Published in final edited form as:

Clin Psychol Rev. 2018 February ; 59: 30–40. doi:10.1016/j.cpr.2017.10.009.

Long-term efficacy of psychotherapy for posttraumatic stress disorder: A meta-analysis of randomized controlled trials[☆]

Alexander C. Kline^{a,*}, Andrew A. Cooper^{a,1}, Nina K. Rytwinski^b, and Norah C. Feeny^a

^aPTSD Treatment and Research Program, Case Western Reserve University, Department of Psychological Sciences, 11220 Bellflower Road, Cleveland, OH 44106-7123, USA

^bWalsh University, School of Behavioral and Health Sciences, 2020 East Maple St., North Canton, OH, USA, 44720

Abstract

Psychotherapies are well established as efficacious acute interventions for posttraumatic stress disorder (PTSD). However, the long-term efficacy of such interventions and the maintenance of gains following termination is less understood. This meta-analysis evaluated enduring effects of psychotherapy for PTSD in randomized controlled trials (RCTs) with long-term follow-ups (LTFUs) of at least six months duration. Analyses included 32 PTSD trials involving 72 treatment conditions ($N = 2935$). Effect sizes were significantly larger for active psychotherapy conditions relative to control conditions for the period from pretreatment to LTFU, but not posttreatment to LTFU. All active interventions demonstrated long-term efficacy. Pretreatment to LTFU effect sizes did not significantly differ among treatment types. Exposure-based treatments demonstrated stronger effects in the post-treatment to LTFU period ($d = 0.27$) compared to other interventions ($p = 0.005$). Among active conditions, LTFU effect sizes were not significantly linked to trauma type, population type, or intended duration of treatment, but were strongly tied to acute dropout as well as whether studies included all randomized patients in follow-up analyses. Findings provide encouraging implications regarding the long-term efficacy of interventions and the durability of symptom reduction, but must be interpreted in parallel with methodological considerations and study characteristics of RCTs.

[☆]Preparation of this manuscript was supported by a grant to Dr. Feeny from the National Institute of Mental Health (R01 MH066348). The funding source had no involvement in any aspect of this manuscript.

^{*}Corresponding author at: Case Western Reserve University, Department of Psychological Sciences, 11220 Bellflower Road, Cleveland, OH 44106-7123, USA. ack63@case.edu (A.C. Kline).

¹Andrew Cooper is now at the Department of Psychology, University of Toronto-Scarborough.

³Given the wide range of dropout rates reported across conditions in our sample, we examined estimates by treatment type, with significant differences in dropout rate emerging, $F(4,50) = 4.25$, $p = 0.005$. Post-hoc comparisons using the Bonferroni correction indicated that cognitive therapy (CT; $k = 5$; mean dropout rate = 4.53%, $SD = 3.99$) had significantly lower dropout rates compared to exposure therapy conditions ($k = 30$, mean dropout rate = 24.08%, $SD = 12.34$, $p = 0.004$), CPT ($k = 6$, mean dropout rate = 28.73%, $SD = 6.24$, $p = 0.005$), and CBT-M ($k = 11$, mean dropout rate = 22.42%, $SD = 10.33$, $p = 0.03$), but not EMDR ($k = 3$, mean dropout rate = 19.08%, $SD = 12.34$, $p = 0.69$). Notably, all CT studies in this sample originate from the United Kingdom, potentially suggesting a sociocultural influence on attrition rates stemming from the nationalized healthcare system present in that country.

Conflict of interest

All authors declare that they have no conflicts of interest.

Contributors

ACK and AAC developed the concept for the study, reviewed the literature, and conducted analyses. NKR and NCF consulted on methodology and the framework of the meta-analysis. All four authors contributed to writing the manuscript and approve of its final version.

Keywords

PTSD; Psychotherapy; Treatment outcome; Long-term follow-up; Meta-analysis

1. Introduction

PTSD is a prevalent, debilitating, and typically chronic disorder associated with significant distress and functional impairment in a number of domains, as well as considerable public health and economic ramifications (e.g., Kessler, Chiu, Demler, & Walters, 2005; Tanielian & Jaycox, 2008). Fortunately, across a range of populations, settings, and trauma types, effective psychotherapy options exist for PTSD. Numerous systematic and meta-analytic reviews (e.g., Bradley, Greene, Russ, Dutra, & Westen, 2005; Cusack et al., 2016; Haagen, Smid, Knipscheer, & Kleber, 2015; Lee et al., 2016; Powers, Halpern, Ferenschak, Gillihan, & Foa, 2010; Watts et al., 2013) demonstrate the efficacy of psychotherapeutic interventions for both civilian and military populations. In particular, a number of trauma-focused cognitive behavioral treatments, including cognitive processing therapy (CPT; Resick & Schnicke, 1993), cognitive therapy (CT; Ehlers, Clark, Hackmann, McManus, & Fennell, 2005), and exposure-based treatments such as prolonged exposure (PE; Foa, Hembree, & Dancu, 2002), have consistently exhibited strong efficacy in the acute phase of treatment and are recommended as first-line treatment options (e.g., Forbes et al., 2010; Institute of Medicine, 2007). Differences among these treatments generally relate to the targeting of key mechanisms of change and the implementation of the cognitive-behavioral techniques specified within the protocol (e.g., stress management and relaxation, psychoeducation, cognitive restructuring, and in-session exposures). In brief, exposure-based therapies (e.g., PE) are defined by the inclusion of some form of repeated exposure to trauma reminders such as the trauma memory aimed at promoting extinction of fear, reduction of avoidance, and changing trauma-related thinking. Cognitive-based interventions, such as CPT and CT, tend to focus more explicitly on identification and modification of dysfunctional beliefs and cognitive patterns following trauma that maintain conditioned fear and patterns of avoidance.

While the efficacy of trauma-focused psychotherapy options has been established in the acute phase, our understanding of the long-term impact of such interventions is more limited. Prior PTSD treatment meta-analyses have primarily focused on acute, immediate outcomes following treatment, with effect sizes calculated using scores from a posttreatment assessment typically conducted shortly after the conclusion of treatment. Additionally, while there is excellent research on spontaneous remission among individuals with PTSD independent of treatment (Morina, Wicherts, Lobbrecht, & Priebe, 2014), there is substantially less literature regarding the long-term impact of interventions for PTSD. Despite robust and well-established support for the efficacy of various psychotherapies in the acute phase of PTSD treatment, no meta-analytic reviews have evaluated the long-term efficacy of these interventions.

Furthermore, though the majority of patients respond to evidence-based treatments for PTSD, some do not. In line with an emerging emphasis on predictors of acute response in

clinical research on PTSD (e.g., Schneider, Arch, & Wolitzky-Taylor, 2015), further study of potential predictors impacting long-term treatment response and maintenance of gains following treatment is necessary. Prior meta-analyses have examined a variety of factors in attempts to identify predictors of acute treatment response, which have included baseline sample characteristics (e.g., population, gender, trauma type Bradley et al., 2005; Watts et al., 2013) and variables tied to treatment components and processes (e.g., intended sessions per protocol, whether interventions are trauma-focused Haagen et al., 2015). However, reliable predictors of response remain elusive (e.g., Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012; Taylor, Abramowitz, & McKay, 2012). Thus, despite the evidence base of several treatments for PTSD, there remains a need to further explore patient-related and treatment-related variables that may predict enduring treatment response.

The goal of the current study was to characterize the long-term outcomes of psychotherapies for PTSD and identify predictors of long-term treatment response. Examining the long-term impact of interventions should provide a more comprehensive understanding of their efficacy beyond short-term, acute-phase symptom reduction. Additionally, understanding long-term outcomes for specific types of interventions may better inform treatment directives and clinical decision making. Indeed, the extent to which brief interventions can effectively and efficiently reduce PTSD severity and maintain gains holds significant public health and economic implications. The current study included a methodologically rigorous, systematic review of randomized controlled psychotherapy trials (RCTs) for PTSD, examining treatment outcomes at minimum six months posttreatment. We also examined potential predictors of long-term outcomes, including sample characteristics and treatment-related factors, with emphasis on clinically and empirically salient factors in psychotherapy research. Further, we also closely examined methodological factors and study characteristics that may impact estimates of effect size (Bradley et al., 2005; Watts et al., 2013), such as rates of attrition in study conditions and analytic methods of studies (e.g., completer vs. intent to treat). As described below, we also took steps to optimize the precision of effect size estimates, excluding studies on the basis of sample size (e.g., Hedges & Pigott, 2001) and risk of bias (e.g., Cuijpers, Straten, Bohlmeijer, Hollon, & Andersson, 2010).

2. Method

2.1. Search strategy

The search process was conducted in two phases. First, the PsycINFO database was searched for articles published from 1980 through 2015, using the search terms “PTSD” OR “post traumatic stress disorder” OR “posttraumatic stress disorder” OR “post-traumatic stress disorder” AND “psychotherapy” or “therapy” AND “treatment” OR “trial” OR “randomized”. Limiters applied in the search were publication year (1980–2015), language (English only), and age group (adulthood, defined as 18 years and older). Following the initial search, reference lists of prior comprehensive meta-analyses of RCTs for PTSD (e.g., Bradley et al., 2005; Cusack et al., 2016; Ehring et al., 2014; Imel, Laska, Jakupcak, & Simpson, 2013; Watts et al., 2013) were closely examined. Any discrepancies between search results and reference lists were recorded and additional studies potentially suitable for inclusion were also closely reviewed.

2.2. Inclusion criteria

The current manuscript focuses exclusively on in-person psychotherapies for PTSD, with pharmacological treatments for PTSD excluded for several salient reasons. While follow-up assessment and long-term outcomes are clearly defined in psychotherapy conditions given the explicit starting and ending times of treatment, “dose” is more difficult to temporally assess in pharmacotherapy interventions, where patients often remain on medications even after the initial phase of assessment is completed. Pill placebo conditions were also therefore excluded as potential control comparisons to active psychotherapies given that pill placebo is a more relevant, representative control to pharmacotherapy interventions. While pill placebo does reflect an inert control condition, this is not a typical psychotherapy control, such as non-directive counseling, relaxation, treatment as usual, or even wait-list control. Telehealth conditions were also excluded in the current meta-analysis given the contrast in therapy modality compared to conventional, face-to-face psychotherapy. Difficulties in interpreting and comparing dropout rates in telehealth conditions relative to in-person psychotherapy are also relevant, with recent meta-analytic evidence suggesting significantly higher rates in teletherapy conditions (Fernandez, Salem, Swift, & Ramtahal, 2015), including studies of PE (Franklin, Cuccurullo, Walton, Arseneau, & Petersen, 2017). For these reasons, we chose to exclusively focus on in-person psychotherapies for PTSD.

Studies were eligible for inclusion if they met the following criteria: (a) the study consisted of adult patients; (b) patients were formally assessed and diagnosed with full PTSD (i.e., not subsyndromal or subthreshold PTSD); (c) patients received in-person psychotherapy with a duration of at least three sessions primarily targeting PTSD severity (i.e., not targeting a specific sub-symptom or comorbid condition), (d) data from a reliable, valid assessment measure (i.e., measures supported by published, peer-reviewed work detailing the construction and psychometrics of the assessment) were available for PTSD severity at pretreatment, posttreatment, and at follow-up assessments at least six months posttreatment, (e) and the study was reported in English. Our follow-up time cutoff of six months was based on criteria established by a recent meta-analysis of long-term outcomes for depression treatments (Karyotaki et al., 2016).

Final inclusion criteria pertained to (f) total sample size of the study, where the randomized N was required to consist of at least 30 patients, and (g) risk of bias. In an attempt to emphasize high-quality RCTs and reduce variance that may be explained by poor design, we prioritized RCTs with adequate sample size and minimal risk of bias. This approach is line with strategies recommended to enhance the utility of meta-analyses (Berlin & Golub, 2014; Fleischhacker, 2017), given the links between sample size, bias, and precision (e.g., Cuijpers et al., 2010; Dechartres, Altma, Trinquart, Boutron, & Ravaud, 2014; Flint, Cuijpers, Horder, Koole, & Munafò, 2015; Hedges & Pigott, 2001; Hedges & Pigott, 2004; Turner, Bird, & Higgins, 2013).

Regarding study sample size, prior meta-analytic reviews of RCTs (e.g., Bradley et al., 2005) and other PTSD meta-analyses, (e.g., Morina et al., 2014; Rytwinski, Scur, Feeny, & Youngstrom, 2013) have set inclusion criteria based on a minimum N at baseline or pre-treatment. Establishment of a minimum N in meta-analysis holds implications for statistical power, a critical factor in both the psychotherapy literature (e.g., Cuijpers et al., 2010; Flint

et al., 2015; Kazdin & Bass, 1989), and meta-analytic research (e.g., Hedges & Pigott, 2001; Hedges & Pigott, 2004; Kraemer, Gardner, Brooks, & Yesavage, 1998; Shadish Jr. & Sweeney, 1991; Turner et al., 2013). Underpowered studies are at greater risk of imprecision (i.e., Type I or II errors) and are also more likely to be published only if they contain statistically significant results, thus contributing to the well-documented “file drawer problem”, or publication bias (Kraemer et al., 1998). Power in meta-analyses is demonstrably enhanced by the exclusion of underpowered studies (Hedges & Pigott, 2001). Additionally, it is important to note that publication bias has inflated estimates of psychotherapy effects, as demonstrated in a recent review of depression interventions (Driessen, Hollon, Bockting, Cuijpers, & Turner, 2015). Further rationale for setting a minimum sample size exclusion criterion a priori pertains to the current focus on long-term outcomes, given that the follow-up period bears a greater risk of attrition over time and therefore may potentially further exacerbate issues posed by small sample sizes.

In regards to risk of bias, ratings of bias and study quality have been found to impact effect size estimates of interventions in meta-analytic reviews, with smaller effect sizes often found in studies of higher quality (Cuijpers et al., 2010). In an effort to calculate optimal estimates of treatment efficacy, studies at high risk of bias were excluded from the current meta-analysis. To assess risk of bias, we mirrored the procedure of Cusack et al. (2016) in their review of PTSD treatments. Risk of bias was evaluated via 12 questions derived from the Agency for Healthcare Quality and Research Methods Guide for Comparative Effectiveness Reviews (Viswanathan et al., 2012) assessing various forms of bias, as reflected by the following concepts: adequate randomization and allocation concealment, uniformity of groups at baseline, masking of independent evaluators, rates of overall and differential attrition, type of analyses used, handling of missing data, use of valid and reliable outcome assessments, and ratings of treatment fidelity. Risk of bias ratings were adopted from those reported by Cusack et al. (2016) when available, and otherwise provided independently by two co-authors (ACK, NKR). Overall agreement between raters was very high. There were only three discrepancies between raters, which were resolved via discussion and consensus among other authors. More comprehensive information regarding specific criteria for ratings can be found in the aforementioned review (Cusack et al., 2016).

2.3. Coding procedure and data extraction

Data from articles were extracted and coded by the primary author (ACK) and a trained, independent second rater. Interrater reliability between coders was excellent ($\kappa = 0.97$). For any necessary study information missing, we contacted study authors to attempt to obtain it. Two variables related to treatment samples were coded, including population type (veterans/active duty versus civilian) and trauma type, based on whether the primary target trauma involved sustained childhood physical or sexual abuse (coded as: all of the sample experienced sustained childhood abuse; none of the sample experienced sustained childhood abuse; or mixed for studies where there were a variety of trauma types and part of the sample experienced sustained childhood abuse).

Conditions were broadly classified as either active treatments (e.g., PE, CPT) or control conditions used in studies as comparison groups (e.g., treatment as usual, waitlist control, or

non-directive control conditions such as supportive counseling or relaxation). Drawing closely from a recent prior review (Cusack et al., 2016), all conditions were also categorized into one of eight groups (five active, three inactive), including (a) cognitive behavioral therapies-mixed (CBT-M) for interventions predominantly including one or a blend of the salient components of CBT, such as cognitive restructuring, psychoeducation, relaxation training, and skills training,² (b) CPT, (c) CT, (d) eye movement desensitization and reprocessing (EMDR; Shapiro, 1995), (e) exposure therapies (i.e., interventions emphasizing exposure to the trauma memory as the principal active treatment component, such as PE and imaginal exposure), (f) non-directive control conditions, where patients were assigned to a non-active, non-trauma focused condition meant to serve as a control comparison (e.g., supportive counseling, relaxation), (g) treatment as usual (TAU), and (h) waitlist control group, where no intervention of any kind was provided to the control group. For studies in which patients in waitlist control conditions ultimately received active treatment as part of study protocol, data from long-term follow-up assessments were excluded from analyses.

Among active treatment conditions, we coded several treatment-related variables. Regarding treatment duration, we coded the intended number of sessions and weeks of each intervention as prescribed by study protocol. Acute treatment dropout rates were also recorded among all active conditions, based on the number of patients who were randomized to a treatment but failed to complete it, excepting those removed administratively (i.e., protocol violation, medical emergency, or death).

Finally, to determine the impact of study design on estimated effect sizes, we coded several methodological variables. Acute analysis type was coded based on whether study authors used intent to treat (ITT) versus completer analyses for their primary reported outcomes during the acute phase of treatment. We also coded a binary variable regarding whether authors followed and assessed all randomized patients through the follow-up phase, or focused only on a subset of patients (e.g., treatment completers or responders). We also coded a binary variable regarding the timing of the final follow-up assessment, at six months posttreatment versus greater than six months posttreatment, to evaluate the potential influence of assessment timing on effect size estimates.

2.4. Analytic strategy

We used continuous outcome measures of PTSD symptom severity as our primary measures to calculate effect sizes. When studies included both self-report and interviewer-based outcome measures, we used the latter in our analyses. When studies included both ITT and completer analyses, we used ITT results in our analyses. Effect sizes were calculated using Cohen's *d* for each condition of each study. Regarding our primary outcomes, we calculated two different effect size estimates from two different time periods: (1) pretreatment through follow-up and (2) posttreatment through follow-up. We focused exclusively on pre-post effect sizes given that variances between pretreatment and posttreatment measures were not

²Table 1 includes information about all conditions that were classified as CBT-M. As articulated by Cusack et al. (2016) in their review, CBT-M conditions typically constituted either specific intervention components typical of CBT delivered separately (e.g., in vivo exposure only) or in sequential augmentation (e.g., skills training & conventional CBT). Compared to other established protocols, these conditions were more variable in terms of intended sessions and use of homework assignments.

expected to be equivalent, and the correlation between pre- and post- measures were not available among included RCTs. Following the recommendations of Morris and DeShon (2002) regarding pre-post effect sizes, we used a modified version of Cohen's d , replacing pooled variance with the pretreatment standard deviation, which is unlikely to be affected by treatment and vary substantially across studies. More specifically, we calculated effect sizes by subtracting time point 2 (follow-up mean) from time point 1 (pretreatment mean or posttreatment mean), and then divided this total by the standard deviation at time point 1. To account for the differences in sample size across conditions and the variable methods of handling attrition, effect sizes were weighted based on pretreatment n of each condition to assign more weight to larger studies, which provides more accurate estimates of effect sizes relative to smaller studies (Borenstein, Hedges, Higgins, & Rothstein, 2010). When data were available for multiple follow-up assessments, data from the last available follow-up assessment were used in analyses and reported. The longest period of reported follow-up is 5 years posttreatment (Resick, Williams, Suvak, Monson, & Gradus, 2012), assessing patients that had been treated in an earlier trial examining CPT and PE (Resick, Nishith, Weaver, Astin, & Feuer, 2002). Because the duration of this follow-up was substantially longer than any other condition included in the current study by a margin of several years, these outcomes at 5-year follow-up were not included in analyses. Follow-up data initially reported in the 2002 paper were included in analyses, however. This is in line with a prior meta-analysis (e.g., Cuijpers, Straten, Warmerdam, & Andersson, 2009), which also removed outliers in terms of timing of follow-up assessments.

All analyses were conducted with SPSS 23 software using macros developed by Wilson (2005). These macros calculated weighted mean effect sizes and also performed mixed effects subgroup analyses mirroring ANOVA (e.g., examining differences in effect sizes based on treatment type) and mixed-model meta-regression analyses for continuous predictor variables (e.g., examining the link between intended duration of treatment and effect sizes). In all cases, models utilized restricted maximum likelihood estimation. Study heterogeneity was assessed via the Q statistic and a corresponding significance test, as well as the I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003). For I^2 , 0% indicates no observed heterogeneity, 25% indicates low, 50% indicates moderate, and 75% indicates high.

3. Results

3.1. Study selection

Steps regarding the study selection and evaluation process for inclusion are presented in Fig. 1. The initial search of the PsycINFO database yielded 3097 citations, which were screened on the basis of title and abstract. Following this initial screening, 223 studies with the potential to meet inclusion criteria were retained. Each abstract was closely reviewed by authors ACK and AAC, followed by full text if necessary to determine eligibility. A total of 192 articles were excluded, with the most common reason for exclusion being diagnostic criteria of the sample (e.g., inclusion of subthreshold PTSD within the RCT sample; trauma-exposure but no formal PTSD diagnosis). As noted, the final step of the selection process included a careful review of reference lists of past meta-analyses of RCTs for PTSD, which

yielded one additional study for inclusion. Ultimately, 32 studies met full inclusion criteria and were included in the current study.

3.2. Study characteristics

Table 1 presents the characteristics of the studies and their respective conditions included in the meta-analysis, comprising 32 studies, 86 total conditions, and 3399 patients. Studies were published between 1998 and 2015. Study sample sizes ranged from 30 to 284 with a median of 92 ($M = 106.22$, $SD = 56.68$). Among included studies, 14 conditions were removed from analyses for the following reasons: nine control conditions did not include follow-up assessments of at least six months posttreatment, two active treatment conditions did not have adequate pretreatment, posttreatment, and follow-up data, one active condition was a teletherapy treatment group, one active condition was a pharmacotherapy group, and one control condition was a pill placebo condition. Taken altogether, 72 (55 active, 17 inactive) conditions ($N = 2935$) met full inclusion criteria and were included in analyses. The mean effect size (d) across all conditions was 1.88 (95% CI [1.68, 2.08]) and 0.17 (95% CI [0.10, 0.23]) for pretreatment to follow-up and posttreatment to follow-up time periods, respectively. The sample size of active and inactive treatment conditions ranged from 15 to 141 ($M = 40.23$, $SD = 22.81$) and 12 to 143 ($M = 42.47$, $SD = 32.76$), respectively. For active treatments, average intended acute treatment duration was 9.73 weeks ($SD = 4.78$) and 11.16 sessions ($SD = 3.44$). Timing of final follow-up assessments in active conditions were at six months ($k = 48$, 66.7%), nine months ($k = 3$, 4.2%), 12 months, ($k = 19$, 26.4%), and 20 months ($k = 2$, 2.8%); thus, approximately one-third of studies had a follow-up assessment greater than six months after treatment concluded. Dropout rates among active treatment conditions ranged from 0 to 47%, with average dropout being 22.2%.² Regarding analysis type during the follow-up phase, 21 studies (65.6%) included all randomized patients in follow-up analyses, while 11 (34.3%) studies only included a subset of randomized patients in follow-up analyses, most often treatment completers.

3.3. Comparisons among treatment types

Analyses indicated substantial heterogeneity within included conditions in both pretreatment to follow-up ($Q_{71} = 2169.55$, $p < 0.0001$, $I^2 = 96.73$) and posttreatment to follow-up ($Q_{71} = 233.32$, $p < 0.0001$, $I^2 = 69.57$) time periods. Random effects models were thus used in all cases (Lipsey & Wilson 2001).

Results for comparisons among treatment conditions are presented in Table 2. Analyses indicated significant differences in weighted effect sizes between active treatments ($d = 2.14$) and control conditions ($d = 1.04$) in the pretreatment to follow-up period ($p < 0.001$), but not the posttreatment to follow-up period ($p = 0.90$). Each subgroup of treatments (exposure, CT, CBT-M, CPT, EMDR) were compared to all other active treatment conditions collapsed together, and analyses indicated no significant differences in effect sizes among active interventions in the pretreatment to follow-up time period ($ps = 0.07$ – 0.50). For the posttreatment to follow-up time period, only exposure therapies demonstrated significantly higher effect sizes ($d = 0.27$) compared to other interventions ($d = 0.05$, $p = 0.005$). Additionally, during the same time period, CBT-M showed significantly lower effect sizes ($d = -0.01$) compared to other interventions ($d = 0.21$, $p = 0.02$).

3.4. Sample-related and treatment-related predictors of effect size

Predictors and subgroup analyses are also presented in Table 2. Among active treatment conditions, population type (i.e., military vs. civilian) was not significantly tied to pretreatment to follow-up effect sizes ($p = 0.82$), but effect sizes were larger for military compared to civilian samples at a trend level in the posttreatment to follow-up time period ($p = 0.06$). Among active treatments, the extent to which childhood physical or sexual abuse was specific as the index trauma was not predictive of pretreatment to follow-up effect size ($p = 0.36$). However, differences emerged at a trend level in the posttreatment to follow-up time period ($p = 0.07$), such that samples entirely comprised of childhood abuse as the index trauma ($k = 5$) showed greater average gains in the post-treatment window ($d = 0.46$) compared to studies focusing exclusively on traumas in adulthood ($k = 27$, $d = 0.12$) or studies with mixed trauma focus ($k = 21$, $d = 0.14$) samples.

Treatment-related variables reflecting intervention duration were also not significantly related to effect size estimates, which included number of sessions (pretreatment to follow-up: $\beta = 0.04$, $p = 0.75$; posttreatment to follow-up: $\beta = 0.01$, $p = 0.93$) and weeks (pretreatment to follow-up: $\beta = 0.07$, $p = 0.63$; posttreatment to follow-up: $\beta = 0.10$, $p = 0.47$) of acute treatment. However, posttreatment to follow-up effect sizes were significantly higher for conditions with longer follow-up periods (i.e., greater than six months duration, as opposed to follow-up assessments at 6 months posttreatment; $p = 0.007$).⁴

3.5. Study design and estimates of effect size

Regarding study design variables and their link to effect sizes, an important pattern of results emerged related to attrition and type of analyses used by study authors. Among active treatment conditions, higher rates of acute treatment dropout predicted attenuated pretreatment to follow-up effect sizes at a trend level ($\beta = -0.25$, $p = 0.06$). However, greater dropout was also associated with significantly larger posttreatment to follow-up effect sizes ($\beta = 0.33$, $p = 0.01$). In the acute phase of treatment, no significant differences emerged in terms of analysis type, with ITT and completer analyses yielding similar effect size for both pretreatment to follow-up ($p = 0.60$) and posttreatment to follow-up ($p = 0.27$). Notably, analysis type demonstrated significant effects in the pretreatment to follow-up period. More specifically, treatment conditions that included only a subset of randomized patients ($d = 2.51$), who were most often treatment completers, demonstrated significantly larger effects compared to studies that included all randomized patients in follow-up analyses ($d = 1.95$), $p = 0.03$. In the posttreatment to follow-up period, effects were in a similar direction but not statistically significant ($p = 0.33$).

To further investigate these findings related to attrition, we ran sensitivity analyses examining type of follow-up analyses used by included studies - a binary variable measuring whether studies followed and assessed all randomized individuals in the follow-up phase vs.

⁴To determine if this finding was an artifact of whether all randomized patients were assessed in the follow-up period, we ran a chi-square with two binary variables: follow-up analyses type and follow-up assessment time point (six months or greater than six months). Timing of follow-up assessments was not significantly linked to follow-up analyses type among active conditions, $\chi^2(1, N = 55) = 2.55$, $p = 0.11$, suggesting that larger effect sizes among studies with assessment time points greater than six months were not a product of the type of follow-up analyses used.

assessed only a subset of patients - as a moderator of the relationship between dropout and effect sizes. Our interaction term was significant for the pretreatment to follow-up model ($p = 0.004$) and at a trend level for the posttreatment to follow-up model ($p = 0.10$). Unpacking these moderation effects, in the pretreatment to follow-up time period, higher rates of acute dropout predicted lower effect sizes among conditions following all randomized individuals ($\beta = -0.45$, $p = 0.002$), but not among conditions only assessing a subset during the follow-up period ($\beta = 0.34$, $p = 0.14$). In the post-treatment to follow-up time period, higher rates of acute dropout predicted higher effect sizes among conditions only assessing a subset of randomized patients during the follow-up period ($\beta = 0.70$, $p < 0.001$), but were not associated with effect size among conditions following all randomized individuals ($\beta = 0.21$, $p = 0.20$). Stated differently, attrition effects were particularly pronounced based on whether or not study authors followed and assessed all individuals initially randomized to treatments.

3.6. Fail-safe N calculation

Potential publication bias was assessed by calculating the fail-safe N (Rosenthal, 1979) adapted by Orwin (1983). This estimates the number of nonsignificant or unpublished studies needed to reduce the aggregate effects observed across studies included in the meta-analysis to a specified criterion. Mirroring contemporary work by Sloan, Feinstein, Gallagher, Beck, and Keane (2013), we set this threshold at 0.01, reflecting an effect that is nearly null. For within-group effects on PTSD, the fail-safe N was 5984 for our main within-group analyses in the pretreatment to follow-up time period, strongly suggesting the absence of publication bias.

4. Discussion

In this systematic review examining the enduring effects of acute psychotherapy for adults with PTSD, the overall pattern of findings is supportive of the long-term impact of active interventions. Notably, patterns of effect sizes generally reinforce the efficacy of evidence-based psychotherapies for PTSD, mirroring outcomes of other reviews focused on the acute phase of treatment (e.g., Bradley et al., 2005; Cusack et al., 2016; Watts et al., 2013) and providing further evidence for a range of efficacious PTSD treatments. Similar to prior reviews, there was not strong evidence for unequivocal superiority of any particular intervention, with comparable effect sizes from baseline through follow-up. Overall, findings of the current study reinforce well-established support for evidence-based, trauma-focused treatments for PTSD and extend them to longer durations following treatment. The current study adds to the strong imperative to use these treatments for PTSD, as these interventions clearly demonstrate enduring effects and maintenance of gains following termination.

While our findings suggest that all psychotherapies led to lasting improvements following treatment, differences between treatment types emerged in analyses focused specifically on the period from posttreatment to follow-up. In this time period, exposure therapies demonstrated significantly larger effect sizes compared to other treatments. Despite a small difference in absolute terms ($d = 0.27$), this finding is particularly noteworthy in the context of generally modest posttreatment gains among treatments and somewhat restricted variance, as reflected by the I^2 estimate. However, we also note that exposure-based treatments are

particularly well-represented in our sample compared to other treatment types, which may affect comparative analyses. Unexpectedly, CBT-M conditions demonstrated significantly smaller effect sizes from posttreatment to follow-up compared to other interventions. One possible explanation for this finding may be due to the substantial heterogeneity of treatment protocols within this category, with some studies involving multiple active therapeutic components (i.e., in vivo exposure, cognitive restructuring, skills training, relaxation, etc.) and others focusing on a single technique. Additionally, it may be that more concentrated, stronger doses of traditional exposure therapy coupled with exposure-based homework assignments may be more effective in fear extinction and the gains continuing following the termination of treatment. Indeed, completion of exposures and stronger homework adherence produces better outcomes in treatment for anxiety disorders (Glenn et al., 2013), possibly signifying a link between exposure practice and consolidation of treatment gains (Cooper et al., 2017). Exposure-focused therapies, with the principal ingredient often being imaginal exposure, demonstrated the most robust continued improvement following treatment, yet it is meaningful that all of the broad range of trauma-focused treatments included in the meta-analysis produced lasting gains.

We found little evidence that study population or treatment structure was linked to overall or enduring treatment effects. Intended duration of treatment was not linked to effect sizes, consistent with prior reviews of acute treatment of PTSD (Haagen et al., 2015). Regarding childhood abuse, the extent to which childhood physical and/or sexual abuse was the index trauma in samples did not robustly impact overall outcomes, although there was a trend toward more robust gains in the post-treatment to follow-up period in the small number of studies focused primarily on childhood traumas ($k = 5$). Population type was not robustly linked to overall or enduring effect size, in contrast to meta-analyses of acute treatment that have found somewhat lower mean effect sizes for samples with a greater percentage of veterans compared to civilians (Bradley et al., 2005; Watts et al., 2013). Debate remains whether civilians achieve stronger PTSD treatment outcomes compared to veterans, as the evidence is inconsistent thus far (Steenkamp, Litz, Hoge, & Marmar, 2015). We note that the current meta-analysis includes a subset of treatment studies for military-related PTSD relative to two recent, broader reviews of PTSD treatment outcomes for veterans (Haagen et al., 2015; Steenkamp et al., 2015). In the comprehensive review completed by Haagen et al. (2015), roughly 67% of included studies were non-randomized, while our focus was on randomized controlled trials. The emphasis of the current meta-analysis on long-term outcomes also excluded a large number of studies in both recent military-related PTSD reviews, where many included studies did not include long-term outcome assessments at least 6 months following treatment. It is therefore recommended that findings regarding long-term outcomes among veterans be interpreted cautiously, given the relatively small number of studies exclusively with veterans in our meta-analysis.

The duration of follow-up assessments among studies ranged from six to 20 months. While pretreatment to follow-up effect sizes were virtually equivalent when comparing assessments at six months ($d = 2.14$) and after six months ($d = 2.13$), effect sizes from post-treatment to follow-up assessing outcomes longer than six months were larger ($d = 0.31$) compared to assessments at the six month time point ($d = 0.09$). Notably, our findings were not an artifact of follow-up analysis type. Though more studies at the six month assessment time point

assessed all randomized individuals - which would presumably be associated with smaller effect sizes - this difference was not statistically significant. Interestingly, our pattern of findings runs counter to the one observed in a recent meta-analysis of long-term depression outcomes, where effect sizes decreased with longer follow-up periods (Karyotaki et al., 2016). This is also particularly noteworthy given the considerable diagnostic overlap between PTSD and depression as well as the high rates of comorbidity between the two disorders (Rytwinski et al., 2013). It is possible that the longer time between posttreatment and follow-up assessments may have provided greater opportunity for skills to be practiced and reinforced and for treatment gains to crystallize. Additionally, these findings may reflect greater opportunity for symptom severity to spontaneously diminish. Given the discrepancy with the recent depression meta-analysis on long-term outcomes, a better understanding of residual symptoms in the months long after treatment termination is needed. However, these findings provide tentative reasons for optimism regarding the durability of treatment gains for evidence-based PTSD treatments.

Our findings related to long-term treatment efficacy should be considered and interpreted in the context of studies' methodological implications. Indeed, our results reflect the significance of analytic approaches, sample size, attrition, and the constellation of factors that comprise study quality or risk of bias. The type of analyses used by study authors was robustly linked to mean effect sizes, with significantly smaller effect sizes from pretreatment through follow-up found in analytic models that included all randomized patients in follow-up analyses ($d = 1.95$) compared to models that only included a subset of patients, often treatment completers ($d = 2.51$). Our findings strongly indicate that the type of analyses employed in RCTs has a robust impact on effect size estimates, and ultimately the implications drawn from treatment research. Indeed, findings reinforce directives from the Institute of Medicine (2007) and other contemporary guidelines in advocating the use of intent to treat analyses in RCTs to provide more precise estimates of treatment effects (for a further discussion, see Ten Have et al., 2008). The most precise estimates are likely to be drawn from analytic models that are intent to treat for not only the acute phase of treatment, but also the follow-up phase as well. Models that do not include all randomized patients are likely to provide overly optimistic, inflated estimates of long-term treatment effects.

In line with this, an informative pattern of findings emerged regarding dropout and analyses type. In the current study, dropout was robustly associated with effect size estimates in two meaningful ways: (a) effect sizes from baseline onward were significantly *lower* when analytic models included all randomized patients and (b) effect sizes from posttreatment onward were significantly *higher* when analytic models did not include all randomized patients. These findings mirror other recent reviews. For example, Watts et al. (2013) also found higher rates of missing data among psychotherapy trials to be associated with higher effect sizes. Additionally, Bradley et al. (2005) observed a strong link between treatment completion and effect sizes, such that fewer patients finishing treatment predicted higher effect sizes in trials. Examining long-term outcomes is particularly conducive to the pattern of effects in our results, as there are multiple possible periods of attrition that can ultimately impact outcomes. More specifically, completers – theoretically more likely to be treatment responders by virtue of receiving a full protocol of treatment – are initially assessed at post-treatment, but then also again throughout the follow-up phase. During follow-up, attrition

effects also likely occur, where responders and patients who have maintained treatment gains are more likely to complete follow-up assessments. Similar to attrition, the sample size of RCTs is another salient factor in interpreting findings of RCTs and reviews. As attrition increases, power and precision decreases, which are further magnified when trials have a small sample size even at the start of treatment (Kazdin & Bass, 1989). The current study used a minimum N criterion to account for this influence, but even with this cutoff, the sample size of individual conditions became smaller due to attrition during treatment and after termination in the follow-up phase. This reduction of sample size likely introduces greater error in effect sizes estimate and ultimately provides overly optimistic effect sizes among studies that conducted completer analyses. Intent to treat analyses, which better handle attrition, are likely to provide a safeguard against these types of effects (Ten Have et al., 2008).

Assessing the risk of bias and quality of RCTs remains an important aspect of conducting meta-analytic reviews, but has frequently been defined in different ways. Cuijpers et al. (2010) noted that nearly 30 scales had been used to assess quality in RCTs. There currently is no standardized metric for determining study quality and risk of bias, but their impact on meta-analytic findings is significant, as studies of higher quality generally show smaller effect sizes (Cuijpers et al., 2010). Indeed, while the above considerations regarding analysis type and study quality do not necessarily limit the findings of meta-analytic reviews, they must be thoroughly considered when interpreting results. Studies that are most likely to yield the most precise estimates of treatment effects are ones of high quality with adequate sample sizes that also used statistical models robust to missing data. These methodological factors – risk of bias, sample size, analysis type – are often related to one another, but there currently no well-established directives or guidelines for handling these issues uniformly in meta-analyses or other comprehensive reviews. It is critical that future reviews continue to take into consideration the aforementioned methodological factors, which are likely to ultimately influence findings and recommendations.

4.1. Strengths and limitations

The current meta-analysis closely adhered to current guidelines for conducting systematic reviews and meta-analyses (e.g., PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2009). To begin, a thorough search was conducted using concrete inclusion criteria based on theoretical and empirical rationale. Numerous steps were taken to minimize potential bias and improve precision of effect size estimates in our review. We focused on gold-standard studies, excluding those judged to have a high risk of bias, as determined via two independent raters. There are many ways to assess study quality. To provide uniformity and boost generalizability, we followed the same rating criteria used by Cusack et al. (2016) to review and assess the most salient study characteristics related to bias in RCTs. Additionally, a cutoff for sample size was established to ensure inclusion of adequately powered RCTs and more precise estimates of effects in the meta-analysis. Primary outcomes for this study were based on validated, psychometrically-sound assessments. Finally, the study included a number of well-researched, evidence-based treatments (e.g., PE, CPT, trauma-focused CBT) and a range of heterogeneous samples that captured a broad range of clinical complexities,

demographic characteristics, populations, trauma history, comorbidities, and index trauma types.

Several limitations should be considered when interpreting findings. First, the timing of follow-up assessments used in analyses were somewhat variable, ranging from six to 20 months. This variability, however, is similar to prior meta-analytic reviews of long-term followup outcomes (e.g., Karyotaki et al., 2016). Additionally, this extended range of follow-up assessment time points provides information regarding the efficacy of treatments and durability of gains nearly two years following the treatment termination. Additionally, very little data were provided among studies regarding whether patients were eligible to receive additional or adjunctive treatments during the follow-up phase, whether as part of the study design (e.g., “booster sessions”) or through another external provider. Few studies offered any detail about their follow-up protocols vis-à-vis these adjunctive or additional treatment sessions, and actual data about attendance was rarely reported, and if detailed, was done so in a nonstandardized fashion. Thus, the lack of data on additional and adjunctive treatment raises questions about the degree to which follow-up periods are comparable across studies and interventions, and the degree to which post-treatment gains in clinical research samples differ from routine clinical care. A systematic method of reporting booster sessions and additional treatment received during follow-up would assist in interpreting the long-term impact of interventions. Relative to the acute phase of treatment, the follow-up phase of RCTs is more naturalistic and less controlled (Crits-Christoph, 1997). Despite this, assessments in the months following treatment termination are still a valid reflection of long-term outcomes and therefore warrant close examination to more comprehensively understand treatment efficacy. We also note that the current meta-analysis does not include unpublished data in its analyses. Our fail-safe N calculation does not suggest the presence of potential publication bias, though the limitations associated with assessment of publication bias - including the fail-safe N calculation - should also be noted (Thornton & Lee, 2000). A final limitation pertains to the challenges associated with categorizing types of active treatments, which ultimately impacts findings. In recent years, there has been a growth in the complexity and variants of interventions, including treatment augmentations, (e.g., exposure plus cognitive restructuring), combination therapies (e.g., exposure plus pharmacotherapy), and phase-based therapies (e.g., skills training in emotion regulation followed by exposure). Navigating decisions regarding classification is a common consideration when conducting reviews (e.g., Cusack et al., 2016); for difficult decisions, the current paper drew from prior reviews and their categorical models (e.g., Bradley et al., 2005; Cusack et al., 2016; Watts et al., 2013). Also, in line with these reviews, the current paper categorized treatment conditions based on the most salient theoretical foundations and core components of each intervention.

5. Conclusion

Findings of the current meta-analysis are in line with PTSD treatment guidelines (e.g., Forbes et al., 2010), which also advocate the use of exposure-based therapies, trauma-focused CBT, cognitive therapy, and EMDR. Taken together with recent meta-analyses of acute PTSD treatments (e.g., Cusack et al., 2016), the current study and its emphasis on long-term outcomes provides additional support for these interventions. Our results clearly

point to sustained clinical success in the period following acute treatment, and in several cases also suggests the prospect of ongoing improvement after treatment has concluded.

Evidence of both the short- and long-term efficacy of these treatments reinforces the necessity of increasing the availability of evidence-based treatments. Recent cost projections regarding US-based veterans alone have reached billions of dollars in terms of lost productivity, impairment, and healthcare service fees (Tanielian & Jaycox, 2008). PTSD is usually chronic and impairing, rarely spontaneously remits (Morina et al., 2014), and often co-occurs with depression (~50%; Rytwinski et al., 2013), resulting in more substantial impairment and negative life effects. In the United States, access to high quality evidence-based care is especially difficult for patients with PTSD who do not live near an academic medical center or have access to service-related care. A major barrier to access stems from the fact that relatively few clinicians consistently use evidence-based treatments for PTSD (e.g., Becker, Zayfert, & Anderson, 2004; Foa, Gillihan, & Bryant, 2013). For instance, a recent study of community based treatment for anxiety highlighted that fewer than 5% of patients received exposure-based treatment, despite a clear empirical basis for its use (Wolitzky-Taylor, Zimmerman, Arch, Guzman, & Lagomasino, 2015). We hope that systematic reviews demonstrating clear evidence of acute and sustained gains from evidence-based psychotherapy will spur greater public awareness of and demand for high quality treatment options, greater motivation for providers to adopt these techniques, and greater incentive for healthcare systems and organizations to aggressively encourage their use as a means of improving patient outcomes and driving down unsustainable costs of ineffectual care.

In conclusion, the results of the current meta-analysis offer encouraging findings regarding the long-term impact and viability of a number of evidence-based psychotherapy treatments for PTSD. Active interventions included demonstrated robust symptom reduction from baseline through long-term follow-up. Additionally, treatment gains were maintained or continued to improve in the follow-up phase. Findings should also be considered in parallel with salient methodological factors, which are critical in interpretation of RCT findings. Efforts such as these, along with the future work into moderators and predictors of treatment response for individuals, will provide a more comprehensive understanding of the long-term effectiveness of psychotherapies for PTSD and how best to translate this knowledge into clinical practice.

Acknowledgments

The authors wish to thank Vincent Pugliese for his diligent review of articles included in this meta-analysis and members of the PTSD Treatment and Research Program for their helpful feedback on this project.

Role of funding sources

This work was supported by the National Institute of Mental Health (R01 MH066348) awarded to NCF. The funding source had no role in study design, analysis, writing or decision to submit this paper for publication.

References

- Becker CB, Zayfert C, Anderson E. A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behaviour Research and Therapy*. 2004; 42:277–292. [http://dx.doi.org/10.1016/S0005-7967\(03\)00138-4](http://dx.doi.org/10.1016/S0005-7967(03)00138-4). [PubMed: 14975770]
- Berlin JA, Golub RM. Meta-analysis as evidence: Building a better pyramid. *JAMA*. 2014; 312:603–606. <http://dx.doi.org/10.1001/jama.2014.8167>. [PubMed: 25117128]
- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*. 2010; 1:97–111. <http://dx.doi.org/10.1002/jrsm.12>. [PubMed: 26061376]
- Bradley R, Greene J, Russ E, Dutra L, Westen D. A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry*. 2005; 162:214–227. <http://dx.doi.org/10.1176/appi.ajp.162.2.214>. [PubMed: 15677582]
- Bryant RA, Mastrodomenico J, Hopwood S, Kenny L, Cahill C, Kandris E, Taylor K. Augmenting cognitive behaviour therapy for post-traumatic stress disorder with emotion tolerance training: A randomized controlled trial. *Psychological Medicine*. 2013; 43:2153–2160. <http://dx.doi.org/10.1017/S0033291713000068>. [PubMed: 23406821]
- Cooper AA, Kline AC, Graham B, Bedard-Gilligan M, Mello PG, Feeny NC, Zoellner LA. Homework “dose,” type, and helpfulness as predictors of clinical outcomes in prolonged exposure in PTSD. *Behavior Therapy*. 2017; 48:182–194. <http://dx.doi.org/10.1016/j.beth.2016.02.013>. [PubMed: 28270329]
- Crits-Christoph P. Limitations of the dodo bird verdict and the role of clinical trials in psychotherapy research: Comment on Wampold et al. (1997). *Psychological Bulletin*. 1997; 122:216–220. <http://dx.doi.org/10.1037/0033-2909.122.3.216>.
- Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: A meta-analysis of study quality and effect size. *Psychological Medicine*. 2010; 40:211–223. <http://dx.doi.org/10.1017/S0033291709006114>. [PubMed: 19490745]
- Cuijpers P, van Straten A, Warmerdam L, Andersson G. Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Depression and Anxiety*. 2009; 26:279–288. <http://dx.doi.org/10.1002/da.20519>. [PubMed: 19031487]
- Cusack K, Jones DE, Forneris CA, Wines C, Sonis J, Middleton JC, Gaynes BN. Psychological treatments for adults with posttraumatic stress disorder: A systematic review and meta-analysis. *Clinical Psychology Review*. 2016; 43:128–141. <http://dx.doi.org/10.1016/j.cpr.2015.10.003>. [PubMed: 26574151]
- Dechartres A, Altma DG, Trinquart L, Boutron I, Ravaud P. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. *JAMA*. 2014; 312:623–630. <http://dx.doi.org/10.1001/jama.2014.8166>. [PubMed: 25117131]
- Driessen E, Hollon SD, Bockting CLH, Cuijpers P, Turner EH. Does publication bias inflate the apparent efficacy of psychological treatments for major depressive disorder? A systematic review and meta-analysis of US National Institutes of Health-funded trials. *PLoS One*. 2015; 10:e0137864. <http://dx.doi.org/10.1371/journal.pone.0137864>. [PubMed: 26422604]
- Ehlers A, Clark DM, Hackmann A, McManus F, Fennell M. Cognitive therapy for post-traumatic stress disorder: Development and evaluation. *Behaviour Research and Therapy*. 2005; 43:413–431. <http://dx.doi.org/10.1016/j.brat.2004.03.006>. [PubMed: 15701354]
- Ehring T, Welboren R, Morina N, Wicherts JM, Freitag J, Emmelkamp PMG. Meta-analysis of psychological treatments for posttraumatic stress disorder in adult survivors of childhood abuse. *Clinical Psychology Review*. 2014; 34:645–657. <http://dx.doi.org/10.1016/j.cpr.2014.10.004>. [PubMed: 25455628]
- Fernandez E, Salem D, Swift JK, Ramtahal N. Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology*. 2015; 83:1108–1122. <http://dx.doi.org/10.1037/ccp0000044>. [PubMed: 26302248]
- Fleischhacker WW. A meta view on meta-analyses. *JAMA Psychiatry*. 2017; 74:684–685. <http://dx.doi.org/10.1001/jamapsychiatry.2017.1167>. [PubMed: 28514464]

- Flint J, Cuijpers P, Horder J, Koole SL, Munafò MR. Is there an excess of significant findings in published studies of psychotherapy for depression? *Psychological Medicine*. 2015; 45:439–446. <http://dx.doi.org/10.1017/S0033291714001421>. [PubMed: 25062429]
- Foa EB, Gillihan SJ, Bryant RA. Challenges and successes in dissemination of evidence-based treatments for posttraumatic stress: Lessons learned from prolonged exposure therapy. *Psychological Science in the Public Interest*. 2013; 14:65–111. <http://dx.doi.org/10.1177/1529100612468841>. [PubMed: 25722657]
- Foa EB, Hembree EA, Dancu CV. Prolonged exposure (PE) manual: Revised version (unpublished manuscript). 2002
- Forbes D, Creamer M, Bisson JI, Crow BE, Foa EB, Friedman MJ, Ursano RJ. A guide to guidelines for the treatment of PTSD and related conditions. *Journal of Traumatic Stress*. 2010; 23:537–552. <http://dx.doi.org/10.1002/jts.20565>. [PubMed: 20839310]
- Franklin CL, Cuccurullo L, Walton JL, Arseneau JR, Petersen NJ. Face to face but not in the same place: A pilot study of prolonged exposure therapy. *Journal of Trauma & Dissociation*. 2017; 18:116–130. <http://dx.doi.org/10.1080/15299732.2016.1205704>. [PubMed: 27348462]
- Glenn D, Golinelli D, Rose RD, Roy-Byrne P, Stein MB, Sullivan G, Craske MG. Who gets the most out of cognitive behavioral therapy for anxiety disorders? The role of treatment dose and patient engagement. *Journal of Consulting and Clinical Psychology*. 2013; 81:639–649. <http://dx.doi.org/10.1037/a00333403>. [PubMed: 23750465]
- Haagen JFG, Smid GE, Knipscheer JW, Kleber RJ. The efficacy of recommended treatments for veterans with PTSD: A meta-regression analysis. *Clinical Psychology Review*. 2015; 40:184–194. <http://dx.doi.org/10.1016/j.cpr.2015.06.008>. [PubMed: 26164548]
- Hedges LV, Pigott TD. The power of statistical tests in meta-analysis. *Psychological Methods*. 2001; 6:203–217. <http://dx.doi.org/10.1037/1082-989X.6.3.203>. [PubMed: 11570228]
- Hedges LV, Pigott TD. The power of statistical tests for moderators in meta-analysis. *Psychological Methods*. 2004; 9:426–445. <http://dx.doi.org/10.1037/1082-989X.9.4.426>. [PubMed: 15598097]
- Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *British Medical Journal*. 2003; 327:557–560. <http://dx.doi.org/10.1136/bmj.327.7414.557>. [PubMed: 12958120]
- Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*. 2012; 36:427–440. <http://dx.doi.org/10.1007/s10608-012-9476-1>. [PubMed: 23459093]
- Imel, ZE., Laska, K., Jakupcak, M., Simpson, TL. Meta-analysis of dropout in treatments for posttraumatic stress disorder; *Journal of Consulting and Clinical Psychology*. 2013. p. 394-404. <http://dx.doi.org/10.1037/a0031474>
- Institute of Medicine. Treatment of PTSD: An assessment of the evidence. Washington, DC: The National Academies Press; 2007.
- Karyotaki E, Smit Y, de Beurs DP, Henningsen KH, Robays J, Huibers MJH, Cuijpers P. The long-term efficacy of acute-phase psychotherapy for depression: A meta-analysis of randomized trials. *Depression and Anxiety*. 2016; 33:370–383. <http://dx.doi.org/10.1002/da.22491>. [PubMed: 27000501]
- Kazdin AE, Bass D. Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*. 1989; 57:138–147. <http://dx.doi.org/10.1037/0022-006X.57.1.138>. [PubMed: 2925964]
- Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of the 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*. 2005; 62:617–627. <http://dx.doi.org/10.1001/archpsyc.62.6.617>. [PubMed: 15939839]
- Kraemer HC, Gardner C, Brooks JO III, Yesavage JA. Advantages of excluding underpowered studies in meta-analysis: Inclusionist versus exclusionist viewpoints. *Psychological Methods*. 1998; 3:23–31. <http://dx.doi.org/10.1037/1082-989X.3.1.23>.
- Lee DJ, Schnitzlein CW, Wolf JP, Vythilingam M, Rasmusson AM, Hoge CW. Psychotherapy versus pharmacotherapy for posttraumatic stress disorder: Systemic review and meta-analyses to determine first-line treatments. *Depression and Anxiety*. 2016; 33:792–806. <http://dx.doi.org/10.1002/da.22511>. [PubMed: 27126398]

- Lipsey, MW., Wilson, DB. Practical meta-analysis. Thousand Oaks, California: Sage Publications; 2001.
- Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine. 2009; 6:e1000097. <http://dx.doi.org/10.1371/journal.pmed.1000097>. [PubMed: 19621072]
- Morina N, Wicherts JM, Lobbrecht J, Priebe S. Remission from post-traumatic stress disorder in adults: A systematic review and meta-analysis of long term outcome studies. Clinical Psychology Review. 2014; 34:249–255. <http://dx.doi.org/10.1016/j.cpr.2014.03.002>. [PubMed: 24681171]
- Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological Methods. 2002; 7:105–125. <http://dx.doi.org/10.1377/1082-989X.7.1.105>. [PubMed: 11928886]
- Orwin RG. A fail-safe *N* for effect size in meta-analysis. Journal of Educational Statistics. 1983; 8:157–159. <http://dx.doi.org/10.2307/1164923>.
- Powers MB, Halpern JM, Ferenschak MP, Gillihan SJ, Foa EB. A meta-analytic review of prolonged exposure for posttraumatic stress disorder. Clinical Psychology Review. 2010; 30:635–641. <http://dx.doi.org/10.1016/j.cpr.2010.04.007>. [PubMed: 20546985]
- Resick, PA., Schnicke, MK. Cognitive processing therapy for rape victims: A treatment manual. Newbury Park, CA: Sage; 1993.
- Resick PA, Williams LF, Suvak MK, Monson CM, Gradus JL. Long-term outcomes of cognitive-behavioral treatments for posttraumatic stress disorder among female rape survivors. Journal of Consulting and Clinical Psychology. 2012; 80:201–210. <http://dx.doi.org/10.1037/a0026602>. [PubMed: 22182261]
- Rosenthal R. The “file-drawer problem” and tolerance for null results. Psychological Bulletin. 1979; 86:638–641. <http://dx.doi.org/10.1037/0033-2909.86.3.638>.
- Rytwinski NK, Scur MD, Feeny NC, Youngstrom EA. The co-occurrence of major depressive disorder among individuals with posttraumatic stress disorder: A meta-analysis. Journal of Traumatic Stress. 2013; 26:299–309. <http://dx.doi.org/10.1002/jts.21814>. [PubMed: 23696449]
- Schneider RL, Arch JJ, Wolitzky-Taylor KB. The state of personalized treatment for anxiety disorders: A systematic review of treatment moderators. Clinical Psychology Review. 2015; 38:39–54. [PubMed: 25795293]
- Shadish WR Jr, Sweeney RB. Mediators and moderators in meta-analysis: There’s a reason we don’t let Dodo birds tell us which psychotherapies should have prizes. Journal of Consulting and Clinical Psychology. 1991; 59:883–893. <http://dx.doi.org/10.1037/0022-006X.59.6.883>. [PubMed: 1837794]
- Shapiro, F. Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures. New York: Guilford; 1995.
- Sloan DM, Feinstein BA, Gallagher MW, Beck JG, Keane TM. Efficacy of group treatment for posttraumatic stress disorder symptoms: A meta-analysis. Psychological Trauma: Theory, Research, Practice, and Policy. 2013; 5:176–183. <http://dx.doi.org/10.1037/a0026291>.
- Steenkamp MM, Litz BT, Hoge CW, Marmar CR. Psychotherapy for military-related PTSD: A review of randomized clinical trials. JAMA. 2015; 314:489–500. <http://dx.doi.org/10.1001/jama.2015.8370>. [PubMed: 26241600]
- Tanielian, T., Jaycox, LH., editors. Invisible wounds of war: Psychological and cognitive injuries, their consequences, and services to assist recovery. Santa Monica, CA: RAND Corporation; 2008.
- Taylor S, Abramowitz JS, McKay D. Non-adherence and non-response in the treatment of anxiety disorders. Journal of Anxiety Disorders. 2012; 26:583–589. <http://dx.doi.org/10.1016/j.janxdis.2012.02.010>. [PubMed: 22440391]
- Ten Have TR, Normand ST, Marcus SM, Brown CH, Lavori P, Duan N. Intent-to-treat vs. non-intent-to-treat analyses under treatment non-adherence in mental health randomized trials. Psychiatric Annals. 2008; 38:772–783. <http://dx.doi.org/10.3928/00485713-20081201-10>. [PubMed: 20717484]
- Thornton A, Lee P. Publication bias in meta-analysis: Its causes and consequences. Journal of Clinical Epidemiology. 2000; 53:207–216. [PubMed: 10729693]

- Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: Examination of underpowered studies in Cochrane reviews. *PLoS One*. 2013; 8:e59202. <http://dx.doi.org/10.1371/journal.pone.0059202>. [PubMed: 23544056]
- Viswanathan, M., Ansari, MT., Berkman, ND., Chang, S., Hartling, L., McPheeters, LM., Treadwell, JR. Assessing the risk of bias in individual studies in systematic reviews of health care interventions. Rockville, MD: Methods Guide for Comparative Effectiveness Reviews; 2012.
- Watts BV, Schnurr PP, Mayo L, Young-Xu Y, Weeks WB, Friedman MJ. Meta-analysis of the efficacy of treatments for post-traumatic stress disorder. *Journal of Clinical Psychiatry*. 2013; 74:541–550. <http://dx.doi.org/10.4088/JCP.12r08225>.
- Wilson, DB. Meta-analysis macros for SPSS. 2005. Retrieved May 2016 from <http://mason.gmu.edu/~dwilsonb/ma.html>
- Wolitzky-Taylor K, Zimmerman M, Arch JJ, Guzman E, Lagomasino I. Has evidence-based psychosocial treatment for anxiety disorders permeated usual care in community mental health settings? *Behaviour Research and Therapy*. 2015; 72:9–17. <http://dx.doi.org/10.1016/j.brat.2015.06.010>. [PubMed: 26159908]

HIGHLIGHTS

- We conducted a meta-analysis of psychotherapy for adults with PTSD.
- Long-term treatment efficacy and follow-up outcomes in RCTs were examined.
- All treatments demonstrated long-term efficacy.
- Attrition and analytic method of RCTs significantly impacted effect size estimates.
- Methodological design must be considered when interpreting RCTs.

Fig. 1.
Flow diagram of study identification and selection process.

Table 1

Characteristics of included randomized controlled trials and treatment conditions.

Study, author, year, and treatment conditions	Follow-up assessments (months)	N	Condition	Intended Number of Sessions	Population	Trauma type	Primary outcome measure	Acute Analyses	Follow-up Analyses
Bryant et al., 2003		58			Civilian	Mixed	CAPS	ITT	ITT
<i>IE</i>	6	20	EX	8					
<i>IE + CR</i>	6	20	EX	8					
<i>SC</i>	6	18	Non-Directive	8					
Bryant et al., 2008		118			Civilian	Mixed	CAPS	ITT	ITT
<i>IE</i>	6	31	EX	8					
<i>IVE</i>	6	28	CBT-M	8					
<i>IE + IVE</i>	6	31	EX	8					
<i>IE + IVE + CR</i>	6	28	EX	8					
Bryant et al., 2013		70			Civilian	Mixed	CAPS	ITT	ITT
<i>Emotion regulation training + CBT</i>	6	36	CBT-M	12					
<i>SC + CBT</i>	6	34	CBT-M	12					
Cloitre et al., 2002		58			Female, civilian	CSA, CPA	CAPS	Comp	Other
<i>STAIR + PE</i>	3, 9	31	EX	16					
<i>WLC*</i>	–	27	–	–					
Cloitre et al., 2010		104			Female, civilian	CSA, CPA	CAPS	ITT	ITT
<i>STAIR + EX</i>	3, 6	33	EX	16					
<i>SC + EX</i>	3, 6	38	EX	16					
<i>STAIR + SC</i>	3, 6	33	CBT-M	16					
Cottraux et al., 2008		60			Civilian	Mixed	PCL	Comp	Other
<i>CBT</i>	8, 20	31	CBT-M	16					
<i>Rogesian supportive therapy</i>	8, 20	29	Non-Directive	16					
Ehlers et al., 2003		85			Civilian	MVA	CAPS	Comp	Other
<i>CT</i>	6	28	CT	12					
<i>Self-help booklet</i>	6	28	Non-Directive	1					
<i>Repeated assessments</i>	6	29	WLC	–					
Ehlers et al., 2014		121			Civilian	Mixed	CAPS	ITT	ITT

Study, author, year, and treatment conditions	Follow-up assessments (months)	N	Condition	Intended Number of Sessions	Population	Trauma type	Primary outcome measure	Acute Analyses	Follow-up Analyses
<i>7-day intensive CT</i>	3, 6	30	CT	12					
<i>Standard weekly CT</i>	3, 6	31	CT	12					
<i>SC</i>	3, 6	30	Non-Directive	12					
<i>WLC*</i>	–	30	–	–					
Foa et al., 1999		96			Female, civilian	Physical assault, sexual assault	PSS-I	Comp	Other
<i>PE</i>	3, 6, 12	25	EX	9					
<i>SIT</i>	3, 6, 12	26	CBT-M	9					
<i>PE + SIT</i>	3, 6, 12	30	EX	9					
<i>WLC*</i>	–	15	–	–					
Foa et al., 2005		190			Female, civilian	Physical assault, sexual assault, CSA	PSS-I	ITT	Other
<i>PE</i>	3, 6, 12	79	EX	12					
<i>PE + CR</i>	3, 6, 12	74	EX	12					
<i>WLC*</i>	–	26	–	–					
Foa, Yusko et al., 2013		165			Civilian, comorbid PTSD and alcohol dependence	Mixed	PSS-I	ITT	ITT
<i>PE + NAL</i>	3, 6	40	EX	18					
<i>PE + PBO</i>	3, 6	40	EX	18					
<i>SC + NAL</i>	3, 6	42	Non-Directive	18					
<i>SC + PBO</i>	3, 6	43	Non-Directive	18					
Hinton et al., 2005		40			Cambodian refugees	Cambodian genocide	CAPS	ITT	ITT
<i>CBT</i>	3, 6	20	CBT-M	12					
<i>WLC, then CBT*</i>	3	20	–	–					
Kubany et al., 2004		125			Female, civilian	Battered women	CAPS	ITT	Other
<i>CTT-BW</i>	3, 6	63	CBT-M	11					
<i>WLC, then CTT-BW*</i>	3, 6	62	–	–					
McDonagh et al., 2005		74			Female, civilian	CSA	CAPS	ITT	Other
<i>CBT</i>	3, 6	29	EX	14					
<i>PCT</i>	3, 6	22	Non-Directive	14					
<i>WLC*</i>	–	23	–	–					
Marks et al., 1998		87			Civilian	Mixed	CAPS	Comp	Other

Study, author, year, and treatment conditions	Follow-up assessments (months)	N	Condition	Intended Number of Sessions	Population	Trauma type	Primary outcome measure	Acute Analyses	Follow-up Analyses
<i>PE</i>	1, 3, 6	23	EX	10					
<i>CR</i>	1, 3, 6	19	CT	10					
<i>PE + CR</i>	1, 3, 6	24	EX	10					
<i>Relaxation</i> *	1, 3	21	–	10					
Morland et al., 2014	3, 6	125			Male, Military	Mixed	CAPS	ITT	ITT
<i>CPT-C in person</i>	3, 6	64	CPT	12					
<i>CPT-C teletherapy</i> *		61	–	12					
Mueser et al., 2008		108			Civilian, co-occurring severe mental illness	Mixed	CAPS	ITT	ITT
<i>CBT</i>	3, 6	54	CBT-M	16					
<i>TAU</i>	3, 6	54	TAU	#					
Mueser et al., 2015		201			Civilian, co-occurring severe mental illness	Mixed	CAPS	ITT	ITT
<i>CBT with CR</i>	6, 12	104	CBT-M	16					
<i>Brief CBT without CR</i>	6, 12	97	Non-Directive	3					
Nacasch et al., 2011		30			Military	Military	CAPS	ITT	ITT
<i>PE</i>	12	15	EX	<i>M</i> = 11					
<i>TAU</i>	12	15	TAU	#					
Nacasch et al., 2015		40			Military	Mixed	CAPS	ITT	ITT
<i>PE (90 min)</i>	6	20	EX	<i>M</i> = 13.2					
<i>PE (60 min)</i>	6	20	EX	<i>M</i> = 13.6					
Neuner et al., 2004		43			Sudanese refugees	Mixed	PDS	ITT	ITT
<i>NET</i>	4, 12	17	EX	4					
<i>SC</i>	4, 12	14	Non-Directive	4					
<i>Psychoeducation</i>	4, 12	12	Non-Directive	1					
Resick et al., 2002		171			Female, civilian	Rape survivors	CAPS	ITT	ITT
<i>CPT</i>	3, 9	62	CPT	12					
<i>PE</i>	3, 9	62	EX	12					
<i>WLC</i> *	–	47	–	–					
Resick et al., 2008		162			Female, civilian	IPV	CAPS	ITT	ITT
<i>CPT</i>	6	56	CPT	12					
<i>CPT-C</i>	6	51	CPT	12					

Study, author, year, and treatment conditions	Follow-up assessments (months)	N	Condition	Intended Number of Sessions	Population	Trauma type	Primary outcome measure	Acute Analyses	Follow-up Analyses
WA	6	55	CBT-M	7					
Resick et al., 2015		108			Military	Military	PSS-I	ITT	ITT
<i>CPT-C</i>	6, 12	56	CPT	12					
<i>PCT</i>	6, 12	52	Non-Directive	12					
Rothbaum et al., 2005		73			Female, civilian	Rape survivors	CAPS	Comp	Other
<i>PE</i>	6	23	EX	9					
<i>EMDR</i>	6	25	EMDR	9					
<i>WLC*</i>	–	24	–	–					
Rothbaum et al., 2014		156			Military	Military	CAPS	ITT	ITT
<i>VR EX + DCS</i>	3, 6, 12	53	EX	6					
<i>VR EX + ALP</i>	3, 6, 12	50	EX	6					
<i>VR EX + PBO</i>	3, 6, 12	53	EX	6					
Schnurr et al., 2007		284			Military	Mixed	CAPS	ITT	ITT
<i>PE</i>	3, 6	141	EX	10					
<i>PCT</i>	3, 6	143	Non-Directive	10					
Sloan et al., 2012		46			Civilian	MVA	CAPS	ITT	ITT
<i>WET</i>	3, 6	22	EX	5					
<i>WLC*</i>	3	24	–	–					
Suris et al., 2013		86			Military	Military sexual trauma	CAPS	ITT	ITT
<i>CPT</i>	2, 4, 6	52	CPT	12					
<i>PCT</i>	2, 4, 6	34	Non-Directive	12					
Tarrier et al., 1999		72			Civilian	Mixed	CAPS	Comp	Other
<i>IE</i>	6, 12	35	EX	16					
<i>CT</i>	6, 12	37	CT	16					
van den Berg et al., 2015		155			Civilian, comorbid psychotic disorder	Mixed	CAPS	ITT	ITT
<i>PE</i>	6	53	EX	8					
<i>EMDR</i>	6	55	EMDR	8					
<i>WLC</i>	6	47	WLC	–					
van der Kolk et al., 2007		88			Civilian	Mixed	CAPS	ITT	Other
<i>EMDR</i>	6	29	EMDR	8					

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Study, author, year, and treatment conditions	Follow-up assessments (months)	N	Condition	Intended Number of Sessions	Population	Trauma type	Primary outcome measure	Acute Analyses	Follow-up Analyses
<i>Fluoxetine</i> [*]	6	30	–	–					
<i>Pill PBO</i> [*]	–	29	–	–					

Note. CAPS = Clinician-Administered PTSD Scale; PDS = Posttraumatic diagnostic scale; PCL = PTSD Checklist; Comp = completer; CSA = childhood sexual abuse; CPA = childhood physical abuse; IPV = interpersonal violence; ITT = intent to treat; TAU = treatment as usual; IE = imaginal exposure; CR = cognitive restructuring; SC = supportive counseling; IVE = in vivo exposure; CBT = cognitive behavioral therapy; EMDR = eye movement desensitization and reprocessing; STAIR = skills training in affect and interpersonal regulation; PE = prolonged exposure; WLC = wait list control; EXP = exposure therapy; DCS = D-cycloserine; PBO = placebo; CT = cognitive therapy; SIT = stress inoculation training; NAL = naltrexone; CPT = cognitive processing therapy; NET = narrative exposure therapy; CTT-BW = cognitive trauma therapy for battered women; PCT = present centered therapy; CPT-C = CPT without WA; WA = written accounts; VR = virtual reality; ALP = alprazolam; WET = written exposure therapy.

^{*} Not included in meta-analysis.

[#] Specific number not provided.

Table 2

Results from subgroup analyses.

Group (# of conditions)	Weighted Mean ES, pretreatment to follow-up	$\bar{Q}$ value ^b	<i>p</i> value	Weighted Mean ES, posttreatment to follow-up	$\bar{Q}$ value ^b	<i>p</i> value
Condition Type		20.65	< 0.001		0.01	0.90
Psychotherapy (55)	2.14			0.16		
Control (17)	1.04			0.17		
Treatment Type ^{ab} (55)						
CBT-M (11)	1.71	3.24	0.07	-0.01	5.07	0.02
vs. other PT (44)	2.24			0.21		
CPT (6)	1.59	2.57	0.11	0.16	0.02	0.90
vs. other PT (49)	2.20			0.18		
CT (5)	2.40	0.45	0.50	0.02	1.12	0.29
vs. other PT (50)	2.11			0.18		
EMDR (3)	2.57	0.74	0.39	0.02	0.72	0.40
vs. other PT (52)	2.11			0.17		
Exposure (30)	2.32	2.75	0.10	0.27	8.07	0.005
vs. other PT (25)	1.92			0.05		
Acute analyses ^a		0.28	0.60		1.22	0.27
ITT (42)	2.10			0.19		
Completer (13)	2.25			0.08		
Follow-up analyses ^a		5.06	0.03		0.97	0.33
All randomized (37)	1.95			0.14		
Subset (18)	2.51			0.22		
Population Type ^a		0.05	0.82		3.49	0.06
Military (10)	2.08			0.32		
Civilian (45)	2.15			0.13		
Childhood trauma primary ^a		2.04	0.36		5.25	0.07
None of sample (27)	1.98			0.12		
Part of sample (21)	2.36			0.14		
Entire sample (5)	2.25			0.46		

Group (# of conditions)	Weighted Mean ES, pretreatment to follow-up	$\bar{Q}$ value ^b	<i>p</i> value	Weighted Mean ES, posttreatment to follow-up	$\bar{Q}$ value ^b	<i>p</i> value
Follow-Up Assessment ^a		0.00	0.97		7.39	0.02
6 months (37)	2.14			0.09		
> 6 months (18)	2.13			0.31		

Note. ES = Cohen's *d* effect size; CBT-M = cognitive behavioral therapy mixed; CPT = cognitive processing therapy; CT = cognitive therapy; EMDR = eye movement desensitization and reprocessing; ITT = intent to treat; PT = psychotherapies.

^aAnalyses include only active treatments.

^b*Q* statistic reflects between-group variance.