TECHNICAL COMMENTARY

NeuRA
Discover. Conquer. Cure.
POST-TRAUMATIC STRESS DISORDER LIBRARY

Diagnosis and screening

Introduction

For a person to be diagnosed with PTSD, at least one stressor is required. Stressors as determined by the DSM-5 include being exposed to threatened death, actual or threatened serious injury, or actual or threatened sexual violence. Examples are direct exposure, witnessing the trauma, or learning that a relative or close friend was exposed to a trauma. Stressors can be encountered in the course of professional duties.

At least one *"intrusion"* symptom is required for a diagnosis of PTSD. These symptoms include unwanted and upsetting memories, nightmares, flashbacks, and emotional distress and/or physical reactivity after exposure to reminders. At least one *"avoidance"* symptom is required. These include avoidance of trauma-related thoughts or feelings and/or avoidance of trauma-related external reminders. At least two *"negative alterations in cognitions and mood"* are required. These include negative thoughts or feelings that began or worsened after the trauma, an inability to recall key features of the trauma, overly negative thoughts and assumptions about oneself or the world, exaggerated blame of self or others for causing the trauma, negative affect, decreased interest in activities, feeling isolated, and difficulty experiencing positive affect. Finally, there needs to be at least two *"hyperarousal"* symptoms, such as irritability or aggression, risky or destructive behavior, hypervigilance, heightened startle reaction, difficulty concentrating, and difficulty sleeping.

Symptoms must not be due to medication, substance use, or other illness. They must last for more than one month and cause significant distress or problems to the individual's daily functioning. While most individuals develop symptoms within three months of the trauma, some symptoms can appear later and persist for months and sometimes years.

The latest World Health Organization's International Classification of Diseases (ICD-11) also includes complex PTSD, which is conceptualised as the core symptoms of PTSD plus disturbances in self organisation, affect dysregulation, negative self-concept, and disturbances in relationships. These disturbances are proposed to be associated with sustained, repeated, or multiple forms of traumatic exposure, reflecting loss of emotional, psychological and social resources under conditions of prolonged adversity.

A variety of tools have been developed to screen for, or diagnose, PTSD. The gold standards for diagnosis are the Clinician-Administered PTSD Scale (CAPS) and the Structured Clinical Interview for DSM-V (SCID-5), PTSD module. There are also a wide range of self-report PTSD measures, including the Primary Care PTSD Screen (PC-PTSD) and the PTSD Checklist (PCL), which are mostly used to monitor PTSD symptom severity, but can also be used for screening and diagnosing PTSD in people who have been exposed to trauma.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2010 that report results for people with a diagnosis of PTSD. Reviews were identified by searching the databases MEDLINE, EMBASE, and PsycINFO. Hand searching reference lists of identified reviews was also conducted. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[1]. Reviews with less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large, there is a dose dependent response or if results are reasonably consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found six systematic reviews that met our inclusion criteria[3-8].

- Moderate to high quality evidence finds a small increase in the severity of PTSD symptoms in people exposed to DSM-5 nominated stressors of actual or threatened death or serious injury or of threat to the physical integrity of self or others than in people exposed to other stressors, such as divorce, financial stress, or minor car accidents.
- Moderate to high quality evidence suggests around 24.5% of people diagnosed with PTSD have a delayed onset (>6 months post trauma), with most of these people experiencing earlier subclinical symptoms. Delayed-onset PTSD is highest in professional groups and in those who experienced combat trauma (prevalence both ~40%).
- Moderate to high quality evidence suggests reasonable sensitivity and good specificity of the Primary Care PTSD Screen (PC-PTSD) and the PTSD Checklist (PCL) for predicting a diagnosis of PTSD.
- Moderate quality evidence finds good diagnostic validity and internal consistency, and reasonable test-retest and external (convergent) validity of the PCL. Authors suggest findings support the construction of the new PCL-5 as a population-nonspecific instrument.
- Moderate to high quality evidence finds the average T score on the Trauma Symptom Checklist for Children is around 50 in youth exposed to traumatic events, which is 15 points less than the clinical cut-off for PTSD on this scale. Factors associated with increased scores include international (vs. U.S.) samples, sexual abuse (vs. neglect, community violence, or complex trauma), female sex, and older age in sexual abuse samples.
- Moderate to low quality evidence finds machine learning techniques (mostly Support Vector Machine learning) using neuroimaging, neuropsychological, or audio data can reasonably predict PTSD in people previously diagnosed using traditional means (mostly CAPS or PCL).

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Bressler R, Erford BT, Dean S

A systematic review of the Posttraumatic Stress Disorder Checklist (PCL)

Journal of Counseling and Development 2018; 96: 167-86

View review abstract online

Comparison	**Psychometric measures of the PTSD Checklist (PCL).**
Summary of evidence	**Moderate quality evidence (large sample, unable to assess consistency or precision, direct) finds good diagnostic validity and internal consistency, and reasonable test-retest and external (convergent) validity of the PCL. Authors suggest findings support the construction of the new PCL-5 as a population-nonspecific instrument.**
PCL-Specific (PCL-S)	
Mean scores in nonclinical populations Total (N = 13,756): 27.91, SD 12.46 Males (N = 863): 26.03, SD 12.32 Females (N = 1,800): 29.28, SD 13.42 Good diagnostic validity Cut off score 32 (N = 3,588): sensitivity = 0.91, specificity = 0.77 Cut off score 40 (N = 2,391): sensitivity = 0.84, specificity = 0.38 Cut off score 45 (N = 2,413): sensitivity = 0.73, specificity = 0.88 Cut off score 50 (N = 2,249): sensitivity = 0.59, specificity = 0.93 Good internal consistency (measuring construct accurately) Total (N = 22,179): α = 0.94, 95%CI 0.93 to 0.95 Intrusion/reexperiencing subscale (N = 889): α = 0.81, 95%CI 0.75 to 0.86 Avoidance subscale (N = 889): α = 0.80, 95%CI 0.73 to 0.86 Hyperarousal subscale (N = 683): α = 0.77, 95%CI 0.70 to 0.85 Dysphoria/numbing subscale (N not reported): α = 0.73, 95%CI 0.59 to 0.87 Moderate test–retest reliability ≤1 month (N = 35): *r* = 0.87, 95%CI 0.54 to 1.0 1–6 months (N = 448): *r* = 0.78, 95%CI 0.69 to 0.87 7–12 months (N = 1,099): *r* = 0.64, 95%CI 0.58 to 0.70 >12 months (N = 129): *r* = 0.77, 95%CI 0.60 to 0.94	

Moderate external (convergent) validity (correlations with other, related scales)
Acute Stress Disorder Severity Scale (N = 334): *r* = 0.54, 95%CI 0.43 to 0.65
Clinician-Administered PTSD Severity Scale (N = 164): *r* = 0.65, 95%CI 0.50 to 0.80
Davidson Trauma Scale (N = 164): *r* = 0.74, 95%CI 0.59 to 0.89
Posttraumatic Stress Diagnostic Scale (N = 164): *r* = 0.78, 95%CI 0.63 to 0.93
Posttraumatic Growth Inventory (N = 590): *r* = 0.37, 95%CI 0.29 to 0.45
Posttraumatic Cognitions Inventory (N = 553) total: *r* = 0.59, 95%CI 0.51 to 0.67
Posttraumatic Cognitions Inventory self: *r* = 0.56, 95%CI 0.48 to 0.64
Posttraumatic Cognitions Inventory world: *r* = 0.49, 95%CI 0.41 to 0.57
Posttraumatic Cognitions Inventory self-blame: *r* = 0.37, 95%CI 0.29 to 0.45
PCL-S subscale intercorrelations
Reexperiencing and avoidance subscales (N = 904): *r* = 0.78, 95%CI 0.72 to 0.84
Reexperiencing and hyperarousal subscales: *r* = 0.77, 95%CI 0.71 to 0.83
Avoidance and hyperarousal subscales: *r* = 0.79, 95%CI 0.73 to 0.85
PCL-S factor structure
Best fit was the four-factor model of reexperiencing, avoidance, hyperarousal, and dysphoria/numbing;
Comparative fit index (CFI, N = 236) = 0.92, root-mean-square error of approximation (RMSEA) = 0.058, standardised root mean square residual (SRMSR) = 0.053
Next best fit was the three-factor model of reexperiencing, avoidance, and hyperarousal;
CFI = 0.86, RMSEA = 0.076, SRMSR = 0.062
PCL-Military (PCL-M)
Mean scores in nonclinical populations
Total (N = 12,685): 26.94, SD 11.81
Good internal consistency
Total (N = 35,912): α = 0.95, 95%CI 0.94 to 0.96
Intrusion/reexperiencing subscale (N = 4,534): α = 0.84, 95%CI 0.81 to 0.86
Avoidance subscale (N = 4,534): α = 0.67, 95%CI 0.64 to 0.67
Hyperarousal subscale (N = 4,207): α = 0.70, 95%CI 0.67 to 0.73
Dysphoria/numbing subscale (N = 4,207): α = 0.85, 95%CI 0.82 to 0.88
Moderate test–retest reliability
1–6 months (N = 1,795): *r* = 0.58, 95%CI 0.53 to 0.63
7–12 months (N = 926): *r* = 0.68, 95%CI 0.62 to 0.74

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Moderate external (convergent) validity
Acute Stress Disorder Severity Scale (N = 334): *r* = 0.54, 95%CI 0.43 to 0.65
Clinician-Administered PTSD Severity Scale (N = 514): *r* = 0.65, 95%CI 0.57 to 0.74
Post-Deployment Readjustment Inventory (N = 215): *r* = 0.90, 95%CI 0.75 to 1.00
Posttraumatic Cognitions Inventory (N = 45) total: *r* = 0.61, 95%CI 0.32 to 0.90
PCL-S subscale intercorrelations
Reexperiencing and total (N = 1,080): *r* = 0.98, 95%CI 0.82 to 0.94
Avoidance and total (N = 1,080): *r* = 0.90, 95%CI 0.84 to 0.96
Hyperarousal and total (N = 753): *r* = 0.89, 95%CI 0.82 to 0.96
Dysphoria/numbing and total (N = 452): *r* = 0.89, 95%CI 0.80 to 0.98
Reexperiencing and avoidance (N = 5,208): *r* = 0.62, 95%CI 0.59 to 0.65
Reexperiencing and hyperarousal (N = 5,208): *r* = 0.56, 95%CI 0.53 to 0.59
Avoidance and hyperarousal (N = 4,881): *r* = 0.49, 95%CI 0.46 to 0.52
Dysphoria/numbing and reexperiencing (N = 4,506): *r* = 0.57, 95%CI 0.54 to 0.60
Dysphoria/numbing and avoidance (N = 4,506): *r* = 0.49, 95%CI 0.46 to 0.52
Dysphoria/numbing and hyperarousal (N = 4,506): *r* = 0.59, 95%CI 0.56 to 0.62
PCL-Civilian (PCL-C)
Mean scores in nonclinical populations
Total (N = 33,539): 25.28, SD 10.08
Males (N = 1,419): 24.95, SD 9.93
Females (N = 1,097): 27.77, SD 12.01
Good internal consistency
Total (N = 77,084): α = 0.94, 95%CI 0.93 to 0.95
Intrusion/reexperiencing subscale (N = 6,870): α = 0.86, 95%CI 0.84 to 0.89
Avoidance subscale (N = 6,870): α = 0.81, 95%CI 0.79 to 0.84
Hyperarousal subscale (N = 6,870): α = 0.85, 95%CI 0.82 to 0.87
Dysphoria/numbing subscale (N = 6,870): α = 0.82, 95%CI 0.77 to 0.86
Moderate test–retest reliability
≤1 month (N = 945): *r* = 0.79, 95%CI 0.73 to 0.85
1–6 months (N = 1,302): *r* = 0.51, 95%CI 0.46 to 0.56
7–12 months (N = 2,511): *r* = 0.52, 95%CI 0.48 to 0.56
>12 months (N = 666): *r* = 0.41, 95%CI 0.33 to 0.49
Moderate external (convergent) validity

Clinician-Administered PTSD Severity Scale (N = 4,023): *r* = 0.76, 95%CI 0.73 to 0.79 Trauma Symptom Checklist-40 (N = 417): *r* = 0.61, 95%CI 0.51 to 0.71 Posttraumatic Growth Inventory (N = 327): *r* = 0.39, 95%CI 0.28 to 0.50 Civilian Mississippi Posttraumatic Stress Disorder Scale (N = 978): *r* = 0.74, 95%CI 0.63 to 0.75 Impact Event Scale (N = 630): *r* = 0.77, 95%CI 0.69 to 0.85 <u>PCL-S subscale intercorrelations</u> Reexperiencing and total (N = 5,375): *r* = 0.89, 95%CI 0.86 to 0.92 Avoidance and total (N = 5,375): *r* = 0.82, 95%CI 0.79 to 0.85 Hyperarousal and total (N = 5,375): *r* = 0.90, 95%CI 0.87 to 0.93 Dysphoria/numbing and total (N = 4,076): *r* = 0.89, 95%CI 0.86 to 0.92 Reexperiencing and avoidance (N = 22,534): *r* = 0.84, 95%CI 0.83 to 0.85 Reexperiencing and hyperarousal (N = 22,534): *r* = 0.78, 95%CI 0.77 to 0.79 Avoidance and hyperarousal (N = 22,534): *r* = 0.76, 95%CI 0.75 to 0.77 Dysphoria/numbing and reexperiencing (N = 20,344): *r* = 0.79, 95%CI 0.78 to 0.80 Dysphoria/numbing and avoidance (N = 20,344): *r* = 0.79, 95%CI 0.78 to 0.80 Dysphoria/numbing and hyperarousal (N = 20,344): *r* = 0.88, 95%CI 0.87 to 0.89

Consistency‡	Authors report data are inconsistent.
Precision§	Precise
Directness‖	Direct

Larsen SE, Pacella ML

Comparing the effect of DSM-congruent traumas vs. DSM-incongruent stressors on PTSD symptoms: A meta-analytic review

Journal of Anxiety Disorders 2016; 38: 37-46

View review abstract online

Comparison	**Severity of PTSD symptoms after exposure to DSM-congruent events vs. DSM-incongruent events.** **DSM-congruent events were those involving actual or threatened death or serious injury, or a threat to the physical integrity of self or others. DSM-incongruent events were those that could potentially cause PTSD but were judged to be less traumatic (minor motor vehicle accidents, divorce, illness,**

	financial problems, etc).
Summary of evidence	**Moderate to high quality evidence (large sample, inconsistent, precise, direct) finds a small increase in the severity of PTSD symptoms in people exposed to actual or threatened death or serious injury, or a threat to the physical integrity of self or others compared to people exposed to other stressors.**
Severity of PTSD symptoms	
A small, significant effect showed more severe PTSD symptoms with DSM-congruent traumas; 22 studies, N = 9,964, *g* = 0.18, 95%CI 0.027 to 0.332, *p* = 0.021 Moderator analyses revealed self-report measures of trauma showed the greatest association with symptom severity.	
Consistency	Authors report data are inconsistent.
Precision	Precise
Directness	Direct

Martinez W, Polo AJ, Zelic KJ

Symptom variation on the trauma symptom checklist for children: a within-scale meta-analytic review

Journal of Traumatic Stress 2014; 27: 655-63

View review abstract online

Comparison	**Assessment of the Trauma Symptom Checklist for Children (mean age = 12.5 years).**
Summary of evidence	**Moderate to high quality evidence (large sample, appears precise, direct, inconsistent) finds the average T score on the Trauma Symptom Checklist for Children is around 50 in youth exposed to traumatic events, which is 15 points less than the clinical cutoff for PTSD on this scale. Factors associated with increased scores include international (vs. U.S.) samples, sexual abuse (vs. neglect community violence or complex trauma), female sex, and older age in sexual abuse samples.**
Trauma Symptom Checklist for Children	

74 studies, N = 14,867, 54 items, overall clinical cutoff = 65

Posttraumatic stress: 48 studies, T score mean = 50.91, 95%CI 50.03 to 51.78

Anger: 39 studies, T score mean = 49.44, 95%CI 48.24 to 50.65

Anxiety: 42 studies, T score mean = 51.25, 95%CI 50.40 to 52.11

Depression: 38 studies, T score mean = 50.73, 95%CI 49.50 to 51.95

Dissociation: 39 studies, T score mean = 50.80, 95%CI 49.80 to 51.79

I^2 values for each subscale ranged from 90.7% to 94.1%

Subgroup analyses showed youth in international samples reported greater posttraumatic stress than youth in U.S. samples for anxiety and depressive symptoms, but not for anger or dissociation.

Sexual abuse was associated with higher symptoms of posttraumatic stress, depression, and anxiety than child abuse/neglect, community violence, or complex trauma.

Females reported higher levels of posttraumatic stress, anxiety, depression, and dissociation than males. Females exposed to sexual abuse reported higher levels of anger, and females exposed to community violence reported higher levels of depression.

Older age was associated with higher posttraumatic stress, anxiety, depression, and dissociation in youth exposed to sexual abuse. Among youth exposed to community violence, older age was associated with increased dissociation only.

Ethnic minority status was associated with lower depression ratings.

Consistency in results	Inconsistent
Precision in results	Appears precise
Directness of results	Direct

Ramos-Lima LF, Waikamp V, Antonelli-Salgado T, Passos IC, Freitas LHM

The use of machine learning techniques in trauma-related disorders: a systematic review

Journal of Psychiatric Research 2020; 121: 159-72

View review abstract online

Comparison	**Assessment of the machine learning techniques for screening and diagnosis of PTSD.**
Summary of evidence	**Moderate to low quality evidence (mostly small samples, direct, unable to assess precision or consistency) finds machine learning techniques (mostly Support Vector Machine learning) using neuroimaging, neuropsychological, or audio data can**

	reasonably predict PTSD in people previously diagnosed using traditional means (mostly CAPS or PCL).

Machine learning techniques

Neuroimaging

1 study (N = 40) assessed fMRI scans in motor vehicle accident victims who had PTSD (CAPS designated) vs. healthy controls and found that Support Vector Machine learning accurately predicted 92.5% of those diagnosed with PTSD. Both sensitivity (90%) and specificity (95%) were excellent. Authors report that limbic structure and prefrontal cortex provided the most discriminant features.

1 study (N = 87) assessed fMRI and DTI scans in male veterans who had PTSD (PCL-5 designated), vs. post-concussion syndrome + PTSD, and vs. trauma-exposed controls, and found that Support Vector Machine learning accurately predicted 84% of those diagnosed with PTSD. Authors report that PTSD was associated with hippocampalstriatal hyperconnectivity.

1 study (N = 57) assessed fMRI and sMRI scans in earthquake survivors who had PTSD (SCID designated), vs. trauma-exposed controls, and vs. healthy controls, and found that Support Vector Machine learning accurately predicted 89% of those diagnosed with PTSD when compared to healthy controls, and 68% of those diagnosed with PTSD when compared to trauma-exposed controls.

1 study (N = 97) assessed sMRI scans in veterans who had PTSD (CAPS designated) vs. trauma-exposed controls and found that Support Vector Machine learning accurately predicted 69% of those diagnosed with PTSD. Sensitivity was moderate (58%), and specificity was good (81%). Authors report that surface area in the right posterior cingulate was selected as an important feature for classification of PTSD.

1 study (N = 150) assessed sMRI scans in earthquake survivors who had PTSD (CAPS and PCL designated), vs. trauma-exposed controls, and vs. healthy controls, and found that Support Vector Machine learning accurately predicted 91% of those diagnosed with PTSD when compared to healthy controls, and 67% of those diagnosed with PTSD when compared to trauma-exposed controls. Both sensitivity (95%) and specificity (87.5%) were excellent.

Neuropsychological tests

1 study (N = 60) assessed questionnaires and sleep assessments in people exposed to sexual assault who had PTSD (CAPS designated), vs. trauma-exposed controls, and vs. healthy controls, and found that Support Vector Machine learning accurately predicted 80% of those diagnosed with PTSD when compared to all controls, and 70% of those diagnosed with PTSD when compared to trauma-exposed controls. Sensitivity was good (87%), and specificity was moderate (65%). Authors report that sleep characteristics were the primary features that could differentiate those with PTSD from those without.

1 study (N = 391) assessed questionnaires in people exposed to mixed traumas who had PTSD (CAPS designated), vs. trauma-exposed controls, and found that Sequential Minimal Optimisation, Multi-Layer Perceptron, or Naive Bayes machine learning techniques accurately predicted 74-79% of those diagnosed with PTSD.

Audio recordings

1 study (N = 52) assessed audio recordings in people on youtube or in hospital who had PTSD (CAPS designated), vs. healthy controls, and found that Support Vector Machine learning, Deep Belief Network, or Transfer Learning techniques accurately predicted 58-75% of those diagnosed with PTSD. Authors report that patients were diagnosed with PTSD by analysing speech signals. 1 study (N = 25) assessed audio recordings in females who had PTSD and panic attacks or agoraphobia found that Support Vector Machine learning, k-nearest Neighbors, or Multi-Layer Perceptron machine learning techniques accurately predicted 82-90% of those diagnosed with PTSD.	
Consistency in results	Unable to assess consistency; no measure is reported.
Precision in results	Unable to assess precision; no CIs are reported.
Directness of results	Direct

Spoont MR, Williams JW, Jr., Kehle-Forbes S, Nieuwsma JA, Mann-Wrobel MC, Gross R

Does This Patient Have Posttraumatic Stress Disorder?: Rational Clinical Examination Systematic Review

JAMA 2015; 314: 501-10

View review abstract online

Comparison	**Sensitivity and specificity of the Primary Care PTSD Screen (PC-PTSD) and the PTSD Checklist (PCL) screening tools for predicting a diagnosis of PTSD.**
Summary of evidence	**Moderate to high quality evidence (large samples, some inconsistency, appears precise, direct) suggests reasonable sensitivity and good specificity of the Primary Care PTSD Screen (PC-PTSD) and the PTSD Checklist (PCL) screening tools for predicting a diagnosis of PTSD.**
Sensitivity and specificity	
PC-PTSD 5 studies, N = 1,100, threshold = ≥3 Sensitivity = 0.69, 95%CI 0.55 to 0.81 Specificity = 0.92, 95%CI 0.86 to 0.95 PCL 6 studies, N = 4,906, threshold 38 to 44	

Sensitivity = 0.70, 95%CI 0.64 to 0.77 Specificity = 0.90, 95%CI 0.84 to 0.93	
Consistency	Some inconsistency
Precision	Appears precise
Directness	Direct

Utzon-Frank N, Breinegaard N, Bertelsen M, Borritz M, Eller NH, Nordentoft M, Olesen K, Rod NH, Rugulies R, Bonde JP

Occurrence of delayed-onset post-traumatic stress disorder: a systematic review and meta-analysis of prospective studies

Scandinavian Journal of Work, Environment & Health 40: 215-29

View review abstract online

Comparison	**Prevalence of delayed-onset PTSD (>6 months post-trauma).**
Summary of evidence	**Moderate to high quality evidence (large sample, inconsistent, appears precise, direct) suggests around 24.5% of people diagnosed with PTSD have a delayed onset, with most experiencing earlier subclinical symptoms. The prevalence of delayed-onset PTSD is highest in professional groups and those who experienced combat trauma.**
Delayed-onset PTSD	
39 prospective studies, N = 30,210 Overall prevalence of PTSD = 19.7%, 95%CI 15.8 to 24.2%, I^2 = 90% Overall prevalence of delayed-onset PTSD = 5.6%, 95%CI 4.3 to 7.3%, I^2 = 91% Proportion of delayed-onset PTSD relative to all cases of PTSD = 24.5%, 95%CI 19.5 to 30.3%, I^2 = 94% Subgroup analyses showed delayed-onset PTSD was higher among professional groups compared to non-professional victims (37.6% vs. 20.3%). It was also higher in those exposed to combat than other trauma types (39.9% vs. 17-26%). There were no moderating effects of early (1-6 months after the trauma) vs. late (>9 months after the trauma) baseline assessment, assessment tool, or other study methods. Authors report that most people with delayed-onset PTSD experienced early subclinical symptoms.	
Consistency	Inconsistent

Precision	Appears precise
Directness	Direct

Explanation of acronyms

CAPS = clinical-administered PTSD scale, CI = confidence interval, DSM = American Psychiatric Association's Diagnostic and Statistical Manual, DTI = diffusion tensor imaging, fMRI = functional magnetic resonance imaging, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), ICD = World Health Organisation's International Classification of Diseases, N = number of participants, *p* = statistical probability of obtaining that result, PC-PTSD = Primary Care PTSD Screen, PCL = PTSD Checklist, PTSD = Post-traumatic stress disorder, SCID = structural clinical interview for the DSM, sMRI = structural magnetic resonance imaging, vs. = versus

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[9].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardsed mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) that allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 and over represents a large effect[9].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[10]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship

between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[9];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[11].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Larsen SE, Pacella ML (2016): Comparing the effect of DSM-congruent traumas vs. DSM-incongruent stressors on PTSD symptoms: A meta-analytic review. *Journal of Anxiety Disorders* 38: 37-46.
4. Spoont MR, Williams JW, Jr., Kehle-Forbes S, Nieuwsma JA, Mann-Wrobel MC, Gross R (2015): Does This Patient Have Posttraumatic Stress Disorder?: Rational Clinical Examination Systematic Review. *JAMA* 314: 501-10.
5. Utzon-Frank N, Breinegaard N, Bertelsen M, Borritz M, Eller NH, Nordentoft M*, et al.* (2014): Occurrence of delayed-onset post-traumatic stress disorder: a systematic review and meta-analysis of prospective studies. *Scandinavian Journal of Work, Environment & Health* 40: 215-29.
6. Bressler R, Erford BT, Dean S (2018): A systematic review of the Posttraumatic Stress Disorder Checklist (PCL). *Journal of Counseling and Development* 96: 167-86.
7. Ramos-Lima LF, Waikamp V, Antonelli-Salgado T, Passos IC, Freitas LHM (2020): The use of machine learning techniques in trauma-related disorders: a systematic review. *Journal of Psychiatric Research* 121: 159-72.
8. Martinez W, Polo AJ, Zelic KJ (2014): Symptom variation on the trauma symptom checklist for children: a within-scale meta-analytic review. *Journal of Traumatic Stress* 27: 655-63.
9. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
10. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
11. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*