, Jin, & Walters, 2005). Instead of being seen as an acute mental health concern, MDD is seen as a chronic disorder that is projected to be the second overall cause of disability by the year 2020 (Murray & Lopez, 1997). In addition, 28.8% of the general population has been diagnosed with an anxiety disorder at least once in their lifetime (Kessler et al., 2005). The total yearly economic burden of anxiety disorders is estimated to be around $46.6 billion (Rosenblatt, 2010).

In response to the pervasive nature of depression and anxiety in the general population, psychotherapy researchers have focused on identifying efficacious treatments for these disorders (Westen & Morrison, 2001). Over the past four decades, this emphasis has led to identification of Empirically Supported Treatments (ESTs), which was an official designation of Division 12 of the American Psychological Association's Task Force for the Promotion and Dissemination of Psychological Procedures (Chambless & Hollon, 1998; Task Force on Promotion and Dissemination of Psychological Procedures, 1995). The field has moved to "evidence-based treatments" (EBTs) to describe therapies that have been demonstrated to be efficacious in randomized clinical trials (RCTs), the design that is the "gold standard" for establishing the viability of a treatment (Westen, Novotny, & Thompson-Brenner, 2004). Several treatments for anxiety and depression have been identified as evidence-based, establishing the value of these treatments for anxiety and depression (e.g., Interpersonal Therapy for depression, Cognitive Behavioral Therapy for generalized anxiety disorder, and Prolonged Exposure for post traumatic stress disorder; see APA Division 12, Society of Clinical Psychology, website for a complete list, http://www.div12.org/PsychologicalTreatments/index.html).

Once efficacy of a treatment has been established in controlled settings (i.e., RCTs), the next logical step to improve the quality of care is to test the effect of the treatment on anxiety and depression in routine care. The primary method for investigating the effect of EBTs in naturalistic settings for anxiety and depression involves a direct comparison of EBTs with services that are currently being delivered in routine care, which are often referred to as treatment-as-usual (TAU). Implementation of this strategy has resulted in mixed findings. On the one hand, for example, Meuser et al. (2008) compared cognitive behavioral treatment (CBT) with TAU for PTSD and found that patients in the CBT condition improved more than patients in TAU. Similarly, Grote et al. (2009) found that EBT outperformed TAU for the treatment of depression. In contrast, Cuijpers, van Lier, van Straten, and Donker (2005) directly compared CBT to TAU for depressed patients and concluded, "On average, patients in both test conditions improved significantly from baseline to posttest, and no significant difference was found between the conditions" (p. 137). These contradictory findings may be attributable to sampling error or systematic differences between studies (e.g., levels of training, supervision, treatment dose, type or nature of the TAU; see Spielmans, Gatlin, & McFall, 2010 for a discussion in the youth treatment literature).

One well-established method of synthesizing seemingly inconsistent results from a corpus of studies is meta-analysis (Cooper, Hedges, & Valentine, 2009; Hunt, 1997; Mann, 1994) Recently, Weisz, Jensen-Doss, and Hawley (2006) conducted a meta-analysis in which EBTs and TAU were directly compared for child and adolescent populations. Their findings indicated that EBTs were more effective than TAU, but they noted aspects of the research design that posed significant confounds: "EBT descriptions were significantly more likely than [TAU] descriptions to note the use of pretherapy training, treatment manuals, and adherence checks" (p. 681). Moreover, several studies in the meta-analysis included EBT therapists who had specialized training and expertise (e.g., trained extensively in the EBT for particular disorders or had training about the disorder itself) vis-à-vis the TAU therapists who received no training. In addition, the heterogeneity of TAU conditions was problematic:

> We considered studies in which participants received medications in addition to therapy, studies in which the EBT was administered in addition to UC [usual care, which is used synonymously with TAU], studies in which a psychotherapy placebo was administered in addition to UC, and studies in which UC involved case management services (e.g., probation and referral) that may or may not have included significant doses of psychotherapy (Weisz et al., 2006, p. 683).

To model the confounds noted in Weisz et al. (2006), Spielmans et al. (2010) conducted a replication of the Weisz et al. meta-analysis and found that EBTs and TAU had similar outcomes when the confounds were controlled. Essentially, Spielmans et al. found that many design operations in the studies favored the EBT (e.g., a greater dose of therapy for EBT patients or supervision of EBT therapists only), which attenuated the magnitude of the effect.

Heterogeneity of TAU conditions in the Weisz et al. meta-analysis highlights the importance of clarifying the research questions being asked when investigating and drawing conclusions from EBT–TAU comparisons. Presumably, researchers are interested in two different research questions: (a) does the implementation of an EBT into existing mental health systems that are delivering psychotherapeutic services by trained professionals improve outcomes (psychotherapy TAU)? or (b) does the implementation of an EBT into existing mental health services, where little or no psychotherapy is provided, improve outcomes (non-psychotherapy TAU)?

A comparison of an EBT to a psychotherapy TAU delivered by mental health therapists with comparable training to the EBT therapists (psychotherapy TAU) provides evidence about the viability of transporting EBTs into existing mental health systems that deliver psychotherapy, whereas EBT–non-psychotherapy TAU comparisons do not. The latter comparison, that is, the comparison of an EBT with other types of services that do not involve psychotherapy (e.g., a referral to a primary care physician), provides evidence related to the question of whether instituting EBTs improves the quality of care of mental health services that do not provide psychotherapy. Consequently, the latter does not provide evidence about the viability

of transporting EBTs into existing mental health systems that deliver psychotherapy. Interpretation of EBT–TAU comparisons requires an in-depth examination of the nature of the TAU to understand what conclusions can be drawn from the evidence.

The purpose of the present study was to examine the relative efficacy of EBTs when compared to TAU for anxiety and depression in adults, while examining possible confounds, including the type of services provided in the TAU (i.e., the proximity of the TAU to psychotherapeutic interventions). We hypothesized that EBT would be superior to TAU (i.e., the overall effect would be significantly greater than zero), but that the various confounds would moderate the effect. In particular, we hypothesized that the EBT–TAU effect would be smaller when the TAU was psychotherapy than when the TAU involved little or no psychotherapy.

2. Method

2.1. Inclusion criteria

In order to be selected for this analysis, studies needed to meet the following criteria: (a) the study directly compared an EBT with TAU; (b) participants were diagnosed with either a depressive or anxiety disorder; (c) the study did not primarily focus on reduction of suicidality; (d) the treatment modality was either group or individual; (e) the article reporting the results was published between the years 1995 and 2009, inclusive, in peer reviewed English language journals; (f) the investigation was randomized; (g) and the article reported quantifiable outcome measures. If medications were provided as part of treatment, either the protocol or access to medication had to be equal in the two conditions. Studies were excluded that examined couples treatment, trials focused on children and adolescents, follow up studies, trials that included "collaborative care" or integrated care (e.g., primary care physician and therapist both gave part of treatment), additive trials (EBT and TAU vs. TAU alone), and at-risk or prevention studies.

Because the focus of this study was to examine the relative efficacy of EBTs vs. TAU for depression and anxiety, we wanted to include EBTs that have been clearly established as efficacious for these disorders, were complete treatment packages (not components or adaptations thereof, e.g., a treatment which only uses "cognitive strategies"), and have been tested in controlled clinical trials. Accordingly, treatments had to satisfy all of the following criteria to meet the eligibility criteria as an EBT for this study: (a) use a manual or have been based on an already developed manual, (b) be established as an EBT for the disorder being treated (e.g., CBT for depression, prolonged exposure for post-traumatic stress disorder), (c) include six or more sessions of face-to-face therapy (no telephone, tele-mental health, etc.), (d) be delivered by a mental health provider with at least a master's degree in a clinical/counseling/social work or related field (e.g., no depression care managers, paraprofessionals), and (e) be individually tailored to the patient (i.e., no self help books, progressive muscle relaxation, etc.).

The term TAU, as used scientifically, is heterogeneous. Over the years, researchers have defined TAU to include anything from a range of unknown and unspecified treatments to active psychotherapies for particular disorders. We decided to include the variety of TAU conditions as individually defined by the respective authors, but to code for the variability in these conditions. Given our interest in examining the nature of the TAU as a comparison treatment, for this study TAU only had to (a) be an intervention of some kind; and (b) be identified by the authors as, "treatment as usual", "standard care", or "usual care" but (c) not involve a condition where the therapists in the TAU were proscribed from certain actions. The latter condition ruled out studies in which TAU therapists were instructed or trained to avoid using certain techniques they might ordinarily use (see e.g., Feske, 2008).

2.2. Literature search

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Standards (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). The final search was conducted on May 7, 2010. A systematic examination of several major databases was completed, including Medline, Academic Search, PsycARTICLES, PsycINFO, PsycCRITIQUE, ERIC, Social Work Abstracts, SocIndex, HealthSource, Nursing/Academic Edition, and CINAHL. Studies were identified through the combination of any of the keywords: *Depression*, *Major Depressive Disorder*, *MDD*, *Dysthymic*, *Anxiety*, *Generalized Anxiety Disorder*, *Social Phobia*, *Specific Phobia*, *PTSD*, *Post Traumatic Stress Disorder*, *Panic Disorder*, *OCD*, *Obsessive Compulsive Disorder*, *Psychotherapy* along with the TAU key words: *Treatment-as-Usual*, *TAU*, *Usual Care*, *Care as Usual*, *Standard Care*.

A total of 2554 initial abstracts were included in the first round of examination (see Fig. 1). Each record was screened in greater detail, which resulted in 149 potentially relevant studies. Next, two doctoral students reviewed the method section (i.e., were blind to the authors, introduction, results, and conclusions) of each record and excluded those which did not meet inclusion criteria. If disagreements arose at this stage, the study was temporarily retained and discussed during the qualitative synthesis stage as outlined by the PRISMA Statement. Twenty-nine studies were included in the final qualitative synthesis. Finally, after rigorous application of the inclusion/exclusion criteria and determining whether the studies contained adequate information for the meta-analysis, 14 studies were included in the final quantitative synthesis.

2.3. Coding study quality

Based on some of the issues raised by Weisz et al. (2006) and Spielmans et al. (2010), the following variables were coded.

2.3.1. Sample size

For each treatment group (EBT and TAU), sample size was recorded for those who were randomized and placed into each group (including intent-to-treat, ITT).

2.3.2. Number of therapists

When provided, the number of therapists for each treatment group was documented.

2.3.3. Dose hours for treatment

We coded the total number of treatment hours for the EBT above and beyond the TAU. This was calculated by multiplying the average number of treatment hours for the specific treatment (e.g., 12 sessions of CPT for PTSD), by the number of individuals who participated in at least one session for the EBT, and subtracting the same values for the TAU. This value could be negative (i.e., when TAU had more dose hours). Based on this number, a study was coded as a "1" if the dose hours for the EBT were greater than the TAU, "0" if both treatments received equal dose. A code of "N/A" was assigned if information was not available.

2.3.4. Training

We calculated number of hours that were provided for training in the interventions. No study provided training for the TAU, therefore the total hours for training were exclusively in favor of the EBT. Based on these values, we created a dichotomous variable, whose value was "1" if the training was provided for the EBT but not for the TAU and "0" if no training was provided in both conditions. If information was not available, it was coded "N/A."

2.3.5. Supervision

We also calculated number of hours of individual supervision provided to the therapists for the EBT above and beyond the TAU. Similar

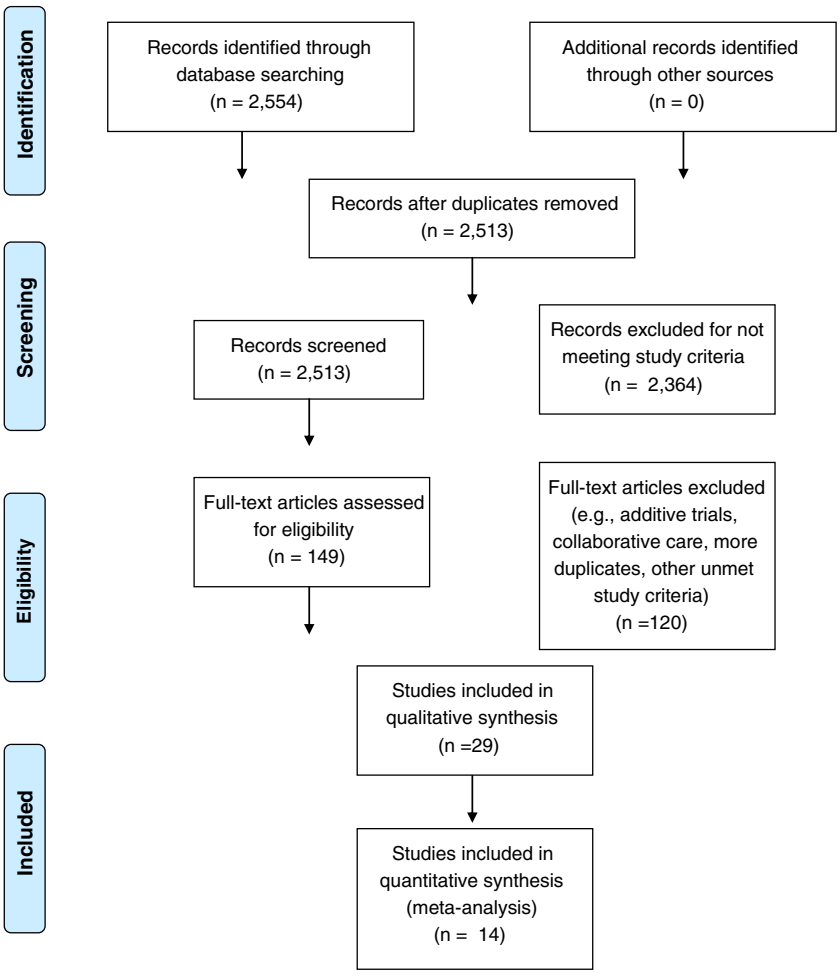

Fig. 1. Identification, screening, eligibility, and inclusion of studies.

to training, none of the studies indicated that supervision was provided for the TAU therapists. If supervision was provided in a group modality, we coded these hours in a separate category than for individual supervision hours. If the information was not provided, the variable was coded as missing (i.e., "N/A").

2.3.6. Adherence

If a study used adherence checks to measure the fidelity with which the therapists were providing the intervention, the study was given a "1". If they explicitly stated that adherence was not checked, this variable was coded as a "0". If the information was not available, the variable was coded as "N/A".

2.3.7. Allegiance

We used a six-point rating scale in order to determine researchers' allegiance to the treatments within the study. The highest level of allegiance was coded if the EBT was developed by the author(s) and if they supervised or trained the therapists ("5"). The next level of allegiance was if the EBT was developed by the authors, but they did not train or supervise the therapists ("4"). Allegiance was coded a "3" if the treatment was advocated by one of the authors and they also supervised/trained the therapists. If the treatment was advocated by the authors but they did not train or supervise therapists, it received a "2". In addition, the same code was used if the author showed no advocacy for the treatment, but provided better trained or more experienced therapists for one treatment over another. The next level of allegiance was coded "1" if the EBT was more fully explained in the introduction and/or method section than the alternative. The lowest level of allegiance was coded "0" if there was no apparent advocacy over one treatment than the other.

2.3.8. TAU proximity to psychological intervention (TAU PPI)

We coded each study based on the degree to which the TAU resembled a genuine psychological intervention. To directly address the question of whether implementation of an EBT in a service delivery system already providing psychotherapy improves outcomes, the TAU would need to be a *psychotherapeutic intervention* (i.e., a treatment the therapists believed had a cogent rationale) delivered by a trained mental health provider with no proscriptions (i.e., requiring that TAU therapists abstain from any actions they ordinarily would use). Moreover, training, supervision, and dose hours would need to be matched to not advantage the EBT. Such studies received a "0" if these conditions were met. The study was coded as "1" if all patients in the TAU modality received a psychotherapeutic treatment delivered by a mental health professional, doing what they usually do with no proscriptions for psychotherapeutic care, *but* where the EBT therapists received additional training and supervision or the EBT patients received a greater dose of treatment. To receive a code of "2", the TAU patients received a treatment of some kind (e.g., medical, psychotherapy, etc.), which was tracked by the researchers in order to ensure that the patients had actually received *some* kind of treatment, but it was still unclear exactly what type and if all participants received the same treatment. In addition, the therapists providing treatment in the TAU had to be different from the therapists providing the EBT. For a code of "3", the TAU patients received a treatment (e.g., medical, psychotherapy, etc.), which was tracked by the

researchers, but was provided by EBT therapists (e.g., therapists were crossed and the therapists knew they were delivering a treatment different from the EBT). The last code "4" was given to a study if the TAU services were suggested or mentioned by the researchers to the patients who were randomized to the TAU (e.g., it was suggested they could present to their primary care physician, PCP, if they desired service), but the services were not tracked, reported, or discussed in the study.

As will be seen, 12 of the 14 studies were either coded as 1 (3 studies) or 4 (9 studies). The former category (viz., TAU PPI = 1) included studies where the TAU participants clearly received psychotherapy. On the other hand, the preponderance of participants in TAU in the latter instance (viz., TAU PPI = 4) did not receive any psychotherapeutic or indeed any mental health treatment. The difference in type of treatment received had direct implications for our analysis. For example, take the TAU that involves making a referral to a PCP. First, it is not clear how many participants would actually present to a PCP for anxiety and depression; even if they did present, less than *four* percent of such patients are referred for mental health services, and of those referred less than 40% are referred to a psychologist (Forrest, Nutting, Starfield, & Von Schrader, 2002). Consequently, for the moderator analysis TAU PPI was dichotomized into two groups, those who received psychotherapy (viz., TAU PPI = 1) and those who most likely did not (TAU PPI = 4).

2.4. Coding process

Four doctoral students were assigned to teams of two and coded records independently. Disagreements were discussed in greater detail before bringing them to the research team for further qualitative analysis. Mean interrater agreement was calculated as percentage agreement, as there was a mix of dichotomous, ordinal, categorical, and continuous variables (agreement if conclusion was the same, e.g., EBT does greater than TAU dose). The mean interrater agreement was .94, implying that the codes were sufficiently well defined and that the coders were able to adequately apply the definitions.

2.5. Analytic strategy

For each study, we extracted means, standard deviations, and number of participants at the measurement nearest posttreatment. These data were used to calculate effect size statistics (Cohen's d) in order to compare the impact of EBT vs. TAU in standard deviation units across studies. Information was used from ITT, if available (viz., Addis et al., 2004; Difede et al., 2007; van Schaik et al., 2006); otherwise, completer samples were used. Measures were segregated into primary (i.e., measures that assessed the primary diagnosis) and secondary. However, several studies did not report secondary measures, and consequently we present only the results of the meta-analysis for primary measures. The within studies effect size for primary measures were aggregated based on an intercorrelation of .5 (Del Re & Hoyt, 2010; Hedges & Olkin, 1985; see Wampold et al., 1997 for fuller explanation).

For each study, mean difference between the EBT and the TAU was standardized in the usual fashion and corrected for bias and the standard error calculated (Hedges & Olkin, 1985). We assumed that the studies in this meta-analysis were sampled from a population of studies and consequently a random effects restricted maximum-likelihood estimator was used (Viechtbauer, 2005). The analysis was conducted using the R statistical software package for meta-analysis 'MAd' (Del Re & Hoyt, 2010). The test of the hypothesis that EBT is superior to TAU involved an unconditional model (not conditioned on study level variables, i.e., moderators). The formula is

$$d_j = \delta_j + v_j^*,$$

where d_j is the estimate of the effect size for study j, δ_j is the true effect for study j, and $v_j^* = v_j + \tau$, where the variance of the within-study errors, v_j, are known and the between-study errors, τ^2, are unknown and estimated based on the studies included in the analysis. Homogeneity was tested with the H statistic, which indexes the deviation of the sampled effects from the grand mean, weighted by the inverse of the variance (Hedges & Olkin, 1985, Raudenbush & Bryk, 2002). H has an approximate χ^2 distribution with $k-1$ degrees of freedom, where k is the number of studies aggregated.

As will be discussed in the Results section, the majority of studies either failed to provide sufficient information to code moderator variables or the distribution was too highly skewed (i.e., most studies had the same value of the moderating variable) and therefore tests of moderation were precluded. However, it was possible to contrast studies that used a TAU that involved psychotherapy (TAU PPI = 1) with TAU that most likely did not involve psychotherapy (TAU PPI = 4), using a between groups test using the model:

$$\delta_j = \gamma_0 + \gamma_1(\text{TAU PPI}) + v_j^*$$

in which TAU PPI was dummy coded (0 if TAU PPI = 1, 1 if TAU PPI = 4) and where γ_0 is the expected effect for a study when the TAU PPI was equal to 1 and γ_1 is the expected difference between TAU PPI = 1 and TAU PPI = 4. The test of the coefficient γ_1 provides the one degree of freedom between groups test within the random effects model context and will be reported as such.

3. Results

Results are presented in two parts: (a) study variables and (b) meta-analytic synthesis.

3.1. Study variables

The values of moderating variables are found in Table 1. Two conclusions from this table are readily apparent. First, for most aspects of study quality, the information needed to assess the overall comparison of EBT and TAU was unreported. When it was reported, it was clear the design favored the EBT. For example, EBT was favored in 11 of the 13 studies that reported information on treatment dose. As well, only seven studies reported data on therapist training. Furthermore, more training was reported in the EBT condition than in the TAU in all seven of these studies. Similarly, eight studies reported information about supervision and all of those reported supervision in the EBT condition but not the TAU condition. Of the 11 studies that reported adherence, nine conducted adherence checks for the EBT. When there was sufficient information to code researcher allegiance, there was a distinct allegiance to the EBT.

Second, TAU PPI for the majority of studies suggests only a few comparisons between EBT and a TAU condition that clearly involved psychotherapy. None of the 14 studies were coded as a zero on TAU PPI, which would have indicated that the TAU was a psychotherapeutic intervention with dose hours, training, and supervision matched to what was provided in the EBT. Three studies were coded as "1", indicating the TAU was psychotherapeutic treatment, but the EBT therapists had additional training and supervision. Two additional studies were coded as a "2" or "3", indicating that the researchers monitored the TAU services and the patients received some type of mental health service, although it was not demonstrably a psychotherapy designed for the disorder. In the one study coded "3", the TAU was provided by EBT therapists who had a clear allegiance to the EBT. In the remaining nine studies coded 4, it was not clear whether or not the TAU patients received any treatment, but it was clear that the treatment received was not predominantly psychotherapy intended to be therapeutic. In many of these cases, patients were referred to

Table 1
Moderator codes with study level variable information.

Authors	TAU PPI	# EBT therapists	# TAU therapists	EBT–TAU dose hours	EBT dose hours>TAU (raw)	EBT train hrs>TAU	EBT train hrs>TAU (raw)	EBT indv. sup hrs	EBT grp sup hrs	Adher checks	EBT alleg
Addis et al. (2004)[a]	1	7	6	1	64	1	36	N/A	42	1	4
Burns et al. (2007)	4	N/A	N/A	1	366	1	N/A	N/A	N/A	N/A	N/A
Cuijpers, van Lier, van Straten and Donker (2005)	1	N/A	N/A	0	−27	1	N/A	N/A	N/A	N/A	N/A
Difede et al. (2007)[a]	4	N/A	N/A	1	165	1	N/A	N/A	N/A	1	N/A
Grote et al. (2009)	4	2	N/A	1	425	1	N/A	320	0	1	4
Kingston, Dooley, Bates, Lawlor and Malone (2007)	3	2	N/A	1	20	1	N/A	N/A	N/A	0	3
Laidlaw et al. (2008)	4	N/A	N/A	1	160	1	N/A	N/A	N/A	0	3
Marcus, Marquis and Sakai (1997)	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	1	N/A
Meuser et al. (2008)	2	7	N/A	1	628	1	N/A	1456	0	1	3
Miranda, Chung and Green (2003)	4	6	N/A	1	412	1	N/A	N/A	N/A	1	4
Stanley et al. (2009)	4	5	N/A	1	331	1	100	N/A	N/A	1	3
Van Schaik et al. (2006)[a]	4	15	N/A	1	690	1	16	N/A	36	N/A	3
Wagner, Zatzick, Ghesquiere and Jurkovich (2007)	4	N/A	N/A	1	21	1	N/A	N/A	N/A	0	N/A
Ward et al. (2000)	4	26	N/A	0	−301	1	N/A	53	N/A	1	N/A

Note: See section "Coding study quality" for description of each variable. N/A = information not available.
TAU PPI = Treatment-as-usual proximity to psychotherapeutic intervention.
TAU PPI 1 = the patients in this TAU modality received *psychotherapy* by a mental health professional, doing what they usually do with no proscriptions for psychotherapeutic care.
TAU PPI 2 = the TAU patients received a treatment of some kind (e.g., medical, psychotherapy, etc.), which was tracked by the researchers in order to ensure that the patients had actually received *some* kind of treatment, but unclear if all participants received the same treatment. In addition, the therapists providing treatment in the TAU had to be separate than the therapists providing the EBT.
TAU PPI 3 = the TAU patients received a treatment of some kind (e.g., medical, psychotherapy, etc.), which was tracked by the researchers, but was provided by EBT therapists (e.g., therapists were crossed).
TAU PPI 4 = if the TAU services were suggested or mentioned by the researchers to the patients randomized to the TAU, but the services were not tracked, reported, or discussed in the study.
[a] Designates intent-to-treat samples; all others completers.

their PCP, but it was unclear whether any of the patients presented for treatment. Moreover, even if they had presented to their PCP, it is unlikely they received any psychotherapy (Forrest et al., 2002).

3.2. Meta-analysis

The studies, their effects on primary measures, and TAU PPI are displayed in Fig. 2, as a forest plot.

The overall effect for EBT versus TAU was $d = .45$, which was significantly larger than zero ($p < .01$) and represents a medium sized effect in favor of EBT. However, the H statistic was 30.78, which, when compared to a χ^2 distribution with $df = 13$, resulted in rejection of the null hypothesis that effects are homogeneous. That is, there was evidence of variability among the studies that was not likely due to sampling error ($p < .01$; $I^2 = .58$, indicating that 58% of the variability in effects is due to true differences among the studies). Thus, there are factors that may account for these differences. It was our goal to model how the various design confounds would account for this variability, but because of the issues with the moderators discussed above, this was not possible (see Table 1).

For the between groups test, the effect for studies in which the TAU was unlikely to be a psychotherapy was .50, which was significantly greater than zero ($k = 9$, $p < .01$), whereas the effect for studies in which the TAU was clearly a psychotherapy was .33, which was not significantly different from zero ($k = 3$, $p = .06$). Although the direction of the difference in effects was in the expected direction (i.e., larger when the TAU was not a psychotherapy), the difference between the two coefficients was not significantly different from zero ($df = 1$, $p = .46$).

4. Discussion

Given prevalence of anxiety and depression in the United States, developing strategies for improving quality of services for these disorders is clearly an imperative. One method for improving mental health services is to transport EBTs established in RCTs to practice settings. Such a decision at the service level, or enactment of a policy that would require such action, should be based on scientific evidence. Comparisons of EBTs with TAU in routine care would provide useful evidence for those who are entrusted to make mental health management decisions and for those who develop and implement mental health policy. It was the purpose of the present study to assess the status of EBT versus TAU research for treatment of depression and anxiety in adults.

One major result of this review is that quality of the comparisons between EBT and TAU is not sufficient to make a strong conclusion about whether EBT is more beneficial than TAU for the treatment of anxiety and depression. For the most part, studies that compared an EBT to a TAU did not report sufficient information for consumers of the research to understand or utilize the research findings. When evaluating differential efficacy of these two modalities of treatment, the comparison between the EBT and the TAU must not be confounded by variables that would be more advantageous to one of the treatments. If the EBT therapists receive supervision, for example, while the TAU therapists do not, superiority of the EBT may be due to the role of supervision in improving outcomes rather than the EBT being more effective than TAU in the absence of supervision. Unfortunately, few studies reported information about training, supervision, treatment dose, adherence checks, or other aspects of delivering the EBT and TAU.

Researchers are interested in two different types of comparisons: EBT vs. psychotherapy TAU or EBT vs. non-psychotherapy TAU. In our meta-analysis, only three out of 14 studies involved TAUs that were demonstrably and consistently psychotherapeutic treatments. Patients in the other 11 studies were either referred to a PCP or provided an unspecified and untracked treatment. Although seeing a PCP may represent practice in naturalistic settings, and in that case, provide data for EBT vs. non-psychotherapy TAU comparisons, the

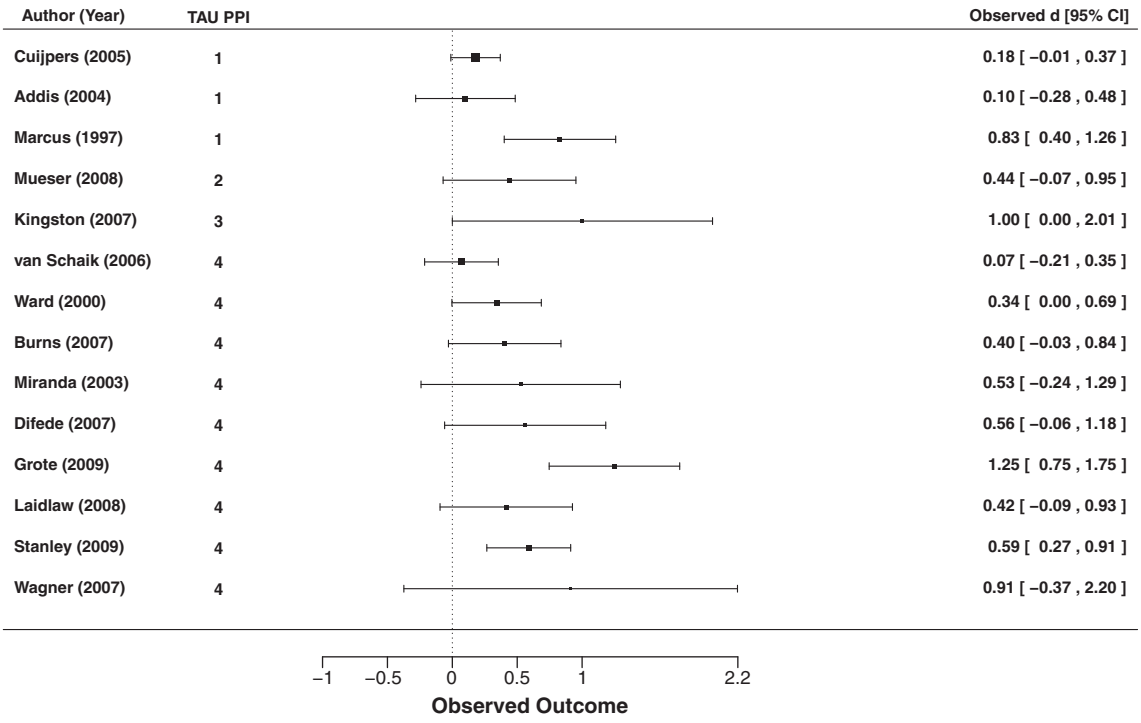

Fig. 2. Forest plot of effects, in order of TAU PPI.

studies in this analysis typically did not report whether they followed up with patients to determine if they presented to the PCP for their complaint, let alone whether they received any treatment. For example, patients in one of the TAU conditions (viz., Ward et al., 2000) were purposefully not referred to treatment—"General practitioners treated patients in this group according to their usual practice, but were asked to refrain from referral for psychological interventions unless this was imperative" (p. 1384). The modal TAU in our study involved no verified psychological service. Nevertheless, the EBTs produced superior outcomes to non-psychotherapy TAUs and thus it appears that implementing EBTs into a system of care that does not involve to a significant degree psychotherapeutic services would generally improve the quality of care.

Assessment of studies during the coding stage for this meta-analysis revealed that the descriptor "TAU" was occasionally used to designate treatments that did not represent any treatment being delivered in routine care. For example, therapists in the Feske (2008) trial were trained in both Prolonged Exposure (PE) and a treatment labeled TAU, for which the therapists were *proscribed* from including specific types of interventions (e.g., imaginal or in-vivo exposure elements of PE). Training therapists to prevent them from using certain therapeutic actions that are typically employed in their practice cannot logically be classified as a TAU.

The three studies that contained TAU conditions that were demonstrably psychotherapy, were still not comparable to the EBT in several ways. In each of these studies, the EBT therapists received additional training and supervision, which provided an advantage to the EBT condition. For example, therapists providing the EBT treatment in the Addis et al. (2004) study were provided a 2-day Panic Control Therapy (PCT) training, were assigned to two training cases, and were individually allotted 2 h of personalized phone consultation from an expert who examined audiotapes of their sessions. These therapists also received 1-hour, bi-weekly group supervision. On the other hand, the therapists in the TAU group received no training, no expert consultation, and no supervision. The EBT in this study was slightly more effective than TAU, but these results must be taken in the context of the advantage provided to the EBT in terms of training, consultation, and supervision. To be conclusive, the TAU therapists would have been provided comparable training and supervision of an equal dose. In any event, the cost/benefit ratio of transporting the EBT to routine care must be considered as the costs of the training, consultation, and supervision in this study and other studies are significant.

Despite the overall meta-analytic result indicating that EBT was more effective than TAU, these findings must be interpreted in the context of several factors. First, the effects for the 14 studies were heterogeneous which suggests that significant study characteristics accounted for the effects obtained. Unfortunately, most study level variables could not be investigated in this meta-analysis. Second, TAUs used in the studies were predominantly "treatments" that did not include any psychotherapy. However, the effect for studies involving non-psychotherapy TAU was relatively large and statistically significant, indicating that EBTs are more effective than services provided to many patients such as a referral to a PCP. For the three studies that included a TAU that was demonstrably a psychotherapy, the effect for EBT vs. TAU was not significantly greater than zero, which indicates insufficient evidence to conclude that EBT is more effective than psychotherapies provided in routine psychological service for depression and anxiety.

4.1. Limitations

There are several limitations that should accompany the conclusions. The small effect between EBT and psychotherapy TAUs could be due to the fact that TAU therapists were already providing an EBT. Yet, this seems unlikely. Stiles, Barkham, Mellor-Clark, and Connell (2007) reported that therapists who indicated that they provided CBT in naturalistic settings did not have better outcomes than those who indicated that they provided humanistic or dynamic therapies. This result was discounted by Clark, Fairburn, and Wessely (2008) who claimed that therapists, even if they reported using an EBT, were failing to provide CBT as it was intended to be delivered:

> Some therapists are likely to 'label' their treatment as falling within a particular approach even if they do not follow the indicated, evidence-based procedures for treating the patient's problems within that approach. Such therapists may have essentially offered a placebo intervention in which non-specific factors (genuineness, warmth and empathy) were the main ingredients. Alternatively, they may have used procedures that are without a theoretical or empirical basis. ….In our experience, such misunderstandings [by practicing therapists] of what CBT comprises are by no means unusual. (p. 631)

Whether or not therapists are delivering EBTs with fidelity in routine care is irrelevant if transportation of an EBT creates little additional benefit, as it appears from the three studies designed to answer this question, particularly given the costs of such transportation.

A second limitation of this study was that a significant portion of the information regarding the moderating variables was not provided in the studies included in this meta-analysis. The significant heterogeneity among the effects indicated that between-study variability was substantial enough to consider the influence of moderating variables and casting doubt on any point estimate of the true effect. However, most moderating analyses were not conducted because either (a) studies lacked sufficient information to determine data for a moderator analysis, or (b) studies did not implement the aspects required to report such data. Despite the inability to conduct such analyses, we were able to control for TAU PPI, which indicated a non-significant difference between EBT and TAU treatments when TAU was psychotherapy. Unfortunately, no studies existed that compared EBT to psychotherapy TAU while controlling for training, supervision, dose, and other factors, despite demonstration that such confounds need to be considered (Spielmans et al., 2010; Weisz et al., 2006).

A third possible limitation to this meta-analysis was that there were too few studies to make confident conclusions about the results. It was expected that many more studies would fit the inclusion criteria to be included for coding and analysis. We were not only surprised by the lack of studies that fit the inclusion criteria, but also by the heterogeneous quality of the TAU conditions that were used as comparison groups. The amount of research required to establish the utility of transporting EBT into routine care for the treatment of anxiety and depression suggests that it is premature to make a recommendation that such transportation would improve the quality of care. Given the cost of such transportation, the scientific perspective is to avoid claiming that the transportation of EBTs for depression and anxiety would improve the quality of care until there is sufficient evidence to believe otherwise.

4.2. Implications

We present implications from this study in two parts: (a) methodological implications and (b) clinical implications. It is clear that when designing EBT–TAU comparisons, researchers need to control for confounding variables as mentioned previously in this study, and as discussed in both Weisz et al. (2006) and Spielmans et al. (2010). Aspects of research design such as equal treatment dose, nesting therapists within treatment, and offering equivalent training and supervision, provide better design characteristics in order to draw conclusions about EBT–TAU comparisons. In most of the studies included in this meta-analysis, it was possible that a patient in the TAU condition received little or no treatment, whereas a patient in the EBT condition received approximately 12 sessions by a therapist who received special training and received supervision during the course of therapy. As it stands from the results of this meta-analysis, it is impossible to know if the EBT treatments show greater effects due to the actual effectiveness of the treatment or because the therapists in that treatment had extra training/supervision, or because the EBT treatment was longer. The current state of the literature as reviewed in this study suggests that more rigorous methods are necessary before any definitive statements can be asserted regarding the differential efficacy of EBT–TAU. In Table 2 we list recommendations for controlling confounds in EBT–TAU comparisons.

Second, it is important that researchers indicate whether they are examining efficacy of an EBT compared to routine psychotherapeutic services or to minimal mental health treatment, such as that which might be obtained through a referral to a PCP. Ambiguity about the research question and what is meant by TAU will be misleading. Therefore, researchers must delineate the difference between the type of EBT–TAU comparison to ensure that policy makers have accurate evidence to make important health management decisions.

Anxiety and depressive disorders are the most prevalent psychological issues addressed in mental health treatment. Developing cost-effective methods of implementing psychosocial interventions in systems of care consequently becomes a critical health policy concern (Lazar, 2010). Clear evidence is needed for mental health policy makers to enact decisions regarding whether or not implementing an EBT in a clinic will improve care above what is already being provided. Currently, there is insufficient evidence to suggest that transporting an EBT to routine care that already involves psychotherapy will improve the quality of services. However, there does appear to be evidence that implementing EBTs into routine care that does not involve psychotherapy would improve the quality of care. Nevertheless, until evidence with regard to EBTs in routine care is more substantial, the clinician treating anxiety and depression, whether administering an EBT or another treatment, should ensure that their services produce outcomes similar to known benchmark for treatments of these disorders (see, e.g. Minami, Wampold, Serlin, Kircher, & Brown, 2007; Minami et al., 2008).

Table 2
Recommendations for controlling confounds in EBT–TAU comparisons.

TAU	• Must be an actual treatment (i.e., patients receive psychological services from a trained provider). • Therapists should have an allegiance to the TAU. • TAU should not be altered to exclude elements of the EBT (i.e., by proscribed actions of TAU therapists). • Format should be same as EBT (e.g., individual or family).
Setting	Setting should be the same for EBT and TAU.
Recruitment	All patients will have sought therapy in the usual way in the routine care setting.
Therapist characteristics	• Therapists should all be individuals who are trained to provide therapeutic interventions (e.g., psychologists, counselors, psychiatrists). • Therapists should be nested with treatment (i.e., deliver either EBT or TAU). • Therapist should have allegiance to the therapy they are delivering. • Therapist randomly assigned to treatments.
Caseload	Therapist caseload should be similar for both EBT and TAU.
Patients	Patients randomized to treatment.
Dose	• Dose differences should be avoided (either by unrestricted dose in both conditions or by restricting does to be the same). • Actual dose should be assessed.
Therapist training and supervision	• Pre-study training dose should be similar between both EBT and TAU (i.e., TAU therapists should get similar dose of training, in the nature of the disorder and skill building, say in the common factors). • Supervision should be same in both treatment modality (e.g., in-person, phone, individual, group). • If an expert is brought in for training/supervision for the EBT, an outside trainer/supervisor for TAU should also be provided.
Adherence	Should be recorded for both EBT and TAU.

References

Addis, M. E., Hatgis, C., Krasnow, A. D., Jacob, K., Bourne, L., & Mansfield, A. (2004). Effectiveness of cognitive behavioral treatment for panic disorder versus treatment as usual in a managed care setting. *Journal of Consulting and Clinical Psychology*, *72*, 625–635, doi:10.1037/0022-006X.72.4.625.

Burns, A., Banerjee, S., Morris, J., Woodward, Y., Baldwin, R., Proctor, R., et al. (2007). Treatment and prevention of depression after surgery for hip fracture in older people: Randomized, controlled trials. *The Journal of the American Geriatrics Society*, *55* (1), 75–80, doi:10.1111/j.1532-5415.2007.01016.x.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7–18, doi:10.1037/0022-006X.66.1.7.

Clark, D. M., Fairburn, C. G., & Wessely, S. (2008). Psychological treatment outcomes in routine NHS services: A commentary on Stiles et al. (2007). *Psychological Medicine*, *38*, 629–634, doi:10.1017/S0033291707001869.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd. ed.). New York: Russel Sage.

Cuijpers, P., van Lier, P. A. C., van Straten, A., & Donker, M. (2005). Examining differential effects of psychological treatment of depressive disorder: An application of trajectory analyses. *Journal of Affective Disorders*, *89*, 137–146, doi:10.1016/j.jad.2005.09.001.

Del Re, A. C., & Hoyt, W. T. (2010). MAd: Meta-analysis with mean differences. R package version 0.8. http://CRAN.R-project.org/package=MAd

Difede, J., Malta, L. S., Best, S., Henn-Haase, C., Metzler, T., Bryant, R., et al. (2007). A randomized controlled clinical treatment trial for world center attach related PTSD in disaster workers. *The Journal of Nervous Mental Disease*, *195*(10), 861–865, doi:10.1097/NMD.0b013e3181568612.

Feske, U. (2008). Treating low-income and minority women with posttraumatic stress disorder: A pilot study comparing prolonged exposure and treatment as usual conducted by community therapists. *Journal of Interpersonal Violence*, *23*, 1027–1040, doi:10.1177/0886260507313967.

Forrest, C. B., Nutting, P. A., Starfield, B., & Von Schrader, S. (2002). Family physicians' referral decisions: Results from the ASPN referral study. *The Journal of Family Practice*, *51*, 215–222, doi:10.1370/afm.732.

Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., et al. (2003). The economic burden of depression in the United States: How did it change between 1990 and 2000? *The Journal of Clinical Psychiatry*, *62*(12), 1465–1475.

Grote, N. K., Swartz, H. A., Geibel, S. L., Zuckoff, A., Houck, P. R., & Frank, E. (2009). A randomized controlled trial of culturally relevant, brief interpersonal psychotherapy for perinatal depression. *Psychiatric Services*, *60*, 313–321, doi:10.1176/appi.ps.60.3.313.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego: Academic Press.

Hunt, M. (1997). *How science takes stock: The story of meta-analysis*. New York: Russell Sage.

Kessler, R., Berglund, P., Demler, O., Jin, R., Merikangas, K., & Walters, E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593–602, doi:10.1001/archpsyc.62.6.593.

Kingston, T., Dooley, B., Bates, A., Lawlor, E., & Malone, K. (2007). Mindfulness based cognitive therapy for residual depressive symptoms. *Psychology and Psychotherapy: Theory, Research and Practice*, *80*, 193–203, doi:10.1348/147608306X116016.

Laidlaw, K., Davidson, K., Toner, H., Jackson, G., Clark, S., Law, et al. (2008). A randomized controlled trial of cognitive behavior therapy vs. treatment as usual in the treatment of mild to moderate late life depression. *International Journal of Geriatric Psychiatry*, *23*, 843–850, doi:10.1002/gps.1993.

Lazar, S. (2010). *Psychotherapy is worth it: A comprehensive review of its cost-effectiveness*. Arlington, VA US: American Psychiatric Publishing, Inc.

Mann, C. C. (1994). Can meta-analysis make policy? *Science*, *266*, 960–962.

Marcus, S. V., Marquis, P., & Sakai, C. (1997). Controlled study of treatment of PTSD using EMDR in an HMO setting. *Psychotherapy*, *34*(3), 307–315, doi:10.1037/h0087791.

Meuser, K. T., Rosenberg, S. D., Xie, H., Jankowski, M. K., Bolton, E. E., Lu, W., et al. (2008). A randomized controlled trial of cognitive-behavioral treatment for posttraumatic stress disorder in severe mental illness. *Journal of Consulting and Clinical Psychology*, *76*(2), 259–271, doi:10.1037/0022-006X.76.2.259.

Minami, T., Wampold, B. E., Serlin, R. C., Hamilton, E. G., Brown, G. S., & Kircher, J. C. (2008). Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Consulting and Clinical Psychology*, *76*, 116–124, doi:10.1037/0022-006X.76.1.116.

Minami, T., Wampold, B. E., Serlin, R. C., Kircher, J. C., & Brown, G. S. (2007). Benchmarks for psychotherapy efficacy in adult major depression. *Journal of Consulting and Clinical Psychology*, *75*, 232–243, doi:10.1037/0022-006X.75.2.232.

Miranda, J., Chung, J. Y., & Green, B. L. (2003). Treating depression in predominantly low-income young minority women: A randomized controlled trial. *The Journal of the American Medical Association*, *290*(1), 57–65, doi:10.1001/jama.290.1.57.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Journal of Clinical Epidemiology*, *62*, 1002–1012, doi:10.1016/j.jclinepi.2009.06.005.

Murray, C. J. L., & Lopez, A. D. (1997). Alternative projections of mortality and disability by cause 1990–2020: Global burden of disease study. *The Lancet*, *349*, 1498–1504, doi:10.1016/S0140-6736(96)07492-2.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Rosenblatt, A. (2010). Psychotherapy in the treatment of anxiety disorders. *Psychotherapy is worth it: A comprehensive review of its cost-effectiveness* (pp. 103–134). Arlington, VA US: American Psychiatric Publishing, Inc.

Spielmans, G., Gatlin, E., & McFall, J. (2010). The efficacy of evidence-based psychotherapies versus usual care for youths: Controlling confounds in a meta-reanalysis. *Psychotherapy Research*, *20*(2), 234–246, doi:10.1080/10503300903311293.

Stanley, M. A., Wilson, N. L., Novy, D. M., Rhoades, H. M., Wagener, P. D., Greisinger, A. J., et al. (2009). Cognitive behavior therapy for generalized anxiety disorder among older adults in primary care: A randomized clinical trial. *The Journal of the American Medical Association*, *310*(14), 1460–1467, doi:10.1001/jama.2009.458.

Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2007). Effectiveness of cognitive-behavioural, person-centered, and psychodynamic therapies in UK primary-care routine practice: Replication in a larger sample. *Psychological Medicine*, doi:10.1017/S0033291707001511.

Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatment: Report and recommendations. *The Clinical Psychologist*, *48*, 2–23.

van Schaik, A., van Marwijk, H., Ader, H., van Dyck, R., de Haan, M., Penninx, B., et al. (2006). Interpersonal psychotherapy for elderly patients in primary care. *American Journal of Geriatric Psychiatry*, *14*, 777–786, doi:10.1097/01.JGP.0000199341.25431.4b.

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293, doi:10.3102/10769986030003261.

Wagner, A. W., Zatzick, D. F., Ghesquiere, A., & Jurkovich, G. (2007). Behavioral activation as an early intervention for posttraumatic stress disorder and depression among physically injured trauma survivors. *Cognitive and Behavioral Practice*, *14*, 31–349, doi:10.1016/j.cbpra.2006.05.002.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bonafide psychotherapies: Empirically, "All must have prizes.". *Psychological Bulletin*, *122*, 203–215.

Ward, E., King, M., Lloyd, M., Bower, P., Sibbald, B., Farrelly, S., et al. (2000). Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: Clinical effectiveness. *British Medical Journal*, *321*, 1383–1388, doi:10.1136/bmj.321.7273.1383.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care. *The American Psychologist*, *61*(7), 671–689, doi:10.1037/0003-066X.61.7.671.

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *69*(6), 875–899, doi:10.1037//0022-006X.69.6.875.

Westen, D., Novotny, C., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, *130*(4), 631–663, doi:10.1037/0033-2909.130.4.631.