RESEARCH ARTICLE

Quality Assessment of Studies Evaluating Cognitive Behavioural Therapy for Insomnia in Populations with PTSD: A Systematic Review

Gunnar Valdemar Grunert[1, 2], Jacob Stig Jarnot Meline[1]

Objective: This systematic review aims to evaluate the quality of intervention studies assessing CBT-I in adult populations with PTSD. Specifically examining overall quality, specific design elements and trends over time. Method: n > 1 studies assessing CBT-I in adult populations with PTSD were included. The Randomized Controlled Trial - Psychotherapy Quality Rating Scale (RCT-PQRS) was used to evaluate study quality. Correlations between overall study quality and publication year were calculated. Results: Nine studies were included. Overall, study quality was rated as moderate. Six design elements were not addressed adequately in most studies. No significant correlation was found. Conclusions: The evidence base is of moderate quality. However, several specific design elements limit their conclusions. Strategies to minimize limitations are discussed.

Keywords: clinical studies as topic, study quality, CBT-I

Posttraumatic stress disorder (PTSD) is a psychiatric diagnosis that an individual can develop after experiencing traumatic experiences. The DSM-5 (American Psychiatric Association, 2013) outlines that, to live up to the criteria for PTSD, a person has to have been exposed to a serious traumatic experience whereafter several symptoms have been developed. The traumatic event is persistently re-experienced, the person engages in avoidance of trauma-related stimuli, and negative thoughts or feelings will begin or worsen after the trauma. Furthermore, the person will experience trauma-related arousal or reactivity that began or worsened

after the trauma. These symptoms will have been present for more than one month, they create distress or functional impairment and are not due to medication, substance use or other illness. Posttraumatic stress disorder (PTSD) is associated with serious health conditions and worsening of psychiatric ailments including alcohol and substance use disorder, depression, and suicidal deaths (American Psychiatric Association, 2013). PTSD is rarely the only psychiatric disorder experienced by the patient, seeing that patients diagnosed with PTSD are 80% more likely to have symptoms that meet diagnostic criteria for at least one other mental disorder, such as depression, anxiety, or insomnia (American Psychiatric Association, 2013; Breslau et al., 1991; Creamer et al., 2001; Kessler et al., 1995). Previous studies have shown that 41% - 91% of patients with PTSD have insomnia (Neylan et al.,

---

[1] Shared first authorship
[2] Inpatient Unit for Eating Disorders, Capital Region of Denmark, Denmark
Corresponding author: Gunnar V. Grunert (vmsp8100@gmail.com)

1998; Ohayon & Shapiro, 2000), and even when PTSD symptoms remit post-treatment, insomnia symptoms usually persist (Zayfert & DeViva, 2004).

Insomnia impacts multiple domains of the patient's life, such as interpersonal and cognitive functioning, for example in the form of decreased energy and inability to concentrate (Benca, 2001; Goff et al., 2007), but it also has a significant financial impact on society, by leading to lower work productivity and increased absenteeism (Fullerton, 2006). Insomnia, in addition to PTSD, is also associated with serious health conditions and worsening of psychiatric ailments including alcohol and substance abuse, depression, and suicidal behaviour (Applewhite et al., 2012; Hoge et al., 2007; McLay et al., 2010; Wright et al., 2012).

Many treatments target insomnia, one of which is cognitive behavioural therapy for insomnia (CBT-I), which The American Academy of Sleep Medicine (AASM) recommends (Schutte-Rodin et al., 2008). The AASM defines CBT-I as a treatment combining different cognitive and behavioural monotherapies indicated for chronic insomnia (Schutte-Rodin et al., 2008). Clinicians' usage of CBT-I for patients with PTSD has become more widespread and studies in populations with PTSD are becoming more common (Pruiksma et al., 2018). An RCT evaluating CBT-I in an adult sample with PTSD found that 41% of participants in their intervention group achieved full remission on subjective sleep measures compared to 0% for a waitlist control group and that these results were maintained at 6-month follow-up (Talbot et al., 2014). In another RCT large significant pre- to post- intervention change was observed for sleep efficiency measured with sleep diaries (Ustinov 2014). In that same study, no significant change was found for the waitlist control group. Although the increase in studies assessing the usefulness of CBT-I for patients with PTSD is useful for determining the value of this approach, it is important to assess the quality of these studies. Study quality is often categorized into different domains with one such categorization being: the description of subjects, definition and delivery of treatment, outcome measurements, data analysis and treatment assignment (Kocsis et al., 2010).

Proper reporting and conducting of different procedures and strategies in these domains ensures quality (Kocsis et al., 2010).

A meta-analysis of a variety of psychotherapies for depression found that study quality moderated treatment effect sizes (Barth et al., 2013). Another meta-analysis examining psychotherapeutic interventions for depression examined the differences between high-quality studies and low-quality studies. Of the 115 randomized controlled trials (RCTs) they located only 11 studies that met the requirements of their quality criteria. The Cohen's *d* effect sizes, when the interventions were compared to control conditions, were 0.74 for all 115 studies and only 0.22 for the high-quality studies (Cuijpers et al., 2010). To further underline the importance of controlling for quality, one meta-analysis of CBT for depression found that across time, the quality of intervention studies increased, and the effect sizes decreased (Johnsen & Friborg, 2015). This may imply that there is a reverse causal relationship between the found effect sizes and the quality of the study.

To date, no systematic reviews have been conducted that specifically assess the quality of studies that evaluate CBT-I in adult populations with PTSD. As this particular research base is still young and quite small, it could be beneficial to gain an understanding of the quality of the evidence base for using these interventions in adult populations with PTSD. Furthermore, this review can help by elucidating aspects of studies that could be improved upon.

The aim of this study is to conduct a systematic review that seeks to answer the following question: What is the quality of the evidence base for CBT-I used in adult populations with PTSD? To critically evaluate the quality of the studies, we aim to answer three separate questions:

1) What is the overall quality of studies?
2) What is the quality of specific study design elements?
3) Are there any trends in the quality of studies over time?

The findings of this review could provide an

evaluation of the quality of current literature and indicate areas to be improved in the design of future intervention studies.

Material and Methods

Eligibility Criteria

The eligibility criteria of this review will be presented according to the PICOS (population, intervention[s], control[s], outcome[s] studies) framework, which can be seen fully described in Table 1 (Cherry & Dickson, 2017). The control(s) section of PICOS was not applicable here, as all studies with n > 1 evaluating CBT-I in adult populations with PTSD were included with no specifications for control groups.

Information Sources

The literature search was conducted in PsycInfo and PubMed. These two databases were originally chosen as they are some of the largest databases of psychological research.

Search Strategy

PsycInfo and PubMed were searched with two similar search strings containing three search blocks (see Supplemental Materials). We applied date, language,

and population limits. We excluded studies published before 1980 since this was the introduction of the DSM-III PTSD diagnosis. We also excluded studies with non-adult or non-human populations. Finally, we only included studies published in English or Danish.

Selection Process

Following the removal of duplicates, the two first authors performed title and abstract screening of all records. Studies were thereafter screened at the full-text level if they met eligibility criteria or if the full text had to be assessed to properly screen the study. When the decisions of the two first authors were inconsistent, disagreements were resolved through discussion. After eligible studies were identified through full-text screening, the reference lists of these studies were checked for any additional eligible studies. Lastly, the first authors of all eligible studies were contacted and asked whether they knew of any other eligible studies that were missed in our search.

Data Collection Process

The two first authors independently extracted relevant data for study, treatment, and participant

Table 1

Population, intervention(s), outcome(s) and study types used in this review.

Population	Intervention(s)	Control(s)	Outcome(s)	Studies
Adult patients either with clinician administered PTSD diagnosis or self-reported symptoms meeting cut-off scores	Cognitive behavioural therapy for insomnia[a,b]	N/A[c]	Insomnia symptoms (Insomnia severity index, sleep diary measurements or other insomnia specific measures)	All intervention studies (e.g. RCTs, pilot studies, chart reviews)

Note.

[a]= Studies were excluded if they examined any kind of treatment that combined conventional CBT-I with elements from other treatments, such as CBT-I + image rehearsal therapy (IRT) for nightmares or CBT-I + prolonged exposure therapy. This was done to focus the review on studies that were solely assessing the isolated effect of CBT-I.

[b]= All intervention formats, such as individual, group, and telemedicine were included, as one of the purposes of this review was to be able to broadly inform future intervention studies of CBT-I in the PTSD population.

[c]= Not applicable. For a more detailed explanation of the four applicable areas see the methods section.

characteristics. Inconsistencies in extracted data were resolved by revisiting study inclusion criteria together and reaching a consensus. Data extraction was purposefully conducted prior to the quality assessment as the process would increase the familiarization with the studies and therefore increase the ease of doing the quality assessment.

The RCT-PQRS was used to assess the quality of the included studies. Additional data items extracted were study type, control group, sample size, treatment name, treatment dosage, format (group/individual), mean age with standard deviations (SD), civilian or veteran population, percentage of sample with a clinician-administered PTSD diagnosis or self-report scores meeting cut-off for PTSD, percentage of women in sample, and concurrent treatment reported as yes/no.

Quality Assessment

Since the commonly used CONSORT tools are designed as checklists and not as tools with continuous rating scales, the RCT-PQRS was used for the purposes of this review (Kocsis et al., 2010). The RCT-PQRS was created by a subcommittee of the American Psychiatric Association for the systematic evaluation of the quality of RCTs assessing psychotherapeutic interventions (Kocsis et al., 2010). The tool has been used to evaluate the quality of clinical studies in a range of different reviews, both with controlled and uncontrolled trials (Gerber et al., 2011; Grenon et al., 2018, 2019; Harb et al., 2013; Johnsen & Thimm, 2018; Keefe et al., 2014, 2020; Kishita & Laidlaw, 2017; Koelen et al., 2014; Lilliengren et al., 2016; Steinert et al., 2017; Suh et al., 2019; Thoma et al., 2012; Zakhour et al., 2020).

The tool contains 25 items, with 24 in the following categories: description of subjects (four items), definition and delivery of treatment (five items), outcome measures (five items), data analysis (five items), treatment assignment (three items), and overall quality of the study (two items). The total score of the first 24 items is typically reported as an overall quality measurement and an omnibus rating (item 25) for the overall quality of the study is also given (Kocsis et al., 2010). Every item is scored as 0 (*inadequate or limited reporting*), 1 (*adequate or brief*

reporting) or 2 (*excellent or full reporting*). The omnibus rating is scored on a scale of 1 (*exceptionally poor*) to 7 (*exceptionally good*).

The two first authors used the RCT-PQRS independently to evaluate the studies and in cases of disagreement, conflicts were resolved by discussion. When lasting disagreements of items or definitions were present, they were resolved through consultation with a senior researcher. In one study using the RCT-PQRS, a cut-off score of 50% of the maximum total scores of the first 24 items was used to indicate when a study met adequate quality (Gerber et al., 2011). That means that if a study receives a score of 1 on the first 24 items or receives equal amounts of scores of 0 and 2, they would be above that cut-off score. The same cut-off score of 50% was used in this review.

Statistical Analysis and Synthesis Methods

SPSS (Version 26) was used to calculate interrater reliability and descriptive statistics, and to perform correlational analyses.

Study, Treatment and Participant Characteristics

Study, treatment, and participant characteristics were synthesized narratively and with the use of descriptive statistics.

Aim 1: Overall Quality of Studies

To synthesize the results of the overall quality assessment, the mean percentage of maximum scores and mean omnibus score across studies were calculated. The maximum score of the RCT-PQRS is normally 48 (Kocsis et al., 2010). However, as several uncontrolled intervention studies are included in this review, several items relating to randomization procedures and control groups are not applicable (items 15, 20, 21, 22, 23). As in previous similar reviews using the RCT-PQRS for both controlled and uncontrolled trials, these items are omitted in the rating of the uncontrolled studies and, therefore limit the overall maximum score to 38 for the uncontrolled intervention studies (Harb et al., 2013). To accommodate the different study types, the percentage of maximum score was used as a metric of overall quality in a study instead of the maximum

score itself. This metric was calculated by dividing the sum of all ratings for that particular study by the maximum possible sum of ratings for that particular study type, which was then multiplied by 100.

Aim 2: Quality of Specific Study Design Elements
The results of specific study design elements as per individual items categorized in the subcategories of the RCT-PQRS were synthesized narratively and tabulated.

Aim 3: Quality Over Time
To assess the quality trends over time, a correlation between the year of publication and the percentage of maximum score, as well as between year of publication and omnibus ratings, were calculated for each study.

Certainty Assessment
To statistically evaluate the interrater reliability of the RCT-PQRS, Kendall's tau-*b* correlation was calculated for the two raters' individual percentages of maximum scores ratings, as well as for omnibus ratings (Akoglu, 2018). Kendall's tau-*b* for the percentage of maximum scores was 0.58 (95% CI [0.51, 0.63], *p* < 0.001). Kendall's tau-*b* for omnibus ratings was 0.62 (95% CI [0.15, 0.86], *p* = 0.039). Typically, Kendall's tau-*b* scores < 0.39, 0.40 to 0.69, 0.7 to 0.99, and 1 are considered estimates of weak, moderate, strong, and perfect reliability, respectively (Akoglu, 2018). Therefore, using Kendall's tau-*b*, the interrater reliability in the current study was moderate for the percentages of maximum total scores and for the omnibus rating.

Results
Search Results
After applying in-database restrictions and removing 71 duplicates, 649 studies were screened. In the title and/or abstract screening process, 598 studies were ineligible and therefore excluded. The percentage of the agreement prior to the discussion of disagreements between authors for title and abstract screening was 90%. 42 studies were excluded during full-text screening, leaving eight studies. No

additional eligible studies were found after hand-searching the reference lists of the eight studies. After contact with the first authors of the eight studies, one additional relevant study was identified and included (Gehrman et al., 2020). In summary, nine studies were included in the review (DeViva et al., 2018; El-Solh et al., 2019; Gehrman et al., 2020; Gellis & Gehrman, 2011; Harb et al., 2019; Laurel Franklin et al., 2018; Owen, 2002; Talbot et al., 2014; Ustinov, 2014). See Figure 1 for a full overview of the screening process.

Figure 1

PRISMA flowchart over the screening process
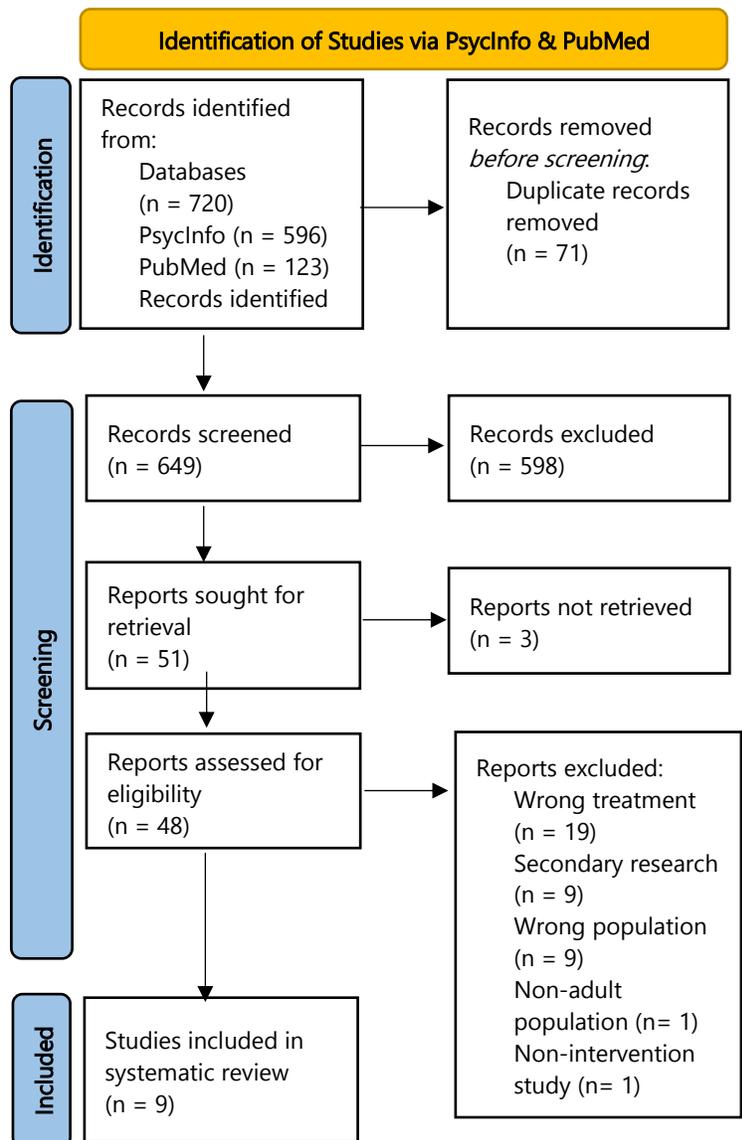
Data Extraction Results
Study Characteristics

Six out of nine studies were RCTs (Gehrman et al., 2020; Harb et al., 2019; Laurel Franklin et al., 2018; Owen, 2002; Talbot et al., 2014; Ustinov, 2014), one study was an uncontrolled trial (Gellis & Gehrman, 2011), and the two remaining studies were chart reviews (DeViva et al., 2018; El-Solh et al., 2019). Control groups utilized in the RCTs were monitoring only waitlist (Talbot et al., 2014; Ustinov, 2014), CBT-I telemedicine (Gehrman et al., 2020; Laurel Franklin et al., 2018), treatment as usual (Owen, 2002), and CBT-I plus image rehearsal therapy (IRT; Harb et al., 2019). Sample sizes ranged between eight and 136, with an average of n = 61. The majority of studies were published in the last decade with only one study published prior to 2010 (Owen, 2002). See Table 2 for complete details regarding study characteristics.

Treatment Type

All studies examined CBT-I in a face-to-face delivery. Six out of nine studies had a control group in the form of telemedicine (Gehrman et al., 2020; Laurel Franklin et al., 2018), CBT-I combined with IRT (Harb et al., 2019), waitlist (Talbot et al., 2014; Ustinov, 2014), or treatment as usual (Owen, 2002). Five studies (El-Solh et al., 2019; Gellis & Gehrman, 2011; Harb et al., 2019; Laurel Franklin et al., 2018; Talbot et al., 2014) examined CBT-I delivered in an individual format and the other four (DeViva et al., 2018; Gehrman et al., 2020; Owen, 2002; Ustinov, 2014) examined it in a group format. The mean (SD) number of treatment sessions was 5.56 (1.17) with a

Table 2

Study and Treatment Characteristics of CBT-I Studies Included in this Review.

Reference	Study type	Control group	Sample size	Treatment name	Dosage	Format
Deviva et al., 2018	Retrospective chart review	None	47	CBT-I	6 sessions	Group
El-Solh et al., 2019	Retrospective chart review	None	136	CBT-I	4 x 30-45 mins	Individual
Gehrman et al., 2020	RCT[a]	CBT-I telemedicine	96	CBT-I	6 x 90 mins	Group
Gellis & Gehrman, 2011	Uncontrolled trial	None	8	CBT-I	5 sessions	Individual
Harb et al., 2019	RCT	CBT-I+IRT[b]	108	CBT-I	6 x 60 mins	Individual
Laurel Franklin et al., 2018	RCT	CBT-I telemedicine	18	CBT-I	5 x 50 mins	Individual
Owen, 2002	RCT	TAU[c]	22	CBT-I	6 x 60 mins	Group
Talbot et al., 2014	RCT	Waitlist	45	CBT-I	8 sessions	Individual
Ustinov, 2014	RCT	Waitlist	65	CBT-I	4 x 60 mins	Group

Note.
[a] = Randomized Controlled Trial.
[b] = Image Rehearsal Therapy.
[c] = Treatment as usual.

Table 3

Participant Characteristics of CBT-I Studies Included in this Review.

Reference	Mean age (SD)	Population	% PTSD diagnosis	% Women (sample size)	Concurrent trauma focused treatment
Deviva et al., 2018	43.5 (13.40)	Veterans	100	9 (47)	Yes
El-Solh et al., 2019	46.9*	Veterans	100	23 (136)	Yes
Gehrman et al., 2020	55.14 (12.2)	Veterans	100	9 (96)	-[b]
Gellis & Gehrman 2011	58.6 (3.0)	Veterans	100	0 (8)	Yes
Harb et al., 2019	37.1 (9.9)	Veterans	100	14 (108)	-
Laurel Franklin et al., 2018	53.8 (12.0)	Veterans	67	0 (18)	-
Owen 2002	52.3 (2.22)	Veterans	100	0 (22)	-
Talbot et al., 2014	37.2 (10.5)	Civilians and Veterans	100	69 (45)	-
Ustinov 2014	53.6[a]	Veterans	87	7 (65)	Yes

Notes:
[a] = Standard deviation (SD) not reported
[b] = Information not reported.

range between four and eight. Three studies (DeViva et al., 2018; Gellis & Gehrman, 2011; Talbot et al., 2014) did not specify the length of sessions. The mean session length in minutes was 59.58 (15.84) minutes with a range between 37.50 and 90 minutes, where this data was available. See Table 2 for the full information.

Participant Characteristics
Most studies evaluated samples consisting entirely of veterans, with only one study involving both civilians and veterans (Talbot et al., 2014). For full information on participant characteristics see Table 3. All studies except two studies (Laurel Franklin et al., 2018; Ustinov, 2014) included a sample where 100% of participants were diagnosed with PTSD. For the two remaining studies, the percentage of participants diagnosed with PTSD in their samples was 67% (Laurel Franklin et al., 2018) and 87% (Ustinov, 2014). In all except one (Talbot et al., 2014) of the studies, the sample was predominantly male, with some being completely male (Gellis & Gehrman,

2011; Laurel Franklin et al., 2018; Owen, 2002). Four of the nine studies reported concurrent trauma-focused treatment, meaning that participants were engaged in other trauma-focused treatments at the same time as the CBT-I intervention (DeViva et al., 2018; El-Solh et al., 2019; Gellis & Gehrman, 2011; Owen, 2002).

Aim 1: Overall Quality of Studies
The mean percentage of maximum total scores for the studies was 57.19 (12.82) with a range between 37.50 and 79.17. Therefore, the mean was above the predefined cut-off score of adequate quality defined as 50% maximum total scores. Seven of the nine studies were above the 50% cut-off score. The mean (SD) omnibus rating for the studies was 4.44 (1.01) with a range between 3 and 6. An omnibus rating of 4 is indicative of average quality. Eight of the nine studies had an omnibus rating of 4 or above. Thus, for most of the studies, both the percentage of maximum total scores and the omnibus rating were rated to have adequate/average quality.

Table 4

Author Ratings for Study Quality of Included Studies Using the RCT-PQRS.

Item of the RCT-PQRS	Uncontrolled Studies (n=3)						RCTs (n=6)						Overall (n=9)					
	Ratings:						Ratings:						Ratings:					
	0[a] %	n	1[b] %	n	2[c] %	n	0 %	n	1 %	n	2 %	n	0 %	n	1 %	n	2 %	n
Description of participants																		
1) Diagnostic method, inclusion and exclusion criteria	0%	0	0%	0	100%	3	17%	1	0%	0	83%	5	11%	1	0%	1	89%	8
2) Reliability of diagnostic methodology	0%	0	100%	3	0%	0	0%	0	100%	6	0%	0	0%	0	100%	9	0%	0
3) Description of relevant comorbidities	0%	0	67%	2	33%	1	33%	2	17%	1	50%	3	22%	3	33%	3	44%	4
4) Numbers of participants screened, included, excluded	0%	0	0%	0	100%	3	33%	2	0%	0	67 %	4	22%	4	0%	2	78%	7
Definition and delivery of treatment																		
5) Treatments described/referenced for replication	0%	0	67%	2	33%	1	0%	0	17%	1	83%	5	0%	5	33%	3	67%	6
6) Treatment fidelity demonstrated	100%	3	0%	0	0%	0	67%	4	0%	0	33 %	2	78%	2	0%	7	22%	2
7) Therapist training/level of experience	0%	10	33%	1	67%	2	17%	1	17%	1	67%	4	11%	4	22%	1	67%	6
8) Therapist supervision during treatment	67%	2	0%	0	33%	1	33%	2	33%	2	33%	2	44%	2	22%	4	33%	3
9) Description of concurrent treatments	0%	0	33%	1	67%	2	0%	0	67%	4	33%	2	0%	2	56%	5	44%	4
Outcome measures																		
10) Validated outcome measure(s)	0%	0	0%	0	100%	3	0%	0	17%	1	83%	5	0%	5	11%	1	89%	8
11) Primary outcome measure(s) specified	0%	0	33%	1	67%	2	0%	0	17%	1	83%	5	0%	5	22%	2	78%	7
12) Outcome assessment by blind raters/with reliability	100%	3	0%	0	0%	0	50%	3	33%	2	17%	1	67%	1	22%	6	11%	1
13) Discussion of safety, adverse events	100%	3	0%	0	0%	0	83%	5	17%	1	0%	0	89%	0	11%	8	0%	0
14) Assessment of long-term posttermination outcome	67%	2	33%	1	0%	0	0%	0	100%	6	0%	0	22%	0	78%	2	0%	0
Data analysis																		
15) Intent-to-treat analyses	-[d]	-	-	-	-	-	33%	2	17%	1	50%	3	-	3	-	-	-	-
16) Description of dropouts and withdrawals	0%	0	67%	2	33%	1	0%	1	83%	5	17%	1	0%	1	78%	0	22%	2
17) Appropriate statistical tests	0%	0	100%	3	0%	0	0%	0	33%	2	67%	4	0%	0	78%	0	22%	2
18) Adequate sample size and power calculation	33%	1	67%	2	0%	0	33%	2	67%	4	0%	0	33%	3	67%	3	0%	0
19) Appropriate consideration of therapist/site effects	100%	3	0%	0	0%	0	67%	4	17%	1	17%	1	78%	1	11%	7	11%	1
Treatment assignment																		
20) A priori justification of comparison group(s)	-	-	-	-	-	-	0%	0	0%	0	100%	6	-	6	-	-	-	-
21) Comparison group(s) from same population/time	-	-	-	-	-	-	0%	0	17%	1	83 %	5	-	5	-	-	-	-
22) Randomized assignment	-	-	-	-	-	-	50%	3	0%	0	50%	3	-	3	-	-	-	-
Overall quality of study																		
23) Balance of allegiance to types of treatment	-	-	-	-	-	-	67%	4	33%	2	0%	0	-	0	-	-	-	-
24) Conclusions justified by sample/measures/analyses	0%	0	33%	1	67%	2	0%	0	67%	4	33%	2	0%	2	56%	5	44%	4

Note.

[a] = limited methods and/or limited description.

[b] = adequate methods or full description.

[c] = full description and adequate methods.

[d] = Not applicable item as the study is an uncontrolled study.

Aim 2: Quality of Specific Study Design Elements

Table 4 provides a breakdown of ratings on each item of the RCT-PQRS for included studies.

Description of Participants (Items 1-4)

Most of the studies fully and appropriately described their diagnostic measurements, eligibility criteria and number of participants included/excluded (items one and four). All the studies briefly documented the reliability of diagnostic measurements but didn't conduct any/enough within-study reliability checks of the diagnostic measurements to achieve excellence (item two). Nearly all the studies either briefly or fully described relevant comorbidities (item three).

Definition and Delivery of Treatment (Items 5-9)

Most studies fully described the treatment being evaluated or cited references that could fully describe it (item five). However, most studies had poor or no adherence reporting with regard to demonstrating the correct delivery of treatment (item six). A small majority of studies fully described therapist training levels and used well-qualified therapists (item seven). A small majority of the studies had limited or excellent descriptions of therapist supervision (item eight). Lastly, all studies had either brief or full descriptions of concurrent treatments (item nine).

Outcome Measures (Items 10-14)

Most studies showed or cited full validations of outcome measures (item ten) and specified the primary outcome measures in advance, though not explicitly (item 11). Most studies had poor or no blinding of outcome raters with regard to the treatment group (item 12) and had poor or no discussions of potential adverse events during treatment (item 13). Most studies had medium-term (2-12 months) follow-ups (item 14), indicating adequacy but not excellence.

Data Analysis (Items 15-19)

The subcategory of data analysis differs for uncontrolled studies and RCTs since intent-to-treat methods (item 15) are not applicable to uncontrolled studies. Half of the RCTs conducted full intent-to-treat analyses with two failing to conduct any full or partial intent-to-treat analyses. A majority of the nine studies had adequate descriptions of dropouts and withdrawals with one study having poor or no descriptions (item 16). Most of the studies used fully appropriate statistics and minimal data dredging in their primary findings (item 17). Most studies had either adequate sample sizes or adequate justifications for sample sizes, with none having both (item 18). Most studies neither discussed nor statistically considered therapist or site effects (item 19).

Treatment assignment (Items 20-22)

The subcategory of treatment assignment is only applicable to the RCTs as it considers randomization processes among other elements. A majority of RCTs had full justifications of comparison groups (item 20) and had comparison groups from the same population and time frame as the treatment group (item 21). Half of the RCTs utilized poor randomization procedures, such as sequential assignment or simple randomization with small samples (Beller et al., 2002). The other half had full and appropriate randomization procedures, such as performing randomization after screening and baseline assessment (item 22).

Overall quality of study (Items 23-24)

The item considering allegiance balancing with regard to treatment by therapists is only applicable for RCTs (item 23). Four of the six RCTs had either poor or no information with regard to allegiance balancing (item 23). The conclusions of most studies were however either partially or fully justified (item 24).

Aim 3: Quality Over Time

The Pearson correlation between year and percentage of maximum total scores was non-significant, $r = 0.56$, $p = 0.116$. In addition to this, the Pearson correlation between omnibus rating and publication year was nonsignificant, $r = 0.52$, $p = 0.154$. This indicates that there was no significant

association between the year of publication and quality ratings.

Discussion
Discussion of Main Findings
Aim 1
As reviews utilizing the RCT-PQRS are not consistent in their metric of quality measurement, it is difficult to compare our results with other reviews. Therefore, the number of studies above the cut-off score, as well as the omnibus ratings will be used as a comparison.

The results of this review indicated that 78% of the studies were above the cut-off score for adequacy. These results are contrary to some research showing that only 57% of 93 RCTs evaluating psychodynamic therapy were rated as of adequate quality (Gerber et al., 2011). However, the findings were more consistent with more recent studies finding rates of studies above the cut-off score like 75%, 77% and in one case 91% (Grenon et al., 2018; Keefe et al., 2020; Steinert et al., 2017).

In our review, the average omnibus rating of studies was average to moderately good. This finding was comparable to the findings of the reviews previously mentioned (Grenon et al., 2018; Keefe et al., 2020). Interestingly, the omnibus ratings for this review were less comparable to those of a similar review assessing the quality of IRT for the other sleep-related symptoms in PTSD, namely PTSD-related nightmares (Harb et al., 2013). In that review, the authors reported the mean omnibus rating to be moderately poor (Harb et al., 2013).

Even though the findings of this review were to some degree comparable to those of other reviews using the RCT-PRQS, the results still showed that most of the studies only just surpassed the cut-off score. This indicates that certain measures could be taken to increase the quality of any future studies evaluating CBT-I in adult populations with PTSD. Additionally, it needs to be noted that there is a substantial difference in the scores between some of the studies. The mean percentage of maximum total scores should, therefore, be interpreted with caution.

Aim 2
As stated previously, low study quality can inflate effect sizes. Therefore, even though most studies were above the cut-off score of adequate study quality, measures should still be taken to ensure optimal quality. Certain aspects of the study design were consistently implemented or reported in ways that resulted in limited scoring for those specific items.

Treatment fidelity checks. Most of the studies failed to conduct proper treatment fidelity checks. Testing treatment fidelity is of importance since it can help explain why prior results weren't replicated or why that particular evaluation found higher effects than prior studies (Temple et al., 2018). Without assessing the treatment fidelity of the study, therapy adherence and integrity cannot be controlled for, which limits the conclusions being drawn. To ensure that the treatment being studied is also the treatment being provided to study participants, systematic treatment fidelity checks should be administered (Temple et al., 2018).

Outcome assessments by blinded raters. Most studies were assessed to have limited outcome assessments in relation to blinding of raters and/or reliability checks. The reason for this in most of the studies was that self-reports were used as outcome measures. By not ensuring blinded outcome raters, the raters, whether clinicians, patients or unblinded raters, could be biased by expectancy effects or knowledge of sham treatments that influence their ratings (Marcus et al., 2006).

Discussion of adverse events during treatment. All except one study failed to report adverse events in the treatments, which only mentioned that there were none (Harb et al., 2019). Unfortunately, one systematic review has shown that the reporting of adverse events is conducted less often in psychotherapy trials than in trials assessing pharmacological treatments (Duggan et al., 2014). However, Spielman et al. (2011) argue that sleep restriction therapy, one of the interventions used in most CBT-I packages, can result in increased daytime sleepiness in the beginning phase of treatment. They underline that patients working in some industries where vehicles or heavy machinery are used have an

increased risk of accidents due to the treatment (Spielman et al., 2011). They also mention that patients with conditions like parasomnia (e.g., sleepwalking or sleep terrors), epilepsy, and sleep-disordered breathing should avoid participating in sleep restriction therapy due to the acute potential side effects of increased symptoms (Spielman et al., 2011).

   Consideration of therapist/site effects. Most studies failed to adequately control for any potential therapist/site effects. Therapist/site effects are defined as differing levels of efficiency between the individual study therapists or study sites. The implications of not properly considering these effects are that confounding variables aren't controlled for, and effects that aren't necessarily due to the specific treatment are concluded to be just that (Wampold & Imel, 2015). This can be avoided by conducting statistical analyses that control for the differing effectiveness of the treatment providers included in the study or by comparing the effectiveness of different sites of the study employing a multisite design (Wampold & Imel, 2015). One such statistical approach would be utilizing the multilevel model (Kahn, 2011), as has been done in three of the articles included in the review (Gehrman et al., 2020; Harb et al., 2019; Laurel Franklin et al., 2018).

   Randomization procedures. Half of the studies had limited randomization procedures. Proper randomization procedures are important to ensure that differences between groups at post-treatment or follow-up due not occur due to systematic differences that existed at baseline (Beller et al., 2002). This means that, when not properly conducted, poor randomization can weaken the ability to conclude that the intervention caused the treatment effect. Adequate or excellent randomization can be ensured using stratified approaches with randomization being conducted after baseline assessments.

   Balancing of allegiance bias. Finally, most RCTs did not discuss or evaluate the balance of allegiance to types of treatment. Allegiance in psychotherapy outcome studies refers to the finding that a correlation has been found between the personal allegiance researchers or study therapists have to one of the treatments being evaluated and the outcomes of that particular treatment (Wampold & Imel, 2015). Multiple explanations, such as implicit bias in study design decisions or expectancy effects on behalf of study therapists have been offered for this finding (Wampold & Imel, 2015). Even though this mostly applies to studies using an active control condition, it has still been recognised by certain researchers that there should be an explicit consideration of allegiance balancing in psychotherapy outcome studies (Wampold & Imel, 2015). This can be implemented by ensuring that study therapists in the different treatment groups are equally motivated towards the respective treatments and that comparative studies use teams with mixed allegiances (Munder et al., 2013).

Aim 3
The finding that the correlation between publication year and percentage of maximum total or omnibus scores was not significant is contrary to that of previous studies. Previous reviews have found quality trends change over time for psychodynamic therapy trials (Kocsis et al., 2010), interventions for emotional distress in breast cancer patients (Temple et al., 2018), and interventions for eating disorders (Grenon et al., 2018). It has been argued that intervention studies and RCTs in particular have improved in quality over time (Kocsis et al., 2010). There are two possible explanations for this association not being found in our review. Firstly, this review only found nine relevant studies whereas previous reviews assessing study quality of intervention studies found 69 relevant studies (Grenon et al., 2018; Kocsis et al., 2010; Temple et al., 2018). Our review was therefore based on a much smaller number of studies than previous reviews. Another explanation is that the included studies span a shorter time period than previous reviews with a time period of 11 years in our included studies whereas the time period in one of the earlier mentioned reviews was 30 years (Grenon et al., 2018; Kocsis et al., 2010; Temple et al., 2018).

Limitations

Only two databases, PsycInfo and PubMed were used for the search. Future reviews could possibly benefit from reviewing additional databases. One such database could be The Published International Literature on Traumatic Stress (PILOTS), that aims to include all articles relevant with regards to PTSD. Furthermore, no search of grey literature, such as unpublished manuscripts or reports not published in scientific journals, was conducted, limiting the scope of our review. Additionally, the RCT-PQRS is used to assess study quality specifically with regard to internal validity and focuses less on external validity in the form of generalizability (Gerber et al., 2011). Future reviews should increase the efforts to evaluate study design elements that are relevant to generalizability. Lastly, one very important limitation of our review is that even though we sought to evaluate the quality of studies evaluating CBT-I in general adult populations with PTSD, nearly all of our studies' samples consisted entirely of veterans. Even though our search seems to indicate that no other studies with adult non-veteran samples exist, it is still important to take note of.

Strengths

The systematic literature searches were thorough, with preliminary work to identify relevant databases and search terms, as well as searching reference lists and contacting the first authors of the included studies. The RCT-PQRS has demonstrated good reliability and validity and is a useful tool for evaluating the quality of clinical studies (Gerber et al., 2011; Kocsis et al., 2010). Furthermore, it was useful to use the same measure across different study types, as has been done in the past, to enable comparison of scores. Lastly, the ICC was deemed to be adequate.

Implications and Future Suggestions

Most of the studies were quantitatively rated to have adequate study quality through their omnibus rating, suggesting that the validity of these studies with regards to the design and conduct could be considered adequate. This would mean that the results of this review support the internal validity of the included studies and their results. However, as our discussions of our findings showcased, there were several important aspects of study quality that were inadequate. The limitations that were discussed, such as lack of treatment fidelity checks, blinded raters, reporting of adverse events, consideration of therapist/site effects, proper randomization procedures, and balancing of allegiance bias indicate impaired internal validity of the studies. This limits the certainty of the findings of the studies and practicing clinicians therefore need to consider the findings with regards to the efficacy of CBT-I in these studies tentative.

We offer six suggestions for future research based on the six items that were found to generally be of inadequate quality in our review:

- We suggest a greater focus on ensuring both the implementation and reporting of treatment fidelity checks during the study.
- The outcome assessments could to a greater extent be conducted by blinded raters.
- The reporting of any adverse events during the interventions could benefit from becoming usual practice.
- Therapist/site effects could statistically be accounted for in future studies to ensure that any potentially confounding variables are controlled for.
- Randomization procedures could be optimized by either using stratified approaches or using third parties to ensure proper randomization.
- Lastly, allegiance bias is a potential threat that can introduce bias to the designing and reporting of the study and therefore needs to be considered. This could be achieved by using mixed teams of researchers or by making sure that the therapists in the different treatment groups have allegiance to the respective techniques of that treatment (Munder et al., 2013).

By implementing these recommendations, the

evidence base for using CBT-I in adult populations with PTSD could be considerably strengthened. Even though it may be too early to try and examine whether the quality of the studies evaluating CBT-I in PTSD populations change over time, the implementation of the proposed changes could hopefully ensure a positive trend of improved study quality with time. As stated earlier, adequate study quality is important if the reported treatment effects and their conclusions are to inform policy-making and evidence-based practice. Finally, future reviews examining the quality of mixed CBT-I interventions are needed.

Conclusion

Intervention studies of cognitive behavioural therapy for insomnia in adult populations with PTSD are generally of adequate quality. This review gives support to the included studies and their results with regard to their internal validity. However, several aspects of study design can be improved. The findings indicated that most intervention studies were generally above the cut-off score for adequate quality on the RCT-PQRS, but only 70% of the studies had a maximum total score of 50% or more. Although most intervention studies were generally of adequate quality, some design elements were lacking in study quality, including the use and reporting of treatment fidelity checks, outcome assessments using blinded raters, reporting adverse events during the treatment, controlling for confounding variables in regard to therapist/site effects, and considering allegiance bias in the designing and reporting of the study. This review found a non-significant correlation between publication year and study quality, which may be due to the relatively recent publication years of included studies and the small sample size. Future studies could improve upon the aforementioned aspects to improve the quality of intervention studies and enhance the ability to draw conclusions from treatment effects.

Conflicts of Interest

No competing interests or financial disclosures are indicated by any of the authors.

Contributorship Statement

Acknowledgements

References

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91-93. https://doi.org/10.1016/j.tjem.2018.08.001

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5 ed.). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Applewhite, L., Keller, N., Borah, A., Applewhite, L., Keller, N., & Borah, A. (2012). Mental health care use by soldiers conducting counterinsurgency operations. *Military Medicine*, *177*(5), 501–506. https://doi.org/10.7205/milmed-d-11-00142

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., . . . Cuijpers, P. (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: A network meta-analysis. *PLoS Med*, *10*(5), e1001454. https://doi.org/10.1371/journal.pmed.1001454

Beller, E. M., Gebski, V., & Keech, A. C. (2002). Randomisation in clinical trials. *The Medical Journal of Australia*, *177*(10), 565–567. https://doi.org/10.5694/j.1326-5377.2002.tb04955.x

Benca, R. M. (2001). Consequences of insomnia and its therapies. *Journal of Clinical Psychiatry*, *62 Suppl 10*, 33–38.

Breslau, N., Davis, G. C., Andreski, P., & Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives Of General Psychiatry*, *48*(3), 216–222. https://psycnet.apa.org/doi/10.1001/archpsyc.1991.01810270028003

Cherry, M. G., & Dickson, R. (2017). Chapter 3: Defining my review question and identifying inclusion and exclusion criteria. In D. a. S. Review: & A. S. s. Guide (Eds.), Doing a systematic review: A student's guide (2. edition ed., pp. 43–59). SAGE.

Creamer, M., Burgess, P., & McFarlane, A. C. (2001). Post-traumatic stress disorder: findings from the Australian National Survey of Mental Health and Well-being. *Psychological Medicine*, *31*(7), 1237–1247. https://doi.org/10.1017/s0033291701004287

Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychological Medicine*, *40*(2), 211–223. https://doi.org/10.1017/s0033291709006114

DeViva, J. C., McCarthy, E., Bieu, R. K., Santoro, G. M., Rinaldi, A., Gehrman, P., & Kulas, J. (2018). Group cognitive-behavioral therapy for insomnia delivered to veterans with posttraumatic stress disorder receiving residential

treatment is associated with improvements in sleep independent of changes in posttraumatic stress disorder. *Traumatology*, *24*(4), 293–300. https://doi.org/10.1037/trm0000152

Duggan, C., Parry, G., McMurran, M., Davidson, K., & Dennis, J. (2014). The recording of adverse events from psychological treatments in clinical trials: Evidence from a review of NIHR-funded trials. *Trials*, *15*, 335–335. https://doi.org/10.1186/1745-6215-15-335

El-Solh, A. A., O'Brien, N., Akinnusi, M., Patel, S., Vanguru, L., & Wijewardena, C. (2019). Predictors of cognitive behavioral therapy outcomes for insomnia in veterans with post-traumatic stress disorder. *Sleep and Breathing, 23*(2), 635–643. https://doi.org/10.1007/s11325-019-01840-4

Fullerton, D. S. (2006). The economic impact of insomnia in managed care: A clearer picture emerges. *The American Journal Of Managed Care, 12*(8 Suppl), S246–252. https://doi.org/10.1001/archpsyc.1991.01810270028003

Gehrman, P., Barilla, H., Medvedeva, E., Bellamy, S., O'Brien, E., & Kuna, S. T. (2020). Randomized trial of telehealth delivery of cognitive-behavioral treatment for insomnia vs. in-person treatment in veterans with PTSD. *Journal of Affective Disorders Reports*, *1*. https://doi.org/10.1016/j.jadr.2020.100018

Gellis, L. A., & Gehrman, P. R. (2011). Cognitive behavioral treatment for insomnia in veterans with long-standing posttraumatic stress disorder: A pilot study. *Journal of Aggression, Maltreatment & Trauma*, *20*(8), 904–916. https://doi.org/10.1080/10926771.2011.613448

Gerber, A. J., Kocsis, J. H., Milrod, B. L., Roose, S. P., Barber, J. P., Thase, M. E., . . . Leon, A. C. (2011). A quality-based review of randomized controlled trials of psychodynamic psychotherapy. *The American Journal Of Psychiatry*, *168*(1), 19–28. https://doi.org/10.1176/appi.ajp.2010.08060843

Goff, B. S., Crow, J. R., Reisbig, A. M., & Hamilton, S. (2007). The impact of individual trauma symptoms of deployed soldiers on relationship satisfaction. *Journal Of Family Psychology*, *21*(3), 344–353. https://doi.org/10.1037/0893-3200.21.3.344

Grenon, R., Carlucci, S., Brugnera, A., Schwartze, D., Hammond, N., Ivanova, I., . . . Tasca, G. A. (2019). Psychotherapy for eating disorders: A meta-analysis of direct comparisons. *Psychotherapy Research*, *29*(7), 833–845. https://doi.org/10.1080/10503307.2018.1489162

Grenon, R., McKenna, A., Maxwell, H., Carlucci, S., Brugnera, A., Schwartze, D., . . . Tasca, G. A. (2018). The quality of randomized controlled trials of psychotherapy for eating disorders. *International Journal of Methods in Psychiatric Research*, *27*(3), e1734. https://doi.org/10.1002/mpr.1734

Harb, G. C., Cook, J. M., Phelps, A. J., Gehrman, P. R., Forbes, D., Localio, R., . . . Ross, R. J. (2019). Randomized controlled trial of imagery rehearsal for posttraumatic nightmares in combat veterans. *Journal Of Clinical Sleep Medicine*, *15*(5), 757–767. https://doi.org/10.5664/jcsm.7770

Harb, G. C., Phelps, A. J., Forbes, D., Ross, R. J., Gehrman, P. R., & Cook, J. M. (2013). A critical review of the evidence base of imagery rehearsal for posttraumatic nightmares: Pointing the way for future research. *Journal of Traumatic Stress*, *26*(5), 570–579. https://doi.org/10.1002/jts.21854

Hoge, E. A., Austin, E. D., & Pollack, M. H. (2007). Resilience: Research evidence and conceptual considerations for posttraumatic stress disorder. *Depression and Anxiety*, *24*(2), 139–152. https://doi.org/10.1002/da.20175

Johnsen, T. J., & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin*, *141*(4), 747–768. https://doi.org/10.1037/bul0000015

Johnsen, T. J., & Thimm, J. C. (2018). A meta-Analysis of group cognitive–behavioral therapy as an antidepressive treatment: are we getting better? *Canadian Psychology*, *59*(1), 15–30. https://doi.org/10.1037/cap0000132

Kahn, J. H. (2011). Multilevel modeling: Overview and applications to research in counseling psychology. *Journal Of Counselling Psychology*, *58*(2), 257–271. https://doi.org/10.1037/a0022680

Keefe, J. R., McCarthy, K. S., Dinger, U., Zilcha-Mano, S., & Barber, J. P. (2014). A meta-analytic review of psychodynamic therapies for anxiety disorders. *Clinical Psychology Review*, *34*(4), 309–323. https://doi.org/10.1016/j.cpr.2014.03.004

Keefe, J. R., McMain, S. F., McCarthy, K. S., Zilcha-Mano, S., Dinger, U., Sahin, Z., . . . Barber, J. P. (2020). A meta-analysis of psychodynamic treatments for borderline and cluster C personality disorders. *Personality Disorders: Theory, Research, and Treatment*, *11*(3), 157–169. https://doi.org/10.1037/per0000382

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry*, *52*(12), 1048–1060. https://doi.org/10.1001/archpsyc.1995.03950240066012

Kishita, N., & Laidlaw, K. (2017). Cognitive behaviour therapy for generalized anxiety disorder: Is CBT equally efficacious in adults of working age and older adults? *Clinical Psychology Review*, *52*, 124–136. https://doi.org/10.1016/j.cpr.2017.01.003

Kocsis, J. H., Gerber, A. J., Milrod, B., Roose, S. P., Barber, J., Thase, M. E., . . . Leon, A. C. (2010). A new scale for assessing the quality of randomized clinical trials of psychotherapy. *Comprehensive Psychiatry*, *51*(3), 319–324. https://doi.org/10.1016/j.comppsych.2009.07.001

Koelen, J. A., Houtveen, J. H., Abbass, A., Luyten, P., Eurelings-Bontekoe Elisabeth, H. M., Van Broeckhuysen-Kloth Saskia, A. M., . . . Geenen, R. (2014). Effectiveness of psychotherapy for severe somatoform disorder: meta-analysis. *The British Journal of Psychiatry*, *204*(1), 12–19. https://doi.org/10.1192/bjp.bp.112.121830

Laurel Franklin, C., Walton, J. L., Raines, A. M., Chambliss, J. L., Corrigan, S. A., Cuccurullo, L.-A. J., . . . Thompson, K. E. (2018). Pilot study comparing telephone to in-person delivery of cognitive-behavioural therapy for trauma-related insomnia for rural veterans. *Journal of Telemedicine and Telecare*, *24*(9), 629–635. https://doi.org/10.1177/1357633X17732366

Lilliengren, P., Johansson, R., Lindqvist, K., Mechler, J., & Andersson, G. (2016). Efficacy of experiential dynamic

therapy for psychiatric conditions: A meta-analysis of randomized controlled trials. *Psychotherapy*, *53*(1), 90–104. https://doi.org/10.1037/pst0000024

Marcus, S. M., Gorman, J. M., Tu, X., Gibbons, R. D., Barlow, D. H., Woods, S. W., & Katharine Shear, M. (2006). Rater bias in a blinded randomized placebo-controlled psychiatry trial. *Statistics In Medicine*, *25*(16), 2762–2770. https://doi.org/10.1002/sim.2405

McLay, R. N., Klam, W. P., & Volkert, S. L. (2010). Insomnia is the most commonly reported symptom and predicts other symptoms of post-traumatic stress disorder in U.S. service members returning from military deployments. *Military Medicine*, *175*(10), 759–762. https://doi.org/10.7205/milmed-d-10-00193

Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review*, *33*(4), 501–511. https://doi.org/10.1016/j.cpr.2013.02.002

Neylan, T. C., Marmar, C. R., Metzler, T. J., Weiss, D. S., Zatzick, D. F., Delucchi, K. L., . . . Schoenfeld, F. B. (1998). Sleep disturbances in the Vietnam generation: Findings from a nationally representative sample of male Vietnam veterans. *American Journal of Psychiatry*, *155*(7), 929–933. https://doi.org/10.1176/ajp.155.7.929

Ohayon, M. M., & Shapiro, C. M. (2000). Sleep disturbances and psychiatric disorders associated with posttraumatic stress disorder in the general population. *Comprehensive Psychiatry*, *41*(6), 469–478. https://doi.org/10.1053/comp.2000.16568

Owen, A. B. (2002). Cognitive-behavioral treatment of insomnia in veterans with post-traumatic stress disorder. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *63*(5-B), 2598.

Pruiksma, K. E., Wachen, J. S., Wardle, S., & Resick, P. A. (2018). Psychotherapy interventions for comorbid sleep disorders and posttraumatic stress disorder. In E. Vermetten, A. Germain, & T. C. Neylan (Eds.), *Sleep and combat-related post traumatic stress disorder* (pp. 277–292). Springer New York. https://doi.org/10.1007/978-1-4939-7148-0_24

Schutte-Rodin, S., Broch, L., Buysse, D., Dorsey, C., & Sateia, M. (2008). Clinical guideline for the evaluation and management of chronic insomnia in adults. *Journal of Clinical Sleep Medicine*, *4*(5), 487–504. https://doi.org/10.5664/jcsm.27286

Spielman, A. J., Yang, C.-M., & Glovinsky, P. B. (2011). Chapter 1 - Sleep restriction therapy. In M. Perlis, M. Aloia, & B. Kuhn (Eds.), *Behavioral treatments for sleep disorders* (pp. 9–19). Academic Press. https://doi.org/10.1016/B978-0-12-381522-4.00001-8

Steinert, C., Munder, T., Rabung, S., Hoyer, J., & Leichsenring, F. (2017). Psychodynamic therapy: As efficacious as other empirically supported treatments? A meta-analysis testing equivalence of outcomes. *American Journal of Psychiatry*, *174*(10), 943–953. https://doi.org/10.1176/appi.ajp.2017.17010057

Suh, H., Sohn, H., Kim, T., & Lee, D.-G. (2019). A review and meta-analysis of perfectionism interventions: Comparing face-to-face with online modalities. *Journal of Counseling Psychology*, *66*. https://doi.org/10.1037/cou0000355

Talbot, L. S., Maguen, S., Metzler, T. J., Schmitz, M., McCaslin, S. E., Richards, A., . . . Neylan, T. C. (2014). Cognitive behavioral therapy for insomnia in posttraumatic stress disorder: A randomized controlled trial. *Sleep: Journal of Sleep and Sleep Disorders Research*, *37*(2), 327–341. https://doi.org/10.5665/sleep.3408

Temple, J., Salmon, P., Tudur-Smith, C., Huntley, C. D., & Fisher, P. L. (2018). A systematic review of the quality of randomized controlled trials of psychological treatments for emotional distress in breast cancer. *Journal of Psychosomatic Research*, *108*, 22–31. https://doi.org/10.1016/j.jpsychores.2018.02.013

Thoma, N. C., McKay, D., Gerber, A. J., Milrod, B. L., Edwards, A. R., & Kocsis, J. H. (2012). A quality-based review of randomized controlled trials of cognitive-behavioral therapy for depression: An assessment and metaregression [Review]. *American Journal of Psychiatry*, *169*(1), 22–30. https://doi.org/10.1176/appi.ajp.2011.11030433

Ustinov, Y. (2014). Treatment of insomnia in veterans with trauma-related disorders: A brief group cognitive-behavioral intervention. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *75*(6-B(E)).

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (Second ed.). Routledge.

Wright, K. M., Foran, H. M., Wood, M. D., Eckford, R. D., & McGurk, D. (2012). Alcohol problems, aggression, and other externalizing behaviors after return from deployment: Understanding the role of combat exposure, internalizing symptoms, and social environment. *Journal of Clinical Psychology*, *68*(7), 782–800. https://doi.org/10.1002/jclp.21864

Zakhour, S., Nardi, A., Levitan, M., & Appolinario, J. (2020). Cognitive-behavioral therapy for treatment-resistant depression in adults and adolescents: A systematic review. *Trends in Psychiatry and Psychotherapy*, *42*. https://doi.org/10.1590/2237-6089-2019-0033

Zayfert, C., & DeViva, J. C. (2004). Residual insomnia following cognitive behavioral therapy for PTSD. *Journal of Traumatic Stress*, *17*(1), 69–73. https://doi.org/10.1023/B:JOTS.0000014679.31799.e7