

## REVIEW

# Safety of psychological interventions for adult post-traumatic stress disorder: meta-analysis on the incidence and relative risk of deterioration, adverse events and serious adverse events

Thole H. Hoppen, Anna S. Lindemann and Nexhmedin Morina

## Background

Attention on harmful effects of psychological interventions for adult post-traumatic stress disorder (PTSD) has increased, yet a comprehensive meta-analysis is lacking.

## Aims

To summarise incidences and relative risks of deterioration, adverse events (AEs) and serious adverse events (SAEs) in trials of psychological interventions for adult PTSD.

## Method

We searched MEDLINE, PsycInfo, Web of Science and PTSDpubs from inception to 21 April 2022 for sufficiently large ( $n \geq 20$ ) randomised controlled trials (RCTs) reporting on the incidence of harms.

## Results

We included 56 RCTs (4230 patients). Incidences of harms were generally low (0–5%). Psychological interventions were associated with decreased risk of deterioration relative to passive (RR = 0.21, 95% CI 0.15–0.28) and active control conditions (RR = 0.36, 95% CI 0.14–0.92). Decreased risk was even more pronounced in sensitivity analyses on trials exclusively delivering treatments face to face. When compared with other psychological interventions, trauma-focused cognitive-behavioural therapy (TF-CBT) was associated with decreased risk of SAEs

(RR = 0.54, 95% CI 0.31–0.95) and with no differential risk of deterioration and AEs.

## Conclusions

The current evidence base suggests that psychological interventions are safe for most adults with PTSD. In none of the analyses were psychological interventions associated with an increased risk of harm compared with control conditions. TF-CBT was found at least as safe as other psychological interventions. Individual face-to-face delivery might be the safest delivery format. However, more data are needed to draw firmer conclusions. We encourage research teams to routinely and thoroughly assess and report the incidence of harms and their causes.

## Keywords

PTSD; psychological intervention; harm; deterioration; adverse event.

## Copyright and usage

© The Author(s), 2022. Published by Cambridge University Press on behalf of the Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Post-traumatic stress disorder (PTSD) is a common, severe and potentially chronic mental disorder that places a large burden on affected individuals and society at large.<sup>1,2</sup> Psychological interventions are the first-line treatment recommendation for adult PTSD in international treatment guidelines.<sup>3</sup> Over the past four decades, dozens of randomised controlled trials (RCTs) have accumulated and meta-analytic summaries robustly attested to the general treatment efficacy of psychological interventions for adult PTSD.<sup>4–7</sup> However, efficacy estimates in meta-analyses are based on comparisons at the group level. Therefore an intervention that has been found effective overall may nevertheless have no effect or even harmful effects for single individuals. Harmful effects may concern the condition at hand (e.g. deterioration of PTSD) or they may concern adverse events (AEs) or serious adverse events (SAEs) during treatment irrespective of improvement, stagnation or exacerbation of the targeted condition. Unlike research on pharmacotherapy for adults with PTSD, where potential harms are frequently monitored and reported,<sup>8–10</sup> research on psychotherapy for PTSD has ignored potential harms for a long time,<sup>11</sup> this is a pattern that is evident in psychotherapy research more generally (i.e. across mental disorders).<sup>12–15</sup> However, various studies have underscored the relevance of investigating potential harms in clinical trials on psychological treatments. For instance, one study found that about 93% of individuals treated with psychotherapy

for obsessive-compulsive disorder reported experiencing at least one adverse event during treatment.<sup>16</sup> Another study found that about 5% of patients with common mental disorders reported long-lasting harmful effects induced by psychological treatment.<sup>17</sup> Two meta-analyses of harmful effects of psychological treatments for depression in RCTs found that 5% of individuals randomised to psychotherapy experienced worse depressive symptomatology after treatment determination<sup>18</sup> and that the risk of deterioration in those treated with psychotherapy was 61% lower relative to those randomised to control conditions.<sup>19</sup> Potential harmful effects of psychological interventions for adult PTSD have also been increasingly investigated in recent years.<sup>11,20–24</sup> In their meta-analysis, Cusack et al (2016)<sup>25</sup> reported that a substantial minority of included trials (27%) reported on the incidence of harms during psychological treatments for adult PTSD. Similarly, Forman-Hoffman et al (2018)<sup>26</sup> reported on the amount of trials reporting various kinds of adverse event, such as suicide attempts, suicidal ideation or self-harmful behaviours, in their comprehensive systematic review on the evidence base for psychotherapeutic and psychopharmacological treatments. However, only one of the previous systematic reviews and meta-analyses quantitatively summarised the incidence of harms. Jayawickreme et al (2014)<sup>11</sup> summarised four RCTs on prolonged exposure therapy for adult PTSD and found that none of the participants treated with

prolonged exposure experienced clinically significant deterioration of PTSD symptoms from pre- to post-treatment, whereas 8.1% of participants in the waiting list conditions did. Consequently, the authors concluded that prolonged exposure is not only an efficacious treatment for adult PTSD but also safe as it reduces the risk of deterioration compared with delaying treatment. To the best of our knowledge, meta-analyses on the incidences of AEs and SAEs are yet lacking.

Against this background, we conducted the first comprehensive meta-analysis summarising the incidences of various kinds of harm (deterioration, AEs and SAEs) in the literature of RCTs on the efficacy of psychological treatments for adult PTSD. Since harms might occur unrelated to treatment, we also calculated the relative risks (RR) compared with passive control conditions as a proxy for the naturalistic incidence of harms and compared with active control conditions to control for effects not related to PTSD treatment. In the light of the above-mentioned previous findings,<sup>11</sup> we hypothesised that psychological treatments would be safer than control conditions. That is, we hypothesised that these would be associated with lower incidences and risks of deterioration, AEs and SAEs. We also investigated whether incidences of harms differed across different families of psychological intervention.

## Method

### Registration and study protocol

We pre-registered the aims and methods of the present work in the PROSPERO database (CRD42020206290, <http://www.crd.york.ac.uk/prospero>) and followed PRISMA guidelines.<sup>27</sup> Two independent raters conducted the literature search and data extractions. Discrepancies were discussed in regular discussions among all authors. We formulated the main research question describing the Population, Intervention, Comparison, Outcome and Study design (PICOS) as follows: in patients with PTSD (P), are psychological interventions (I), compared with passive and active control conditions (C), associated with lower rates/risk of deterioration of PTSD symptoms, AEs and SAEs (O) in RCTs (S)? Whenever possible, we also compared the safety profile among different families of psychological intervention. The PRISMA-P checklist can be found in Supplement 1 of the supplementary material, available at <https://doi.org/10.1192/bjp.2022.111>.

Ethics approval was not required for this meta-analysis as no new data were collected.

### Identification and selection of studies

Trials were included if they met the following inclusion criteria: (a) participant allocation was randomised (i.e. the trial was an RCT); (b) a psychological intervention was compared with at least one passive control condition, active control condition or psychological intervention from another treatment family (see below); (c) PTSD was the primary diagnosis and treatment target; (d) at least 70% of participants were diagnosed with PTSD at baseline on the basis of a clinical interview;<sup>4,5</sup> (e) total  $n \geq 20$  (i.e.  $n \geq 10$  per condition); and (f) the sample was mostly adult (i.e. mean  $\geq 18$  years old). Trials were excluded if they either (a) failed to report on any harms (i.e. deterioration, AEs & SAEs) or (b) exclusively involved participants with a comorbidity of PTSD and substance use disorders or traumatic brain injury.<sup>4</sup> Other comorbidities were allowed provided that inclusion criterion (c) was met. For the timespan from database inception to 22 September 2020, we relied on our previous work.<sup>4</sup> For the timespan thereafter and up until 21 April 2022 we performed a new systematic literature search utilising a similar but broader search strategy. In brief, we searched MEDLINE,

PsycInfo, Web of Science and PTSDpubs using various search terms for PTSD and treatment, including the American Psychological Association's Thesaurus of Psychological Index Terms and Medical Subject Headings (MeSH) terms in multi-field searches. The full search string for each database can be found in Supplement 2. We did not apply any language restrictions. We also systematically reviewed 176 related review articles (Supplement 3) for further eligible trials, as well as the reference lists of included trials and the Veterans Affairs (VA) trial repository. Given the focus of the current analysis, we excluded trials that failed to report on any harms, which explains the significantly lower number of included trials in the present work compared with our previous work.<sup>4</sup>

### Risk of bias assessment

Risk of bias of included trials was assessed using eight dichotomous quality criteria based on the Cochrane Collaboration criteria for assessing the methodological validity of clinical trials<sup>28</sup> and authoritative criteria for empirically supported psychological interventions.<sup>29</sup> Trials received a score (versus no score) when: (a) the entire sample met diagnostic criteria for PTSD at baseline, (b) the treatment was manual-based, (c) therapists were trained or had extensive experience with the manual, (d) the treatment integrity was formally checked, (e) intention-to-treat (ITT) results were reported, (f)  $n \geq 50$ , (g) the randomisation was performed by an independent party and (g) outcome evaluators were masked ('blinded') (self-report-based outcome assessment also received a positive score). Thus, trials could yield a quality sum score between 0 and 8. This scale has been used repeatedly in meta-analyses on efficacy to assess risk of bias.<sup>4,30</sup>

### Coding of trial characteristics

We divided treatment conditions into four categories: trauma-focused cognitive-behavioural therapy (TF-CBT), eye movement desensitisation and reprocessing (EMDR), trauma-focused other interventions (tf-Other, e.g. imagery rehearsal therapy) and non-trauma-focused other interventions (non-tf-Other, e.g. present-centred therapy). Control conditions were divided into either passive control conditions (e.g. waiting lists) or active control conditions (e.g. care as usual). See the first column of Supplement 4 for all categorisations. To avoid data dependencies in the overarching analysis (i.e. across psychological interventions) we had to choose for one comparison in multi-armed trials. In these cases, the main comparison as indicated by the original authors of the respective RCT was prioritised. For all trials we extracted general study characteristics such as the definition of harms or PTSD measure used, the country in which the study was conducted or the type(s) of trauma, as well as relevant information for the moderator and sensitivity analyses (see below).

### Coding of outcomes and potential moderators

Some trials failed to specify their definition of deterioration.<sup>31,32</sup> Since unsystematic assessments of harms may substantially bias meta-analytic results, we only included trials giving a quantitative definition of deterioration. Deterioration was defined as exacerbation of PTSD between two assessment points (i.e. post-treatment versus pre-treatment assessment, and follow-up versus pre-treatment assessment) and not continuously during treatment (e.g. from the first to the second session) in all but one trial. That trial assessed PTSD exacerbations continuously throughout the treatment (i.e. session by session).<sup>33</sup> Since these operationalisations differ and incidences are likely to differ as a consequence, we also performed sensitivity analyses without the mentioned study. It is noteworthy that symptom exacerbations early in treatment are

quite common in psychological interventions such as exposure-based treatments but research suggests that symptom worsening at this stage is not associated with worse treatment outcomes.<sup>34</sup> In terms of adverse effects, we distinguished between AEs (i.e. involving aversive but non-lethal states such as increased severity of a comorbid mental disorder) and SAEs (i.e. involving potentially lethal states such as acute suicidality) and analysed these separately as recommended.<sup>13</sup> Furthermore, we extracted all information on reported (potential) causes of harms in an effort to quantitatively summarise the extent to which harms were deemed treatment-related versus treatment-unrelated. Our main focus was on harms during treatment (i.e. between the pre- and post-treatment assessments). However, some trials also reported on the incidence of harms between the pre-treatment and follow-up assessments, which we also summarised. In case harms were reported for several pre-treatment to follow-up timespans, we only included the data of the longest follow-up. We aimed to summarise the incidence of harms in three ways: (a) proportions as a percentage per group, (b) relative risks (RR) of harms between groups and (c) numbers needed to treat (NNTs), the latter being a metric that might be easier to interpret from a clinician's viewpoint.<sup>35</sup> The NNT denotes the number of patients that need to be treated in the experimental group (compared with the given comparator) to avoid one event of harm (e.g. one patient deteriorating).

### Statistical analysis

Analyses were conducted using the packages *meta* (version 4.16-2<sup>36</sup>) and *metafor* (version 3.0-2<sup>37</sup>) in R version 3.6.1<sup>38</sup> for Windows. In general, analyses were only conducted when the evidence base for a given comparison was sufficiently large (i.e.  $k \geq 4$ ).<sup>4</sup> Random-effects analyses were conducted since we expected large heterogeneity in outcomes.<sup>39</sup> We analysed incidence of harms separately per timespan (see above) and comparison group. Three different comparisons were distinguished: (a) psychological interventions versus passive control conditions, (b) psychological interventions versus active control conditions and (c) a given family of psychological interventions (e.g. TF-CBT) versus another family of interventions. TF-CBT was the only family of interventions with enough trials reporting on harms to warrant sub-analyses. Whenever possible (i.e.  $k \geq 4$ ), we performed sensitivity analyses involving: (a) only trials with face-to-face delivery of intervention (i.e. excluding technology-based interventions), (b) only trials with individual delivery of intervention (i.e. excluding group, couple or mixed formats) and (c) only trials with a definition targeting clinically significant deterioration (i.e. excluding liberal definitions of deterioration). This last-mentioned analysis was performed since some trials defined deterioration as any pre- to post-treatment symptom increase (i.e. an increase of  $\geq 1$  score<sup>40</sup>), which also includes non-clinically significant symptom increase. See the third column of Supplement 4 for the categorisation of definitions into clinically significant (i.e. conservative) or non-clinically significant (i.e. liberal) deterioration.

To summarise proportions, we conducted random-effects meta-analyses on Freeman–Tukey double arcsine transformed prevalence proportions with the inverse variance method.<sup>41</sup> For forest plots, we calculated Agresti–Coull 95% confidence intervals.<sup>42</sup> To calculate RR, the proportion of participants in the treatment group who experienced a given harm (e.g. deterioration) was divided by the proportion in the control group:  $RR = 1$  indicates equal risk,  $RR < 1$  indicates a lower risk in the treatment group compared with the control group and  $RR > 1$  indicates a higher risk than in the control group. We calculated 95% confidence intervals of RR around  $I^2$  using the non-central chi-squared-based approach.<sup>43</sup>

To estimate heterogeneity in reported incidences of harms we calculated  $Q$ -statistics and  $I^2$ -statistics. The latter indicates the

magnitude of heterogeneity in percentages, with higher values indicating higher heterogeneity. We estimated  $\tau^2$ -statistics using the restricted maximum likelihood method as a metric for the between-study variance.<sup>44</sup> To calculate NNT, 1 is divided by the risk difference (i.e. difference between group proportions). As recommended,<sup>45</sup> we performed the Egger's test of asymmetry to detect potential publication bias<sup>46</sup> only when the evidence base was sufficiently large (i.e.  $k \geq 10$ ). We used the trim-and-fill method and reported asymmetry-adjusted results when significant asymmetry was indicated.<sup>47</sup> When we detected outliers, we conducted outlier-adjusted analyses and checked for discrepancies in results. We defined outliers as extraordinarily high or low incidences of harms (i.e. observations at least 3.3 standard deviations above or below the pooled proportion<sup>48</sup>).

When the evidence base was sufficiently large (i.e.  $k \geq 10$ ), we performed moderator analyses on trial quality (see 'Risk of bias assessment' above), the total treatment length in minutes (i.e. number of sessions times session length), age (i.e. sample mean), the proportion of female participants, the proportion of participants on psychotropic medication during the trial and the proportion of participants with a current comorbid depressive illness as potential moderating variables in meta-regressions.

## Results

### Selection and characteristics of studies

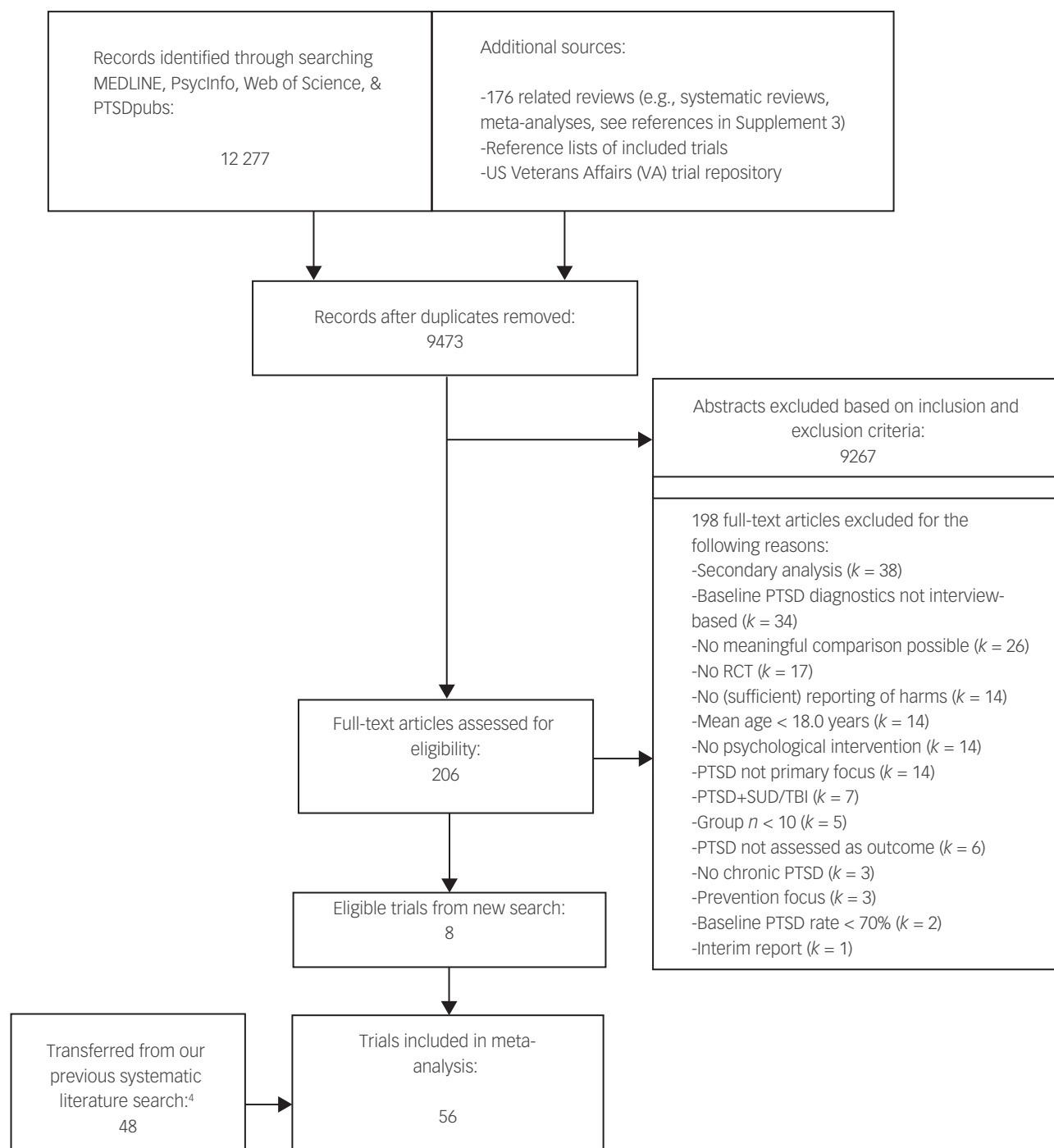
A total of 5711 hits remained after the deletion of duplicates. After the removal of 5526 hits that did not meet inclusion criteria, 185 hits remained for the thorough full-text screen. Eight new trials were identified and another 48 trials were referred from the aforementioned previous search.<sup>4</sup> The detailed study selection flow is shown in the PRISMA flowchart in Fig. 1. The vast majority of trials (94%) were conducted in the geographically Western world and involved TF-CBT (78%). About an even number of trials compared a psychological intervention with passive control conditions ( $k = 21$ ) or active control conditions ( $k = 20$ ). Most trials (67%) utilised the Clinician-Administered PTSD Scale<sup>49</sup> as their primary outcome measure and delivered their intervention in an individual format (81%) rather than a group, couple or mixed format. Across the 56 eligible trials, 4230 participants completed the post-treatment assessment. On average, participants were 40.78 years old at baseline (s.d. = 8.63) and a little more than half (58%) were females. Only 50% of the trials reported on concurrent intake of psychotropic medication. Across these trials, about half of participants (51%) were taking concurrent psychotropic medication during the trial (s.d. = 25%). The average total treatment duration was 882.76 min (s.d. = 359.81) with a mean of 11.51 treatment sessions (s.d. = 3.98). TF-CBT was the only family of interventions with a sufficiently large evidence base to warrant sub-analyses. See Supplement 4 for a comprehensive overview of characteristics of included trials and Supplement 5 for their references.

### Risk of bias

Quality of trials was moderate to high, with a mean quality sum score across trials of 6.56 (out of 8) (s.d. = 0.98). In none of the moderator analyses was trial quality found to be a significant predictor of harm incidences (see below). The quality ratings per trial are reported in Supplement 11.

### Pre- to post-treatment deterioration

Most trials that systematically assessed and reported on deterioration used the reliable change index of Jacobson & Truax.<sup>50</sup> In the trials comparing psychological interventions with passive control



**Fig. 1** PRISMA flowchart of study selection. PTSD, post-traumatic stress disorder; RCT, randomised controlled trial; SUD, substance use disorder; TBI, traumatic brain injury.

conditions, psychological interventions were associated with very low incidences of pre- to post-treatment deterioration, with less than 1% of participants deteriorating (Table 1). Incidences were substantially higher for passive control conditions, with about 11% of participants reporting deterioration between pre- and post-treatment assessment, resulting in a 79% decreased risk for psychological interventions as compared with passive control conditions (RR = 0.21, 95% CI 0.15–0.28). Although heterogeneity in outcomes was low for psychological interventions, it was high and significant for passive control conditions. About ten patients would need to be treated with a psychological intervention compared with passive control conditions

(i.e. no intervention provided) to avoid one patient suffering deterioration. Note, however, that the NNT should be interpreted in the light of generally low incidences of deterioration even when no treatment was offered. None of these trials reported on reasons for deterioration, making a distinction between treatment-related versus treatment-unrelated deterioration impossible. The sensitivity analyses on trials exclusively targeting clinically significant deterioration and trials exclusively delivering the intervention face to face as well as the sub-analyses on TF-CBT versus passive control conditions yielded largely the same pattern of results as the main analyses (see Supplement 6 for an expanded version of Table 1).

**Table 1** Incidence and relative risk of pre- to post-treatment deterioration and pre-treatment to follow-up deterioration

Comparison	k	Proportion <sup>a</sup>	$I^2$	95% CI	$I^2$	NNT	$k^b$	RR	95% CI	$I^2$
<b>Pre- to post-treatment</b>										
Psychological interventions v. PCC	8	0.83%	0.00	0.00–3.32	0.00	9.76	6	0.21	<b>0.15–0.28</b>	0.00
	8	11.08%	0.04	2.82–22.88	82.20***					
Psychological interventions v. ACC	9	0.38%	0.00	0.00–2.21	0.00	25.06	8	0.36	<b>0.14–0.92</b>	0.00
	9	4.37%	0.01	1.10–9.07	46.20					
TF-CBT v. other interventions	11	4.27%	0.01	1.10–8.81	60.20	294.12	8	1.05	0.59–1.86	0.00
	11	4.61%	0.01	1.68–8.51	40.60					
<b>Pre-treatment to follow-up</b>										
Psychological interventions v. ACC	8	2.77%	0.01	0.00–8.50	59.40*	19.69	6	0.61	0.14–2.69	31.30
	8	7.85%	0.02	1.91–16.43	68.70**					
TF-CBT v. ACC	5	1.09%	0.01	0.00–5.80	25.20	14.95	4	0.39	0.02–8.66	55.90
	5	7.78%	0.02	1.02–18.46	66.70*					
TF-CBT v. ACC, f-to-f data only	4	0.00%	0.00	0.00–2.49	0.00	9.45	4	0.13	<b>0.02–0.81</b>	0.00
	4	10.58%	0.02	1.94–23.49	55.40					
TF-CBT v. other interventions	5	0.74%	0.00	0.00–3.75	0.00	30.12	4	0.39	0.13–1.21	0.00
	5	4.06%	0.01	0.15–11.07	47.20					

ACC, active control conditions; f-to-f data only, only includes trials with face-to-face delivery of intervention (rather than mainly or fully internet/technology-based delivery);  $I^2$ , amount of unexplained variance (in %), including the magnitude of statistical significance of the corresponding  $Q$ -statistic as indicated by the asterisks;  $k$ , number of trials included in the analysis for the given comparison; NNT, number needed to treat (i.e. to avoid one case of deterioration when the given treatment is compared with the given control group/other intervention); other interventions, non-TF-CBT interventions; PCC, passive control conditions; TF-CBT, trauma-focused cognitive-behavioural therapy.

a. Proportion refers to the proportion of participants per group who deteriorated (i.e. worsening of post-traumatic stress disorder symptoms).

b. Trials with zero cases/deteriorators for both groups did not contribute to the RR calculation (i.e. absolute risk = 0.00 in both groups) and are therefore not part of  $k$ .

\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . Data in **bold** indicate that the 95% CI of the RR excludes the null (i.e. differential risk).

In the trials comparing psychological interventions with active control conditions, the risk of pre- to post-treatment deterioration was significantly lower for psychological interventions. More specifically, these had a 64% decreased risk of deterioration (RR = 0.36, 95% CI 0.14–0.92), with about 0.5% of participants receiving a psychological intervention deteriorating compared with about 5% in the active control conditions. Although heterogeneity in outcomes was low for psychological interventions, it was moderate (but non-significant) for active control conditions. None of these trials reported on reasons for deterioration. Results remained largely the same for the sensitivity analysis on trials exclusively targeting clinically significant deterioration and the sensitivity analysis on trials exclusively delivering interventions face to face (Supplement 6). The same pattern of results emerged for the sub-analyses on TF-CBT versus active control conditions (Supplement 6).

In the trials comparing TF-CBT with other psychological interventions, incidences of pre- to post-treatment deterioration were generally somewhat higher (4–5%) for both groups than in trials comparing TF-CBT with control conditions. No differential risk of pre- to post-treatment deterioration was observed between TF-CBT and other psychological interventions in the main analysis. Heterogeneity in incidences was moderate but insignificant for both TF-CBT and other interventions. The number of trials allowed for a formal publication bias check and moderator analyses (i.e.  $k \geq 10$ ). The Egger's test did not indicate significant asymmetry. Furthermore, no significant moderations were observed (Supplement 7). Resick et al<sup>33</sup> was identified as an outlier, but only for TF-CBT (Supplement 8) and the analyses were therefore re-run without this trial. Since Resick et al was the only trial delivering interventions in a non-individual format (i.e. group format), the outlier-adjusted analysis is identical to the sensitivity analysis on trials exclusively delivering interventions individually. Results by and large remained the same (Supplement 6). Heterogeneity, however, dropped from moderate to low for both groups. Resick et al was also the only trial reporting on reasons for deterioration, with most cases of deterioration deemed treatment-related for group TF-CBT (73%) and only about one-third of cases of deterioration deemed treatment-related for group present-centred therapy (33%). The sensitivity analysis on trials exclusively targeting clinically significant deterioration yielded very similar results, with no differential risks observed.

### Pre-treatment to follow-up deterioration

Fewer participants receiving psychological interventions (2.77%) reported deterioration from pre-treatment to follow-up assessment than participants in active control conditions (7.85%), yet the 39% decreased risk was not statistically significant (RR = 0.61, 95% CI 0.14–2.69) (Table 1). Heterogeneity in incidences were moderate in both groups. No trial reported on reasons for deterioration. Results remained by and large the same in the sensitivity analysis on trials exclusively targeting clinically significant deterioration. Similar results were found for the sub-analysis on TF-CBT versus active control conditions (Supplement 6). However, a significantly decreased relative risk was found for the sensitivity analysis on trials exclusively delivering TF-CBT in face-to-face format (i.e. excluding trials with mainly or fully internet/technology-based delivery) compared with active control conditions (RR = 0.13, 95% CI 0.02–0.81). Heterogeneity in incidences was low for TF-CBT and moderate for active control conditions. Although TF-CBT was associated with somewhat decreased incidences of pre-treatment to follow-up deterioration compared with other psychological interventions (about 1 v. 4%), the RR was not statistically significant (RR = 0.39, 95% CI 0.13–1.21). The only trial reporting on deterioration causes was again Resick et al,<sup>33</sup> who did not deem any cases of deterioration during follow-up to be related to treatment. That is, the treatment-related deterioration they observed from pre- to post-treatment appeared to be transient. Sensitivity analyses were precluded owing to lack of trials. On average, follow-up deterioration was assessed 5.42 months after treatment determination (s.d. = 3.37 months).

### Pre- to post-treatment adverse events

In trials comparing psychological interventions with passive control conditions, both psychological interventions and passive control conditions were associated with very low incidences of AEs during the timespan between the baseline and post-treatment assessment, with about 1–2% of participants reporting AEs (Table 2). Heterogeneity in incidences was high for psychological interventions as well as passive control conditions. The number of trials allowed for a formal publication bias check and moderator analyses (i.e.  $k \geq 10$ ). The Egger's test did not indicate significant asymmetry. For psychological interventions, younger age was

**Table 2** Incidence and relative risk of adverse events occurring pre- to post-treatment and pre-treatment to follow-up

Comparison	k	Proportion <sup>a</sup>	I <sup>2</sup>	95% CI	I <sup>2</sup>	NNT	k <sup>b</sup>	RR	95% CI	I <sup>2</sup>
<b>Pre- to post-treatment</b>										
Psychological interventions v. PCC	10	1.51%	0.02	0.00–6.31	76.80***	294.12			n.a. (k = 3)	
Psychological interventions v. PCC, outlier-adjusted	9	0.18%	0.00	0.00–2.31	38.40	769.23			n.a. (k = 2)	
Psychological interventions v. PCC, f-to-f data only	9	1.75%	0.02	0.00–7.21	79.10***	238.10			n.a. (k = 3)	
Psychological interventions v. PCC – f-to-f data only, outlier-adjusted	8	0.29%	0.01	0.00–3.08	46.10	769.23			n.a. (k = 2)	
TF-CBT v. PCC	8	2.14%	0.03	0.00–8.64	80.90***	185.19			n.a. (k = 3)	
TF-CBT v. PCC, outlier-adjusted	7	0.50%	0.01	0.00–4.49	53.40*	666.67			n.a. (k = 2)	
TF-CBT v. PCC, f-to-f data only	7	2.58%	0.03	0.00–10.31	83.30***	147.06			n.a. (k = 3)	
TF-CBT v. PCC – f-to-f data only, outlier-adjusted	6	0.74%	0.02	0.00–6.06	61.20*	588.24			n.a. (k = 2)	
Psychological interventions v. ACC	8	1.77%	0.01	0.00–5.54	42.30	-62.11			n.a. (k = 3)	
TF-CBT v. other interventions	10	5.52%	0.03	0.82–13.02	85.50***	25.25	6	0.67	0.29–1.56	68.50**
TF-CBT v. other interventions, individual data only	9	4.49%	0.03	0.22–12.15	86.10***	30.77	5	0.69	0.22–2.21	74.80**
<b>Pre-treatment to follow-up</b>										
Psychological interventions v. ACC	5	0.55%	0.01	0.00–4.86	52.00	-384.62			n.a. (k = 2)	
Psychological interventions v. ACC, f-to-f data only	4	0.00%	0.00	0.00–1.09	0.00	5000.00			n.a. (k = 1)	
TF-CBT v. other interventions	4	3.79%	0.04	0.00–16.60	86.10***	53.48			n.a. (k = 2)	

ACC, active control conditions; f-to-f data only, only includes trials with face-to-face delivery of intervention (rather than mainly or fully internet/technology-based delivery); I<sup>2</sup>, amount of unexplained variance (in %), including the magnitude of statistical significance of the corresponding Q-statistic as indicated by the asterisks; individual data only, only includes trials with individual delivery format of intervention (rather than group or mixed format); k, number of trials included in the analysis for the given comparison; n.a., not applicable; NNT, number needed to treat (i.e. to avoid one participant suffering at least one adverse event); other interventions, non-TF-CBT interventions; PCC, passive control conditions; TF-CBT, trauma-focused cognitive-behavioural therapy.

a. Proportion refers to the proportion of participants per group who experienced at least one adverse event in the given timespan (rather than total number of adverse events).

b. Trials with zero cases (i.e. no participants suffering at least one adverse event) for both groups did not contribute to the RR calculation (i.e. absolute risk = 0.00 in both groups).

\*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.001.

associated with significantly higher incidences of AEs (Supplement 9). All other moderator analyses were non-significant, including those on passive control conditions. The calculation of an aggregated RR was precluded because too few trials reported differential risks. In other words, most trials observed zero incidences of AEs for both groups and therefore did not contribute to the calculation of the aggregated RR. Foa et al<sup>51</sup> was identified as an outlier in both groups (Supplement 8) and we reran the analysis without this trial (Table 2 shows the outlier-adjusted results). This exclusion decreased the incidence of AEs to <0.5% in both groups and decreased heterogeneity substantially in both groups. Foa et al was also the only trial reporting on causes of AEs, with most cases deemed treatment-unrelated for both TF-CBT (61%) and passive control conditions (92%). A sensitivity analysis on trials exclusively delivering psychological interventions face to face as well as an outlier-adjusted repetition of this sensitivity analysis yielded largely similar results (Table 2). The sub-analyses on TF-CBT versus passive control conditions and the associated sensitivity analyses yielded a very similar picture of results, i.e. low incidences of AEs in both groups and a substantial reduction in heterogeneity in both groups when Foa et al was excluded.

In trials comparing psychological interventions with active control conditions, incidences of AEs during the treatment phase were also low and marginally higher for psychological interventions (about 1.8% versus about 0.2% respectively), with moderate heterogeneity in incidences for psychological interventions and low heterogeneity for active control conditions. None of these trials reported on reasons for AEs. Similar to the results on deterioration, trials on TF-CBT versus other psychological interventions yielded somewhat higher incidences of AEs compared with trials comparing TF-CBT with control conditions. TF-CBT was related to somewhat lower incidences of AEs during the treatment phase than other psychological interventions (5.5% v. 9.5% respectively), but this was not a statistically significant difference in risks (RR = 0.67, 95% CI 0.29–1.56). Heterogeneity was large and highly significant in both groups. Foa et al was one of the two trials reporting on causes of AEs, with about even percentages of treatment-related AEs for TF-CBT and present-centred therapy (39% v. 50% respectively). The other trial, Kearney et al,<sup>52</sup> did not deem any AEs to be treatment-related in either group (TF-CBT versus loving kindness meditation). Results remained by and large the same for the sensitivity analysis on trials exclusively delivering interventions individually, the only small difference being that incidences of AEs were marginally lower for both

individually delivered TF-CBT (4.5%) and individually delivered other psychological interventions (7.7%).

### Pre-treatment to follow-up adverse events

Between the baseline and (last) follow-up assessment, both the psychological interventions and the active control conditions yielded very low aggregated incidences of AEs (<1%), with low to moderate and insignificant heterogeneity (Table 2). Again, trials reporting a differential risk between the two groups were too few to allow for the calculation of RR. Rates and heterogeneity of AEs dropped to zero in the sensitivity analysis on trials exclusively involving face-to-face interventions. As in the analyses above, trials comparing TF-CBT with other psychological interventions found higher incidences of AEs than trials comparing TF-CBT with control conditions, although incidences were still low (about 4% for TF-CBT and 6% for other interventions). Heterogeneity was large and highly significant for both groups. The small number of trials precluded sensitivity analysis. Across trials investigating the incidence of AEs from pre-treatment to follow-up assessment, the average timespan was 5.61 months (s.d. = 3.81 months).

### Pre- to post-treatment serious adverse events

All trials comparing psychological interventions with passive control conditions were for TF-CBT. TF-CBT as well as passive control conditions yielded very low aggregated incidences of SAEs during the timespan between the baseline and post-treatment assessment, with about 1–2% of participants reporting SAEs in both groups (Table 3). Moderate and significant heterogeneity in incidences was found for passive control conditions and zero heterogeneity for TF-CBT. No significant differential risk was observed. Three trials reported on causes of SAEs.<sup>20,53,54</sup> The authors and/or internal review boards of all three trials deemed reported SAEs to be unrelated to treatment. The sensitivity analysis on trials exclusively delivering TF-CBT face to face yielded very similar results. Similarly, trials comparing psychological interventions with active control conditions yielded low aggregated incidences of SAEs for both groups (i.e. ≤1%). Heterogeneity was zero for both groups and too few trials found differential risks to allow for the calculation of RR. Also, the low number of trials precluded any sensitivity or sub-analyses. One trial<sup>55</sup> reported on causes of SAEs and deemed none of the SAEs to be related to treatment. Lastly, trials comparing TF-CBT with other interventions yielded low aggregated incidences of pre- to post-treatment SAEs for both groups (2–3%), with low and insignificant heterogeneity

for TF-CBT and large and highly significant heterogeneity for other psychological interventions. Although incidences were low for both groups, they were significantly different in magnitude. TF-CBT was associated with a 46% decreased risk of SAEs compared with other psychological interventions (RR = 0.54, 95% CI 0.31–0.95). The authors and/or internal review boards of five trials reported on causes of SAEs, of which three<sup>20,55,56</sup> deemed all SAEs to be unrelated to treatment and the remaining two<sup>57,58</sup> deemed most SAEs (80% and 67% respectively) to be unrelated to treatment. A sensitivity analysis on trials exclusively delivering interventions in individual format found very similar results. No significant moderations were observed in the main analysis or in the sensitivity analysis (Supplement 10). Trials on pre-treatment to follow-up SAEs were too few to warrant a meta-analytic review.

## Discussion

We aimed to quantitatively summarise the evidence base on the incidences and relative risks of deterioration, AEs and SAEs in psychotherapy trials for adult PTSD relative to control conditions and among different psychological interventions.

The first main finding concerns the lack of reporting of harms in this line of research. Out of a total of 157 potential RCTs, only 56 (i.e. 36%) reported on harms. Reporting on (potential) causes of harms was even more rare. Nonetheless, the current evidence base hints at the safety of psychological interventions for adult PTSD. In none of our analyses were psychological interventions associated with higher risk than control conditions. In some but not all analyses, psychological interventions were associated with significantly decreased risks of harms compared with control conditions. Generally, incidences of harms were low, with only a small minority (0–5% for most analyses) of treated patients experiencing a worsening of their PTSD (i.e. deterioration), AEs or SAEs throughout or shortly after treatment.

The current evidence base further suggests that TF-CBT is about as safe as other psychological interventions when it comes to deterioration and AEs and safer in terms of SAEs. These findings are essential given that many therapists assume that exposure techniques used in TF-CBT are more distressing than other interventions.<sup>59,60</sup> Based on the current evidence base, TF-CBT appears at least as safe as other psychological interventions, including those without a trauma focus.

We further found preliminary evidence for higher incidences of harms in group and technology-based treatments compared with

**Table 3** Incidence and relative risk of serious adverse events occurring pre- to post-treatment

Comparison	k	Proportion <sup>a</sup>	$\tau^2$	95% CI	$I^2$	NNT	$k^b$	RR	95% CI	$I^2$
Psychological interventions/TF-CBT v. PCC	8	1.20%	0.00	0.09–3.08	0.00	–5000	4	0.56	0.11–2.91	3.30
Psychological interventions/TF-CBT v. PCC, f-to-f data only	8	1.18%	0.01	0.00–4.51	51.60*					
	7	1.13%	0.00	0.00–3.02	0.00	357.14		n.a. (k = 3)		
	7	1.41%	0.01	0.00–5.27	57.90*					
Psychological interventions v. ACC	5	0.46%	0.00	0.00–2.68	0.00	185.19		n.a. (k = 2)		
	5	1.00%	0.00	0.00–3.63	0.00					
TF-CBT v. other interventions	12	1.74%	0.00	0.63–3.23	15.70	196.08	7	0.54	<b>0.31–0.95</b>	0.00
	12	2.25%	0.01	0.43–5.01	69.20***					
TF-CBT v. other interventions, individual data only	11	2.03%	0.00	0.84–3.58	8.40	172.41	7	0.54	<b>0.31–0.95</b>	0.00
	11	2.61%	0.01	0.57–5.64	68.80***					

ACC, active control conditions; f-to-f data only, only includes trials with face-to-face delivery of intervention (rather than mainly or fully internet/technology-based delivery);  $I^2$ , amount of unexplained variance (in %) including the magnitude of statistical significance of the corresponding Q-statistic as indicated by the asterisks; individual data only, only includes trials with individual delivery format of intervention (rather than group or mixed format); k, number of trials included in the analysis for the given comparison; n.a., not applicable; NNT, number needed to treat (i.e. to avoid one participant suffering at least one serious adverse event); other interventions, non-TF-CBT interventions; PCC, passive control conditions; TF-CBT, trauma-focused cognitive-behavioural therapy.

a. note that proportion refers to the proportion of participants per group who experienced at least one serious adverse event in the given timespan (rather than total number of serious adverse events).

b. note that trials with zero cases (i.e. participants suffering at least one serious adverse event) for both groups did not contribute to the RR calculation (i.e. absolute risk = 0.00 in both groups).

\* $P < 0.05$ , \*\*\* $P < 0.001$ . Data in **bold** indicate that the 95% CI of the RR excludes the null (i.e. differential risk).

individual face-to-face treatments. In various sensitivity analyses, the exclusion of trials with group or technology-based delivery resulted in significantly decreased incidences and a shift from no differential risks to differential risks. However, only few trials delivering psychological interventions in group or technology-based formats and reporting on harms have been published so far, which precludes firm conclusions. Future research needs to investigate whether technology-based delivery is associated with greater risk of harm relative to face-to-face treatments. In the light of the COVID-19 pandemic and a sharp increase in technology-based delivery of psychological interventions,<sup>61</sup> this is a pressing topic. Similarly, we found some evidence for group treatments being more harmful than individual treatments. The trial by Resick et al<sup>33</sup> was an outlier in the analysis on pre- to post-treatment deterioration, with an extraordinarily high proportion of patients in both groups (i.e. group cognitive processing therapy, defined as TF-CBT, and group present-centred therapy, defined as non-trauma-focused intervention) experiencing deterioration. About two in three instances of pre- to post-treatment deterioration were deemed treatment-related in the TF-CBT group, whereas only one in three instances were deemed treatment-related in the present-centred therapy group. However, only one incidence of deterioration during follow-up was reported in both conditions and both were treatment-unrelated. Resick et al concluded that induced harms did not appear to endure. It is noteworthy that the trial by Resick et al differs from the other trials in that deterioration was assessed continuously throughout treatment (i.e. from session to session) and not just between two timepoints (e.g. pre- and post-treatment assessments). The higher incidence of deterioration in Resick et al compared with the other trials may at least partly be explained by their unique operationalisation. The non-enduring deterioration found (i.e. zero deteriorators during follow-up) is in line with research suggesting that initial symptom exacerbation during the early stages of exposure-based treatments tends to be transient and to not be associated with treatment outcome.<sup>34</sup> Altogether, few trials with group-based treatments reported on harms and the study by Resick et al signifies the need for more research in this regard.

Lastly, we found some evidence for younger patients reporting more AEs during psychological treatment. However, this result just met the statistical significance level and no such moderation was found for the analyses on incidences of deterioration. More data are needed to draw firmer conclusions.

### Comparison with the literature

Meta-analytic reviews on the present topic are scarce. However, the meta-analysis by Jayawickreme et al<sup>11</sup> reported very similar results to ours. In their summary of four RCTs, Jayawickreme et al found zero incidence of deterioration for patients treated with prolonged exposure compared with 8.1% incidence for patients randomised to waiting list control conditions. In our much broader meta-analysis involving considerably more trials and various psychological interventions (prolonged exposure being one of them), we also found significantly lower incidences and risk of PTSD deterioration in patients receiving a psychological intervention (0.83%) compared with those in waiting list control conditions (11.08%). Comparisons can also be drawn with psychological treatment for other mental disorders. As mentioned above, in some studies with patients with depression or obsessive-compulsive disorder, as many as 53% and 93% of participants respectively reported experiencing one or more AEs event during treatment.<sup>16,62</sup> We found incidences of AEs during psychological treatment of 2–3% for trials comparing with control conditions and of 5–11% for trials comparing different psychological interventions. This striking difference in incidences

might be explained by the fact that both studies (on depression and on obsessive-compulsive disorder) assessed AEs retrospectively through an online survey, leading to various sources of potential bias (e.g. memory bias, self-selection bias in responders). The incidences in our meta-analysis were prospectively assessed. Similar results were found for prospective assessments of harms in the field of depression. Like Cuijpers et al in their meta-analysis of RCTs on psychological treatments for adult depression, we found incidences of pre- to post-treatment deterioration of around 5% (or lower)<sup>18</sup> and significantly decreased risk when compared with passive or active control conditions.<sup>19</sup> Lastly, a recent study reported that psychological interventions are also not associated with greater symptom exacerbation when compared with psychotropic medications in the treatment of adult PTSD.<sup>34</sup>

### Clinical implications

Psychological interventions are the first-line treatment recommendation in all influential treatment guidelines.<sup>3</sup> The presented results are reassuring by underscoring the safety in addition to the well-documented efficacy of psychological interventions for adult PTSD. Results might be particularly reassuring for trauma-focused interventions such as TF-CBT, since their safety profile has repeatedly been questioned. Generally, current results signify the need to complement the informed consent procedure by informing patients about potential harms during or following psychological treatment. Crucially, the present work further enhances transparency in this process by enabling comparisons of the incidence of harms between psychological interventions and control conditions. The current evidence base of RCTs suggests that postponing treatment is associated with higher incidences and risks of symptom exacerbation and adverse events than immediate treatment of PTSD with a psychological intervention. Accordingly, patients should also be informed about potential harms if they decide to postpone treatment. It is noteworthy that several longitudinal studies found that PTSD was associated with an increased risk of developing physical illnesses such as type 2 diabetes,<sup>63,64</sup> cardiovascular diseases<sup>65,66</sup> and obesity.<sup>67</sup> Another study found that the risk of developing type 2 diabetes was significantly decreased only in patients who experienced a clinically meaningful symptom reduction of PTSD by means of psychological and/or pharmacological treatment.<sup>68</sup> Therefore, clinicians are encouraged to inform patients not just about the risks and benefits of psychological interventions but also about the potential mental and physical health risks of postponing treatment. Such holistic information helps patients to make more informed and balanced treatment choices.

### Recommendations for future research



The validity of the current results is limited by the fact that most of the potential RCTs did not report on harms. Accordingly, research teams are encouraged to systematically and thoroughly assess and report on harms using validated and standardised measures including assessments of causes and correlates of harms. A more sophisticated understanding of why a minority of patients do suffer harms could help in harm prevention, drop-out prevention and other treatment obstacles, such as lacking motivation.

### Strengths and limitations

To the best of our knowledge, the present work presents the first meta-analysis that comprehensively summarised the incidence of harms in RCTs on the efficacy of psychological interventions for adult PTSD. Only one previous meta-analysis summarised incidences quantitatively, with the limitation that it mainly focused on prolonged exposure and consisted of only four trials.<sup>11</sup> With



our broad focus across several kinds of harm and across a broad range of psychological interventions, our findings add to the previous work in the area. Moreover, we followed gold standard guidelines in meta-analytic work<sup>27</sup> to maximise the validity of our work. The present work, however, naturally comes with its limitations. First, the number of trials reporting on harms was relatively low when compared with the general evidence base, with only about one in three trials reporting on harm(s). Consequently, power was undermined. However, we stuck to recommendations in meta-analytic research and only performed calculations when the number of trials was sufficiently large to give adequate power. Second, there was substantial variability in applied definitions and assessments of AEs between studies – a picture that is evident in psychotherapy research more generally.<sup>13,69</sup> Definitions of SAEs appeared to be somewhat more consistent, which is also in line with a recent systematic review on psychotherapy study protocols more generally.<sup>69</sup> Although definitions of deterioration also varied, research suggests that patient-based definitions of clinically significant change in symptoms overlap considerably with clinician-based definitions utilised in clinical trials.<sup>70</sup> That is, patients and clinicians appear to agree mostly on what meaningful improvement (and deterioration) constitutes. Nevertheless, it would have been desirable to conduct more homogeneous sub-analyses per definition of harms. However, the current evidence base is too slim to allow such sub-analyses with sufficient statistical power. More data are needed for more fine-grained approaches. The presented results are to be interpreted in the realm of varying definitions and cannot and should not be used to make risk predictions for single individuals. Rather, they give an overall estimate of incidences and risks of harms at the group level in RCTs and should be articulated as such (e.g. in the informed consent procedure).

**Thole H. Hoppen** , PhD, Institute of Psychology, University of Münster, Germany; **Anna S. Lindemann**, BSc, Institute of Psychology, University of Münster, Germany; **Nexhmedin Morina** , PhD, Institute of Psychology, University of Münster, Germany

**Correspondence:** Thole H. Hoppen. Email: [thoppen@uni-muenster.de](mailto:thoppen@uni-muenster.de)

First received 4 Mar 2022, final revision 8 Jul 2022, accepted 20 Jul 2022

## Supplementary material

Supplementary material is available online at <https://doi.org/10.1192/bjp.2022.111>.

## Data availability

Data availability is not applicable to this article as no new data were created or analysed in this study.

## Acknowledgements

The authors thank Svenja Heinrich, Johanna Seroka and Donna Ivo Petrova for their support during the systematic literature search and data extractions.

## Author contributions

T.H.H. and N.M. conceptualised the meta-analysis. T.H.H. and A.S.L. performed the literature search and data extractions. T.H.H. performed the statistical analyses and wrote the first draft of the manuscript as well as revised versions of the manuscript. All three authors contributed to and have approved the final manuscript.

## Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

## Declaration of interest

None.

## References

- Kessler RC, Aguilar-Gaxiola S, Alonso J, Benjet C, Bromet EJ, Cardoso G, et al. Trauma and PTSD in the WHO World Mental Health Surveys. *Eur J Psychotraumatol* 2017; **8**(5): 1353383.
- 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. *Clin Psychol Rev* 2014; **34**: 249–55.
- Martin A, Naunton M, Kosari S, Peterson G, Thomas J, Christenson JK. Treatment guidelines for PTSD: a systematic review. *J Clin Med* 2021; **10**(18): 4175.
- Morina N, Hoppen TH, Kip A. Study quality and efficacy of psychological interventions for posttraumatic stress disorder: a meta-analysis of randomized controlled trials. *Psychol Med* 2021; **51**(1): 1–11.
- Lewis C, Roberts NP, Andrew M, Starling E, Bisson JI. Psychological therapies for post-traumatic stress disorder in adults: systematic review and meta-analysis. *Eur J Psychotraumatol* 2020; **11**(1): 1729633.
- Schwartz D, Barkowski S, Strauss B, Knaevelsrud C, Rosendahl J. Efficacy of group psychotherapy for posttraumatic stress disorder: systematic review and meta-analysis of randomized controlled trials. *Psychother Res* 2019; **29**: 415–31.
- Coventry PA, Meader N, Melton H, Temple M, Dale H, Wright K, et al. Psychological and pharmacological interventions for posttraumatic stress disorder and comorbid mental health problems following complex traumatic events: systematic review and component network meta-analysis. *PLoS Med* 2020; **17**(8): e1003262.
- Brady K, Pearlstein T, Asnis GM, Baker D, Rothbaum B, Sikes CR, et al. Efficacy and safety of sertraline treatment of posttraumatic stress disorder: a randomized controlled trial. *JAMA* 2000; **283**: 1837–44.
- Marshall RD, Beebe KL, Oldham M, Zaninelli R. Efficacy and safety of paroxetine treatment for chronic PTSD: a fixed-dose, placebo-controlled study. *Am J Psychiatry* 2001; **158**: 1982–8.
- Mithoefer MC, Wagner MT, Mithoefer AT, Jerome L, Doblin R. The safety and efficacy of  $\pm$ 3,4-methylenedioxymethamphetamine-assisted psychotherapy in subjects with chronic, treatment-resistant posttraumatic stress disorder: the first randomized controlled pilot study. *J Psychopharmacol* 2011; **25**: 439–52.
- Jayawickreme N, Cahill SP, Riggs DS, Rauch SAM, Resick PA, Rothbaum BO, et al. Primum non nocere (first do no harm): symptom worsening and improvement in female assault victims after prolonged exposure for PTSD. *Depress Anxiety* 2014; **31**: 412–9.
- Lilienfeld SO. Psychological treatments that cause harm. *Perspect Psychol Sci* 2007; **2**: 53–70.
- Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. *Clin Psychol Psychother* 2013; **20**: 286–96.
- McKay D, Jensen-Doss A. Harmful treatments in psychotherapy. *Clin Psychol* 2021; **28**: 2–4.
- Nutt DJ, Sharpe M. Uncritical positive regard? Issues in the efficacy and safety of psychotherapy. *J Psychopharmacol* 2008; **22**: 3–6.
- Moritz S, Fieker M, Hottenrott B, Seeralan T, Cludius B, Kolbeck K, et al. No pain, no gain? Adverse effects of psychotherapy in obsessive-compulsive disorder and its relationship to treatment gains. *J Obsessive Compuls Relat Disord* 2015; **5**: 61–6.
- Crawford MJ, Thana L, Farquharson L, Palmer L, Hancock E, Bassett P, et al. Patient experience of negative effects of psychological treatment: results of a national survey. *Br J Psychiatry* 2016; **208**: 260–5.
- Cuijpers P, Karyotaki E, Ciharova M, Miguel C, Noma H, Furukawa TA. The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: a meta-analysis. *Acta Psychiatr Scand* 2021; **144**: 288–99.
- Cuijpers P, Reijnders M, Karyotaki E, Wit Ld, Ebert DD. Negative effects of psychotherapies for adult depression: a meta-analysis of deterioration rates. *J Affect Disord* 2018; **239**: 138–45.
- van den Berg DPG, Bont Pd, van der Vleugel BM, Roos C, Jongh A, van Minnen A, et al. Prolonged exposure vs eye movement desensitization and reprocessing vs waiting list for posttraumatic stress disorder in patients with a psychotic disorder: a randomized clinical trial. *JAMA Psychiatry* 2015; **72**: 259–67.
- Varker T, Kartal D, Watson L, Frejajah I, O'Donnell M, Forbes D, et al. Defining response and nonresponse to posttraumatic stress disorder treatments: a systematic review. *Clin Psychol* 2020; **27**(4): e12355.

- 22 Fonzo GA, Federcheno V, Lara A. Predicting and managing treatment non-response in posttraumatic stress disorder. *Curr Treat Options Psychiatry* 2020; **7**: 70–87.
- 23 ter Heide FJJ, Mooren TM, van de Schoot R, de Jongh A, Kleber RJ. Eye movement desensitisation and reprocessing therapy v. stabilisation as usual for refugees: randomised controlled trial. *Br J Psychiatry* 2016; **209**: 311–8.
- 24 Litz BT, Rusowicz-Orazem L, Doros G, Grunthal B, Gray M, Nash W, et al. Adaptive disclosure, a combat-specific PTSD treatment, versus cognitive-processing therapy, in deployed marines and sailors: a randomized controlled non-inferiority trial. *Psychiatry Res* 2021; **297**: 113761.
- 25 Cusack K, Jonas DE, Forneris CA, Wines C, Sonis J, Middleton JC, et al. Psychological treatments for adults with posttraumatic stress disorder: a systematic review and meta-analysis. *Clin Psychol Rev* 2016; **43**: 128–41.
- 26 Forman-Hoffman V, Middleton JC, Feltner C, Gaynes BN, Weber RP, Bann C, et al. *Psychological and Pharmacological Treatments for Adults with Posttraumatic Stress Disorder: A Systematic Review Update (Comparative Effectiveness Review No. 207)*. Agency for Healthcare Research and Quality, 2018.
- 27 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264–9.
- 28 Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley-Blackwell, 2008.
- 29 Chambless DL, Hollon SD. Defining empirically supported therapies. *J Consult Clin Psychol* 1998; **66**: 7–18.
- 30 Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychol Med* 2010; **40**: 211–23.
- 31 Gersons BP, Carlier IV, Lamberts RD, van der Kolk BA. Randomized clinical trial of brief eclectic psychotherapy for police officers with posttraumatic stress disorder. *J Trauma Stress* 2000; **13**: 333–47.
- 32 Cottraux J, Note I, Yao SN, Mey-Guillard C, Bonasse F, Djamoossian D, et al. Randomized controlled comparison of cognitive behavior therapy with Rogerian supportive therapy in chronic post-traumatic stress disorder: a 2-year follow-up. *Psychother Psychosom* 2008; **77**: 101–10.
- 33 Resick PA, Wachen JS, Mintz J, Young-McCaughan S, Roache JD, Borah AM, et al. A randomized clinical trial of group cognitive processing therapy compared with group present-centered therapy for PTSD among active duty military personnel. *J Consult Clin Psychol* 2015; **83**: 1058–68.
- 34 Walker RSW, Marks EH, Jaeger J, Duax JM, Feeny NC, Zoellner LA. Imaginal exposure exacerbation revisited: deconstructing patient characteristics associated with worse reactions to the initiation of imaginal exposure in PTSD. *Behav Res Ther* 2020; **135**: 103747.
- 35 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995; **310**: 452–4.
- 36 Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019; **22**: 153–60.
- 37 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Soft* 2010; **36**(3): 1–48.
- 38 R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation, 2015. Available from: <http://www.R-project.org/>.
- 39 Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: d549.
- 40 Cloitre M, Koenen KC, Cohen LR, Han H. Skills training in affective and interpersonal regulation followed by exposure: a phase-based treatment for PTSD related to childhood abuse. *J Consult Clin Psychol* 2002; **70**: 1067–74.
- 41 Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. *J Epidemiol Community Health* 2013; **67**: 974–8.
- 42 Agresti A, Coull BA. Approximate is better than “exact” for interval estimation of binomial proportions. *Am Stat* 1998; **52**: 119–26.
- 43 Ioannidis JPA, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ* 2007; **335**: 914–6.
- 44 Schwarzer G, Carpenter JR, Rücker G. *Meta-analysis with R*. Springer, 2015.
- 45 Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011; **343**: d4002.
- 46 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.
- 47 Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; **56**: 455–63.
- 48 Tabachnick BG, Fidell LS. *Using Multivariate Statistics* (6th edn). Pearson, 2013.
- 49 Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, et al. The development of a clinician-administered PTSD scale. *J Trauma Stress* 1995; **8**: 75–90.
- 50 Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; **59**: 12–9.
- 51 Foa EB, McLean CP, Zang Y, Rosenfield D, Yadin E, Yavris JS, et al. Effect of prolonged exposure therapy delivered over 2 weeks vs 8 weeks vs present-centered therapy on PTSD symptom severity in military personnel: a randomized clinical trial. *JAMA* 2018; **319**: 354–64.
- 52 Kearney DJ, Malte CA, Storms M, Simpson TL. Loving-kindness meditation vs cognitive processing therapy for posttraumatic stress disorder among veterans: a randomized clinical trial. *JAMA Netw Open* 2021; **4**(4): e216604.
- 53 Bryant RA, Kenny L, Rawson N, Cahill C, Joscelyne A, Garber B, et al. Efficacy of exposure-based cognitive behaviour therapy for post-traumatic stress disorder in emergency service personnel: a randomised clinical trial. *Psychol Med* 2019; **49**: 1565–73.
- 54 Monson CM, Fredman SJ, Macdonald A, Pukay-Martin ND, Resick PA, Schnurr PP. Effect of cognitive-behavioral couple therapy for PTSD: a randomized controlled trial. *JAMA* 2012; **308**: 700–9.
- 55 Nidich S, Mills PJ, Rainforth M, Heppner P, Schneider RH, Rosenthal NE, et al. Non-trauma-focused meditation versus exposure therapy in veterans with post-traumatic stress disorder: a randomised controlled trial. *Lancet Psychiatry* 2018; **5**: 975–86.
- 56 Chard KM. *A Comparison of Cognitive Processing Therapy (CPT) Versus Present Centered Therapy (PCT) for Veterans*, NCT00607815 (last updated 26 June 2018). clinicaltrials.gov, 2018 (<https://clinicaltrials.gov/ct2/show/study/NCT00607815>).
- 57 Schnurr PP, Friedman MJ, Engel CC, Foa EB, Shea MT, Chow BK, et al. Cognitive behavioral therapy for posttraumatic stress disorder in women: a randomized controlled trial. *JAMA* 2007; **297**: 820–30.
- 58 Surís A, Link-Malcolm J, Chard K, Ahn C, North C. A randomized clinical trial of cognitive processing therapy for veterans with PTSD related to military sexual trauma. *J Trauma Stress* 2013; **26**: 28–37.
- 59 Markowitz JC, Petkova E, Neria Y, van Meter PE, Zhao Y, Hembree E, et al. Is exposure necessary? A randomized clinical trial of interpersonal psychotherapy for PTSD. *Am J Psychiatry* 2015; **172**: 430–40.
- 60 van Minnen A, Hendriks L, Olff M. When do trauma experts choose exposure therapy for PTSD patients? A controlled study of therapist and patient factors. *Behav Res Ther* 2010; **48**: 312–20.
- 61 Inchausti F, MacBeth A, Hasson-Ohayon I, Dimaggio G. Telepsychotherapy in the age of COVID-19: a commentary. *J Psychother Integr* 2020; **30**: 394–405.
- 62 Moritz S, Nestoriuc Y, Rief W, Klein JP, Jelinek L, Peth J. It can't hurt, right? Adverse effects of psychotherapy in patients with depression. *Eur Arch Psychiatry Clin Neurosci* 2019; **269**: 577–86.
- 63 Roberts AL, Agnew-Blais JC, Spiegelman D, Kubzansky LD, Mason SM, Galea S, et al. Posttraumatic stress disorder and incidence of type 2 diabetes mellitus in a sample of women: a 22-year longitudinal study. *JAMA Psychiatry* 2015; **72**: 203–10.
- 64 Agyemang C, Goosen S, Anujoo K, Ogedegbe G. Relationship between post-traumatic stress disorder and diabetes among 105 180 asylum seekers in the Netherlands. *Eur J Public Health* 2012; **22**: 658–62.
- 65 Boscarino JA. Posttraumatic stress disorder and physical illness: results from clinical and epidemiologic studies. *Ann N Y Acad Sci* 2004; **1032**: 141–53.
- 66 Burg MM, Brandt C, Buta E, Schwartz J, Bathulapalli H, Dziura J, et al. Risk for incident hypertension associated with PTSD in military veterans, and the effect of PTSD treatment. *Psychosom Med* 2017; **79**(2): 181–8.
- 67 Scherrer JF, Salas J, Lustman PJ, van den Berk-Clark C, Schnurr PP, Tuerk P, et al. The role of obesity in the association between posttraumatic stress disorder and incident diabetes. *JAMA Psychiatry* 2018; **75**: 1189–98.
- 68 Scherrer JF, Salas J, Norman SB, Schnurr PP, Chard KM, Tuerk P, et al. Association between clinically meaningful posttraumatic stress disorder improvement and risk of type 2 diabetes. *JAMA Psychiatry* 2019; **76**: 1159–66.
- 69 Klatte R, Strauss B, Flückiger C, Färber F, Rosendahl J. Defining and assessing adverse events and harmful effects in psychotherapy study protocols: a systematic review. *Psychotherapy* [Epub ahead of print] 20 Jan 2022. Available from: <https://10.1037/pst0000359>.
- 70 Matthews SR, Elizabeth M, Roberts LN, Nixon RDV. Client versus clinicians' standards of clinically meaningful change and the effects of treatment expectations on therapeutic outcomes in individuals with posttraumatic stress disorder. *Behav Ther* 2022; **53**: 560–70.

