The mental health needs of serving and ex-Service personnel:

A systematic review

TECHNICAL ANNEX

FiMT forces in mind trust
SUCCESSFUL SUSTAINABLE TRANSITION

COMMUNITY FUND

NatCen Social Research

At **NatCen Social Research** we believe that social research has the power to make life better. By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

NatCen Social Research
35 Northampton Square
London EC1V 0AX
T 020 7250 1866
www.natcen.ac.uk

A Company Limited by Guarantee
Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland (SC038454)
This project was carried out in compliance with ISO20252

Authors: Daniel Phillips, Anna Marcinkiewicz, Robert Wishart, Emma Forsyth, Anysia Nguyen, Sarah Lynch-Huggins, Fiona Gogescu, Adam Gilbert, Sokratis Dinos, Martina Vojtkova, John Eyers, David Denney

Date: 07/04/2020

Prepared for: Forces in Mind Trust (FiMT)

Funding: This systematic review was funded by the Forces in Mind Trust.

Contents

Technical Annex Overview

This Technical Annex provides additional methodological detail and detailed findings not included in the main systematic review report (Phillips et al. 2020a). It is divided into six main sections:

1. **Systematic Review Methodology**: the first chapter provides methodological details such as detailed inclusion criteria, full lists of resources searched and search syntax, templates for the extraction of data from included studies and tools used to critically appraise included studies, detailed summary of approach to research synthesis.

2. **Critical Appraisal**: the second chapter provides full details on the critical appraisal process, tools used to assess studies, overall results of the critical appraisal and a study-by-study summary of the critical appraisal of all studies included in the review.

3. **Evidence Map Methodology:** the third chapter summarises the methodology for the Evidence Map.

4. **Stakeholder Interview methodology:** the fourth chapter provides additional detail on our approach to sampling, interviewing and the synthesis of findings from the stakeholder interviews.

5. **Effectiveness findings: additional reporting:** the fifth chapter provides forest plots for all meta-analyses and further information about the assessment of publication bias

6. **Tables of characteristics**

7. **References**

1 Methodology

See Chapter 3 of the main report for a short overview of the study methodology.

1.1 Inclusion criteria

Table A 1.1 through Table A 1.4 summarise the inclusion criteria designed to determine inclusion or exclusion in the review.

Table A 1.1 General inclusion criteria applying to all evidence

Category	Inclusion criteria
The condition of interest	We defined mental health issues as including both mental health conditions and behaviours. **Conditions** We included conditions listed in ICD-11 (2019) under 'Mental, behavioural or neurodevelopmental disorders'. These included the following disorders: • Mood [affective] disorders including bipolar affective disorder, depressive episodes and other mood disorders; • Personality disorders; • Impulse control disorders; • Disorders due to substance use or addictive behaviours; • Disorders specifically related to stress, including post-traumatic stress disorder and complex post-traumatic stress disorder, • Anxiety or fear-related disorders, including generalised anxiety disorder and panic disorder; • Schizophrenia or other primary psychotic disorders. Evidence relating solely to physical health conditions was *not* included in the review. We included evidence drawing links between mental and physical health. **Behaviours** We also included mental health-related behaviours not included in ICD-11 categorisation, such as: • Gambling addictive behaviours • Alcohol and substance misuse • Violent or aggressive behaviours, including domestic violence • Suicide and attempted suicide • Intentional self-harm and parasuicide
Populations of interest	Our populations of interest cover the 'Armed Forces Community', defined in (MOD, 2018d) as Serving Personnel, Regular and Reservists, Veterans and military families. We further define these populations so that all evidence had to focus on at least one of: • Serving personnel;

	• ex-Service personnel; • Mobilised and non-mobilised Reservists; • Early Service Leavers; • National Service veterans; • Families of the above groups of people. We define families broadly to include parents, siblings, children (including adopted children and step-children or those from 'blended' families), carers, and spouses and partners (including same-sex spouses and partners) We did not include studies that provide evidence relating to government officials only, for example studies relating to MOD civil servants.
Publication type	We included both published and unpublished (grey) literature in English.

Table A 1.2 Inclusion criteria for the prevalence domain

Category	Inclusion criteria
Publication date	We included studies published from October 2012 onwards, to allow for a small overlap between the period covered by this review and the period covered by the MHF (2013) review that it updates.
Study focus	We included studies focusing on the **prevalence** of mental health conditions and behaviours as set out in Table A 1.1, General Inclusion Criteria.
Study design	We included the following study designs: • Studies that self-identify as evidence reviews • Prospective and retrospective cohort studies (i.e. longitudinal studies that follow a sample over time to see how their individual characteristics and differences affect outcomes in question) • Case-control studies (i.e. studies comparing different outcomes in two groups based on their supposed causal characteristics and factors) • Cross-sectional studies (i.e. studies that analyse data across populations at a specific timepoint of interest)
Types of measure	We included studies that provided estimates of prevalence rates of mental health conditions and behaviours, either as a percentage, or as the number of cases per 10,000 or 100,000 people. Prevalence rates were based on at least one of the following: • Conditions diagnosed by a medical professional; • Self-reported symptoms and/or self-reported diagnosis; • Use of screening tools or questionnaires such as the Alcohol Use Disorders Identification Test[1] (AUDIT) or Post-Traumatic Stress Disorder Checklist[2] (PCL) researchers to determine prevalence.
Geographic location	We only included evidence relating to UK populations.

[1] https://www.who.int/substance_abuse/publications/audit/en/
2 https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp

Table A 1.3 Inclusion criteria for the experience domain

Category	Inclusion criteria
Publication date	We included studies published from October 2012 onwards, to allow for a small overlap between the period covered by this review and the period covered by the MHF (2013) review that it updates.
Study focus	To be included in the review, a study had to report on at least one of the following experiences, as they relate to conditions and behaviours as set out in Table A 1.1, General Inclusion Criteria: • Experiences of mental health conditions and behaviours; • Experiences of help-seeking, disclosure and accessing treatment for mental health problems; • Experiences of mental health stigma and/or other barriers or enablers to seeking treatment or support
Study design	We included quantitative, qualitative or mixed-methods studies that use primary data, such as survey data, to assess experience. We also included qualitative studies focusing on the subjective experience of individuals, including in-depth interviews, focus groups and other qualitative approaches such as observations or ethnographic research. We also included studies that self-identify as systematic reviews.
Types of measure	Any measure of conditions and behaviours listed in Table A 1.1, General Inclusion Criteria.
Geographic location	We only included evidence relating to UK populations.

Table A 1.4 Inclusion criteria for the effectiveness domain and evidence map

Category	Inclusion criteria
Publication date	We included studies published from October 2012 onwards or studies included in the Mental Health Foundation (2013) and that met other criteria for the **Effectiveness** domain.
Study focus (intervention types)	We included studies providing evidence on **interventions** to address mental health conditions and behaviours, as set out in Table A 1.1, General Inclusion Criteria.
Study types	We included primary studies with an experimental or quasi-experimental design. These include the following: randomised controlled trials, regression discontinuity designs, natural experiments and instrumental variable estimation, and studies with pre and post-intervention outcomes data for an intervention and comparison groups that control for confounding through statistical matching, difference-in-differences (or fixed- or random-effects models with an interaction term between time and intervention for baseline and follow-up observations) and interrupted time series. We also include studies that self-identify as systematic reviews and meta-analyses.
Study comparisons	We included studies that make comparisons between an intervention group and populations that receive 'business as usual' service access (sometimes called standard care), or a different form of intervention.
Types of measure (outcome)	We did not exclude studies based on the type of outcomes they report. However, we synthesised evidence on our primary outcomes of interest only. Primary outcomes of interest had to be related to mental health conditions and behaviours outlined in Table A 1.1, General Inclusion Criteria: The Condition of Interest. They included:

	• symptoms, intensity or number of episodes of condition or behaviour; • self-reported satisfaction or mental health and wellbeing status; • number of diagnoses made as part of a prevention intervention
Geographic location	We included UK based evidence and primary studies from Australia, Canada, Denmark, France, Germany, Israel, Netherlands, New Zealand, Norway, Sweden and USA, and reviews or meta-analyses synthesising international literature.

1.2 Search

1.2.1 Databases searched

The following academic databases were searched for relevant studies. The databases were selected in collaboration with our advisory group and our search strategy expert. We considered the number of relevant results the databases were likely to return and did not search databases that are no longer updated, such as the Centre for Reviews and Dissemination, and those which we did not think would return unique results not captured from our other databases.

Table A 1.5 List of databases and dates searched

Database	Date searched
Cochrane Library	13/08/2018
EBSCO Cumulative Index to Nursing and Allied Health Literature	29/08/2018
Embase	23/08/2018
Epistemonikos	29/08/2018
MEDLINE	13/08/2018
PsycINFO	24/08/2018

1.2.2 Websites/online repositories searched

A range of relevant websites and online repositories in the fields of military research, mental health, and government departments were searched to supplement our database searches. We consulted with the review's advisory group and selected the websites of organisations that are known to produce and publish research on military mental health. The search strategy we used is explained in section 1.2.4. Table A 1.6 below outlines the websites we searched and dates of the searches:

Table A 1.6 List of websites and dates searched

Name	Link	Date searched
Army Families Federation	https://aff.org.uk/about-aff/research-papers/	31/07/2018
Centre for Mental Health	https://www.centreformentalhealth.org.uk/Pages/Category/publications?Take=3	30/07/2018
Combat Stress	https://www.combatstress.org.uk/about-us/research	27/07/2018
Department of Health & Social Care	https://www.gov.uk/government/publications?keywords=&publication_filter_option=all&topics%5B%5D=all&departments%5B%5D=department-of-health-and-social-care&official_document_status=all&world_locations%5B%5D=all&from_date=&to_date=	31/07/2018
Economic and Social Research Council	https://esrc.ukri.org/research/	23/08/2018
EThOS	http://ethos.bl.uk/AdvancedSearch.do;jsessionid=0CD5721370A2A5BD4EB852EF43D42646?new=1	31/07/2018
Forces in Mind Trust	http://www.fim-trust.org/reports/	30/07/2018
Kings Centre for Military Health Research	https://www1.kcl.ac.uk/kcmhr/pubdb/	30/07/2018
King's Fund	https://www.kingsfund.org.uk/publications	31/07/2018
Mental Health Foundation	https://www.mentalhealth.org.uk/publications/listing?search=armed+forces&field_focus_section_target_id=All	30/07/2018
Ministry of Defence (MoD)	https://www.gov.uk/government/organisations/latest?keywords=mental+health&organisations%5B%5D=ministry-of-defence&page=6	31/07/2018
Navy Families Federation	https://nff.org.uk/	21/08/2018
Networked Digital Library of Theses	http://search.ndltd.org/index.php#	31/07/2018
ProQuest Dissertation and Theses Database	https://search.proquest.com/results/3CA87C0D923241DFPQ/1?accountid=14510	04/09/2018
Royal Air Force Families Federation	https://www.raf-ff.org.uk/page/2/?s=mental+health	21/08/2018
Royal British Legion	https://www.britishlegion.org.uk/get-involved/campaign/public-policy-and-research/research-and-reports/	22/08/2018
Substance Abuse & Mental Health Services Administration	https://www.samhsa.gov/ebp-resource-center (NREPP no longer available)	28/08/2018
Veterans and Families Research Hub	https://www.vfrhub.com/vfr-research-search-results/?fwp_theme=mental-health	30/07/2018

Veterans UK	https://www.gov.uk/government/publications?departments%5B%5D=veterans-uk	30/07/2018
Veterans' Gateway	https://support.veteransgateway.org.uk/app/category_list/kw/mental%20health	21/08/2018

1.2.3 Example of full search string

Below is an example of the review search strategy, as applied to MEDLINE. The search string was divided into two parts – terms designed to return studies relating to our prevalence and experience domains, and terms relating to our effectiveness domain. The terms reflected the different inclusion criteria relating to each domain. Medline search strings included MeSH terms (medical subject headings) donated with a forward slash (/) and then search terms, denoted by terms in parentheses.

Ovid MEDLINE(R) and In-Process & Other Non-Indexed Citations, and Daily

1 military personnel/ or veterans/ (49927)
2 (veteran* or ex-service* or "service leaver*" or service-leaver*).ti,ab,kw. (30400)
3 ("armed forces" or "armed services" or "military personnel" or "service personnel" or servicemen or "service men" or servicewomen or "service women" or "forces personnel" or reservist*).ti,ab,kw. (11073)
4 (naval or navy or army or "air force" or airforce or RAF).ti,ab,kw. (34328)
5 or/1-4 (96914)
6 Mental Health/ (31512)
7 exp Mental Disorders/ (1132078)
8 exp Substance-Related Disorders/ (259829)
9 "trauma and stressor related disorders"/ or adjustment disorders/ or stress disorders, traumatic/ or combat disorders/ or psychological trauma/ or stress disorders, post-traumatic/ or stress disorders, traumatic, acute/ or persian gulf syndrome/ or gambling/ or domestic violence/ or child abuse/ or child abuse, sexual/ or elder abuse/ or intimate partner violence/ or spouse abuse/ or battered women/ (80578)
10 (((wife or wives or spouse* or partner* or child* or family or families or domestic) adj3 (abus* or assault* or violen* or rape* or beat* or batter* or coerc* or harass* or victimi* or ill-treat* or perpetrat* or misogyn*)) or "intimate terrorism" or IPV or ((forc* or unwanted) adj3 (sex* or intercourse))).ti,ab,kw. (37443)
11 (gambling or gamble* or betting or wagering).ti,ab,kw. (8283)
12 ("post traumatic stress disorder*" or PTSD or "combat disorder*").ti,ab,kw. (23199)
13 ((drug* or opioid or substance* or alcohol) adj3 (abus* or misus* or depend* or addict*)).ti,ab,kw. (112844)
14 (((mental* or psychiatr*) adj3 (ill* or disorder* or health)) or (psycholog* adj3 adapt*)).ti,ab,kw. (205605)
15 ((combat adj3 (exposure or stress or disorder*)) or "symptomatic ill health" or "war syndrome" or "gulf syndrome" or somatoform).ti,ab,kw. (4835)
16 exp Somatoform Disorders/ (18192)
17 self-injurious behavior/ or self mutilation/ or suicide/ or suicidal ideation/ or suicide, attempted/ (59392)
18 (suicid* or self-harm* or self-injur* or self-mutilat*).ti,ab,kw. (73253)
19 mental health services/ or child guidance/ or community mental health services/ or exp counseling/ or emergency services, psychiatric/ or exp psychotherapy/ or adaptation, psychological/ (338484)
20 (("mental health" or psychiatr* or counsel* or psychotherap* or "family health") adj2 service*).ti,ab,kw. (25717)
21 "delivery of health care"/ or "delivery of health care, integrated"/ or health services accessibility/ or help-seeking behavior/ or patient acceptance of health care/ or "health services needs and demand"/ (227020)
22 (((deliver* or access* or accept* or need* or demand*) adj3 ("health care" or healthcare or "health service*")) or ((help or treatment) adj2 seek*)).ti,ab,kw. (80947)
23 or/6-22 (1805665)

24 exp United Kingdom/ or ("united kingdom" or UK or britain or british or english or scottish or scots or welsh or england or scotland or wales or "northern ireland" or ulster).ti,ab,kw. (597023)
25 5 and 23 and 24 (932)
26 limit 25 to yr="2012 -Current" (316)
27 (australia* or canada or canadi* or denmark or danish or france or french or german* or israel* or netherlands or dutch or "new zealand*" or sweden or swedish or USA or "united states" or american or african-american).ti,ab,kw. (1130136)
28 exp canada/ or exp united states/ or israel/ or exp france/ or exp germany/ or netherlands/ or exp denmark/ or sweden/ or exp australia/ or new zealand/ (1932790)
29 24 or 27 or 28 (3002840)
30 ("quasi experiment*" or quasi-experiment* or "random* control* trial*" or "random* trial*" or RCT or (random* adj3 allocat*) or matching or "propensity score" or PSM or "regression discontinuity" or "discontinuous design" or RDD or "difference in difference*" or difference-in-difference* or "diff in diff" or DID or "case control" or cohort or "propensity weighted" or propensity-weighted or "interrupted time series" or (before adj5 after) or (pre adj5 post) or ((pretest or "pre test") and (posttest or "post test")) or "research synthesis" or "scoping review" or "rapid evidence assessment" or "systematic literature review" or "Systematic review" or "Meta-analy*" or Metaanaly* or "meta analy*" or "Control* evaluation" or "Control treatment" or "instrumental variable*" or heckman or IV or ((quantitative or "comparison group*" or counterfactual or "counter factual" or counter-factual or experiment*) adj3 (design or study or analysis)) or QED).ti,ab,kw. (3234402)
31 clinical trial/ or clinical trial, phase i/ or clinical trial, phase ii/ or clinical trial, phase iii/ or clinical trial, phase iv/ or controlled clinical trial/ or randomized controlled trial/ or pragmatic clinical trial/ or controlled clinical trials as topic/ or non-randomized controlled trials as topic/ or randomized controlled trials as topic/ or pragmatic clinical trials as topic/ or case-control studies/ or retrospective studies/ or controlled before-after studies/ or interrupted time series analysis/ or random allocation/ or cohort studies/ or follow-up studies/ or longitudinal studies/ or prospective studies/ or retrospective studies/ or propensity score/ (2735896)
32 evaluation studies/ or program evaluation/ (286462)
33 or/30-32 (5206493)
34 5 and 23 and 29 and 33 (4519)
35 limit 34 to yr="2012 -Current" (2151)
36 26 and 33 (116)

For the Cochrane Library, the following search strategy was used:

#1 MeSH descriptor: [Military Personnel] this term only 820
#2 MeSH descriptor: [Veterans] this term only 790
#3 (veteran* or ex-service* or "service leaver*" or service-leaver*):ti,ab,kw 4496
#4 ("armed forces" or "armed services" or "military personnel" or "service personnel" or servicemen or "service men" or servicewomen or "service women" or "forces personnel" or reservist*).:ti,ab,kw 1276
#5 (naval or navy or army or "air force" or airforce or RAF):ti,ab,kw 1981
#6 {OR #1-#5} 8178
#7 MeSH descriptor: [Mental Health] this term only 1142
#8 MeSH descriptor: [Mental Disorders] explode all trees 61453
#9 MeSH descriptor: [Substance-Related Disorders] explode all trees 13277
#10 Any MeSH descriptor 3412
#11 (((wife or wives or spouse* or partner* or child* or family or families or domestic) NEAR/3 (abus* or assault* or violen* or rape* or beat* or batter* or coerc* or harass* or victimi* or ill-treat* or perpetrat* or misogyn*)) or "intimate terrorism" or IPV or ((forc* or unwanted) NEAR/3 (sex* or intercourse))):ti,ab,kw 2617
#12 (gambling or gamble* or betting or wagering):ti,ab,kw 680
#13 ("post traumatic stress disorder*" or PTSD or "combat disorder*"):ti,ab,kw3270
#14 ((drug* or opioid or substance* or alcohol) NEAR/3 (abus* or misus* or depend* or addict*)):ti,ab,kw 13127
#15 (((mental* or psychiatr*) NEAR/3 (ill* or disorder* or health)) or (psycholog* NEAR/3 adapt*)):ti,ab,kw 24787
#16 ((combat NEAR/3 (exposure or stress or disorder*)) or "symptomatic ill health" or "war syndrome" or "gulf syndrome" or somatoform):ti,ab,kw 839
#17 MeSH descriptor: [Somatoform Disorders] explode all trees 579
#18 MeSH descriptor: ["self-injurious behavior"] explode all trees 1100

#19 (suicid* or self-harm* or self-injur* or self-mutilat*):ti,ab,kw 4216
#20 MeSH descriptor: ["mental health services"] explode all trees 27417
#21 (("mental health" or psychiatr* or counsel* or psychotherap* or "family health") NEAR/2 service*):ti,ab,kw 615
#22 MeSH descriptor: ["delivery of health care"] explode all trees 4274
#23 (((deliver* or access* or accept* or need* or demand*) NEAR/3 ("health care" or healthcare or "health service*")) or ((help or treatment) NEAR/2 seek*)):ti,ab,kw 3337
#24 {OR #7-#23} 104148
#25 MeSH descriptor: ["United Kingdom"] explode all trees 5683
#26 ("united kingdom" or UK or britain or british or english or scottish or scots or welsh or england or scotland or wales or "northern ireland" or ulster):ti,ab,kw 32702
#27 (australia* or canada or canadi* or denmark or danish or france or french or german* or israel* or netherlands or dutch or "new zealand*" or sweden or swedish or USA or "united states" or american or african-american):ti,ab,kw 132210
#28 MeSH descriptor: [canada] explode all trees 34959
#29 {OR #25-#28} 167635
#30 ("quasi experiment*" or quasi-experiment* or "random* control* trial*" or "random* trial*" or RCT or (random* NEAR/3 allocat*) or matching or "propensity score" or PSM or "regression discontinuity" or "discontinuous design" or RDD or "difference in difference*" or difference-in-difference* or "diff in diff" or DID or "case control" or cohort or "propensity weighted" or propensity-weighted or "interrupted time series" or (before NEAR/5 after) or (pre NEAR/5 post) or ((pretest or "pre test") and (posttest or "post test")) or "research synthesis" or "scoping review" or "rapid evidence assessment" or "systematic literature review" or "Systematic review" or "Meta-analy*" or Metaanaly* or "meta analy*" or "Control* evaluation" or "Control treatment" or "instrumental variable*" or heckman or IV or ((quantitative or "comparison group*" or counterfactual or "counter factual" or counter-factual or experiment*) NEAR/3 (design or study or analysis)) or QED):ti,ab,kw 392893
#31 MeSH descriptor: ["clinical trial"] explode all trees 170781
#32 {OR #30-#31} 480182
#33 #6 and #24 and #29 and #32 with Cochrane Library publication date between Jan 2012 and Aug 2018 **237**

1.2.4 Search terms for grey literature search

We developed a concise version of our full search string to enable us to search a range of websites and online repositories that differ in their search function and do not allow the utilisation of complex search strings. Some websites, such as the Veteran and Families Research Hub have a dedicated repository for research on military mental health. Others, such as the Ministry of Defence, required a more intricate combination of search terms and the use of Boolean operators to return relevant results. Key search terms used across the websites included “mental health”, '"armed forces", "service personnel", “military” and “veterans”.

1.3 Screening

We used dedicated systematic review software called Abstrackr to help us organise the search results returned from the database search and facilitate the screening process[3]. As the volume of evidence identified through the search was large, we made use of machine learning in the Abstrackr software to identify the studies likely to be relevant to our review. Abstrackr learns from inclusion and exclusion decisions made by the research team to predict relevant records (Gates et al. 2018). These papers were then prioritised for screening which allowed us to concentrate our resources and made the review process more efficient. For the results returned from websites and online repositories, screening was conducted manually against the inclusion criteria.

[3] www.abstrackr.cebm.brown.edu

All search results were screened at two stages, at title and abstract and at full-text. At title and abstract, studies that appeared to be relevant and studies for which inclusion was not clear were included for full-text review. The studies we included at title and abstract were then retrieved and downloaded from the academic journal or website. Reviewers then considered the full-text of the paper against the inclusion criteria to decide whether to include the paper in the review. For the prevalence and experience domains, where a primary study was synthesised in a review, we excluded the study and took the decision to summarise the findings at review level. For the Effectiveness domain, we did the opposite, checking all includable reviews for primary studies that met our inclusion criteria so that they could be included in our meta-analysis and excluding reviews from our synthesised findings.

All screening tools were piloted before use to promote inter-coder reliability. After piloting, all screening was undertaken by a single reviewer. Reviewers recorded the reason for exclusion at full-text and all texts included at this stage were checked by a second reviewer before proceeding to data extraction. This final stage of screening ensured that the final selection of studies addressed the review's research questions and met the inclusion criteria for study design.

1.4 Data Extraction

Data was extracted from studies included in the review. As the scope of the review was broad, covering different areas of interest and types of evidence, the approach to data extraction differed across the three domains and a bespoke template was created for each domain in Excel. The data extraction sheets were developed to capture all the relevant information from the included studies, including basic descriptive information and study characteristics, as well as the key findings and data that could be used to answer the research questions.

Table A 1.7 summarises the data extracted for our prevalence and experience domains.

Table A 1.7 Data extraction template for prevalence and experience domains

Category	Sub-category	Description
General/Descriptive information	Population	Select all groups the study covers: Serving personnel; Ex-service personnel; reservists; early service leavers; national serviceman; families of the above groups; other
	Further comments on population	Use this box: to further define what family members are included in the study (families include parents, siblings, children, carers, and spouses and partners) or to describe a population if 'other' is ticked on the dropdown above
	Conditions	Select the conditions covered by the study: Mood [affective] disorders; Personality disorders; Impulse control disorders; Disorders due to substance use or addictive behaviours; Disorders specifically related to stress; Anxiety or fear-related disorders, including generalised anxiety disorder and panic disorder; Schizophrenia or other primary psychotic disorders; Other
	Further comments on conditions	Use this box: to further define the condition. If 'other', provide further details here

	Behaviours	Select the behaviours covered by the study: Substance misuse; Gambling addictive behaviours; Violent or aggressive behaviours, including domestic violence; Suicide and attempted suicide; Intentional self-harm and parasuicide; Other
	Further comments on behaviours	Use this box: to further define the behaviour. If 'other', provide further details here
	Study design	Select the design that best describes the study methodology •Evidence review or meta-analysis •Primary study (including secondary analysis of primary data)
	Primary: further comments on design	For all primary studies, provide additional details -Sample size -Sample strategy (random sample, purposive sample) -Is the study cross-sectional (single time-point) or longitudinal (multiple timepoints) -Does the study make a comparison between two or more groups. If so, describe.
	Synthesis: further comments on design	For all synthesis studies, provide additional details Describe the study design based on the information available in the study -What does the study describe itself as? Meta-analysis, systematic review, rapid review etc -Does the review set out its inclusion criteria -Does the review clearly describe its search strategy (search terms used)? -Does the review clearly describe the resources searched (databases, websites etc)? -Does the review quality appraise the studies it includes?
Prevalence	Estimates of prevalence	Quantitative measure(s) of prevalence for mental health conditions and behaviours in the UK or relating to UK service and ex-service personnel. Provide detail. For example, rates are likely to be reported either as a fraction, as a percentage, or as the number of cases per 10,000 or 100,000 people.
	Prevalence for sub-groups	Provide any differences in reported prevalence for different population groups or demographic groups. For primary studies include information relating to primary research rather than contextual information from other studies.
	Factors (risk and protective) associated with prevalence	Describe any factors reported by studies that explain variations in the prevalence of mental health conditions and behaviours (to include risk or protective factors). Risk factors are those that make incidence of a condition or behaviour more likely. Protective factors are those that make them less likely.
	Life-cycle stage	Describe which stages of the life narrative model evidence relates to and explain your choice: -pre-service -in-service -post-service

		It is possible to select multiple options. If unclear, provide explanation why
	Estimates of prevalence for other contexts	Provide any estimates of prevalence used as comparisons – e.g. for the general population or for other countries or contexts (not the UK). If the research refers to rates of prevalence in other contexts, populations (e.g. police, firefighters etc).
Experience	Experiences of mental health conditions and behaviours	What is the experience of mental health problems among UK former and current Service personnel and their families? Experiences may include: -Evidence related to the lived experience of mental health conditions e.g. including subjective understandings of what it is like, coping strategies etc. [specify which mental health condition/behaviour the experience relates to] -Evidence related to the impact of having a mental health condition or behaviour on quality of life (e.g. relationships, employment, finances etc.)
	How experiences during/ after military affect mental health	How do experiences during and after military service affect mental health problems and behaviours among UK former and current Service personnel and their families? (specify the sub-population) Experiences that are associated with mental health conditions and behaviours Examples After: lack of transition support, negative relationship change During: exposure to combat, traumatic experiences during deployment -State whether service/non-service related
	Types of support and treatment accessed/available	-What form does help seeking take? e.g. what types of support and treatment are military personnel and their families using? Include both types of help seeking support and treatment accessed/available
	Experience of help seeking behaviours	What is the experience of seeking and receiving treatment? -Describe experiences of seeking help and receiving mental health treatment/services e.g.: positive/negative experiences perceptions of available support gaps in support
	Experiences of stigma	Describe experiences of stigma -Evidence of stigma as a barrier to help seeking -How do military personnel and their families experience stigma? -Include how the study defines stigma (if relevant)
	Barriers and enablers to help seeking	Describe any other barriers and enablers to help seeking, other than stigma.

		-Enablers - What are the factors that make it possible/easier for service personnel and their families to seek support? -Barriers - Why do some service personnel and their families choose not to/are unable to seek support/treatment? e.g. lack of awareness, perceptions of MH care, reluctance to accept help/self-reliance NB: Distinguish between barriers/enablers to access and barriers/enablers to use if relevant -State whether the experiences are pre-, during or after service and whether they are service/non-service related
	Experiences of disclosure	What evidence is there regarding experiences of disclosure within the armed forces, both to superior officers and colleagues and to friends/family members? -Describe experiences of disclosure
	Life cycle stage	Describe which stages of the life narrative model evidence relates to and explain your choice: • pre-service • in-service • post-service It is possible to select multiple options. If unclear, provide explanation why. Provide page reference(s)
	Other	

For effectiveness, various types of data were extracted from the included papers. This included descriptive data such as publication data, country and intervention type; methodological information on study design, analytical methods and type of comparison (see Table A 1.8); and quantitative data for our outcomes of interest (see Table A 1.9).

Table A 1.8 Effectiveness data extraction template

Category	Description/guidance on coding
Countries	Select from dropdown menu. Select multiple options if necessary: UK, Australia, Canada, Denmark, France, Germany, Israel, Netherlands, New Zealand, Norway, Sweden, USA
Populations	Select all options that apply from drop-down: Serving: Regulars, Serving: Mobilised and non-mobilised Reservists, Veterans, Early Service Leavers, National Servicemen, Families of the above groups of people, Other
Populations description	Open answer. Use this box if 'other' was selected in 5.1 or to provide additional information.
Study design (select multiple options if necessary)	Select all options that apply from drop-down: Randomised Controlled Trial (RCT), Difference-in-Differences (DID), Instrumental Variables (IV), Regression Discontinuity Design (RDD), Matching, including Propensity Score Matching (PSM), Interrupted time series, Natural experiment, Other

Further comments on study design	Open answer. Use this box if 'other' was selected in 6.1 or to provide additional information. Some papers use multiple methods to analyse data. Please report all and do not pick only one. If multiple designs are selected, state here whether a paper used **multiple** methods to analyse data or **combined** two or more methods (for example, as when matching and DID are combined)
Intervention category (ordered as per map framework)	Select all options that apply
Intervention description	Open answer. Briefly describe the programme or intervention. This description should explain your selection from the intervention categories. This description should be around 4 sentences long. Describe all components of the intervention. Record the programme name **in bold** if any is provided. Provide page numbers for particularly relevant sections
Outcomes (ordered as per map framework)	Select all options that apply
Outcomes descriptions/definitions	Open answer. For each outcome in this category, provide a definition around a sentence or two long. Provide a page number. If there are multiple outcomes in any one category, provide a description of each. This section should basically explain why you have chosen a given outcome category. Provide a description of each outcome that does not fall into any of the categories above here. Again, each of these should be a sentence or two long and should be page referenced
Comparisons	Select all options that apply from drop-down: Intervention Vs Control (no intervention) Intervention Vs Alternative Intervention (active alternative) Intervention Vs 'Business as usual' Other
Other notes	Open answer. Use this section to highlight any additional relevant information or to make any notes if you would like a second opinion on any aspect (record what aspects and any related info.)

Table A 1.9 Effectiveness data extraction template for calculating effect sizes

Description	Question
Country in which programme was implemented	What country was the programme implemented in?
Design type	What type of study design is used? 1= Randomised controlled trial (RCT) (experiment with random assignment to households/individuals) 2= Cluster-RCT 3= Quasi-RCT (experiment with quasi-random assignment to households/individuals) 4= Cluster-quasi-RCT 5= RDD (quasi-experiment with discontinuity assignment) 6 = CBA (quasi-experiment with baseline and endline data collection) 7=Natural experiment

	8= Interrupted time series 9=Other (state in parentheses)
Methods used for analysis	Which methods are used to control for selection bias and confounding? 1=PSM 2=Covariate matching 3=DID 4=IV-regression 5=Heckman selection model 6= Fixed effects regression 7= Other regression 8=Other (state in parentheses)
	Which page(s) contain the effect size data?
Outcome	1=Yes 2=No 3=Partially
Definition of outcome	What definition of outcome x given, note page number
SD Outcome in treatment post intervention	State SD of post intervention outcome measure for control group
SD Outcome in control post intervention	State SD of post intervention outcome measure for treatment group
SE Outcome in treatment post intervention	State **standard error** of post intervention outcome for control group
SE Outcome in control post intervention	State **standard error** of post intervention outcome for treatment group
SD pooled across treatment and control	Pooled standard deviation
Outcome in control post intervention	State result of post intervention outcome for control group
Outcome in treatment post intervention	State result of post intervention outcome for treatment group
	Which page(s) or tables contain the sample size data?
Sample size metric	Sample size unit of analysis 1= individuals 2= Households 3= Groups (e.g. unit?) 4= Other 5= Not clear
Sample size baseline control	State sample size at baseline
Sample size baseline treatment	State sample size at baseline
Number with outcome in control post intervention	State sample size post intervention
Number with outcome in treatment post intervention	State sample size post intervention

OLS	OLS used? 1=Yes 2=No
Logistic	Logistic used? 1=Yes 2=No
Type of logistic	What type of logistic regression? 1=binomial 2=multinomial
GLS/WLS	GLS or WLS used? 1=Yes 2=No
Poisson	Poisson regression used? 1=Yes 2=No
other regression types	Other regression type used? Specify
multilevel models	Is this a multilevel model? 1=Yes 2=No
number of predictors	How many predictors/covariates (not including the intercept) are in the model?
continuous outcome	Is the outcome continuous? 1=Yes 2=No
dichotomous outcome	Is the outcome dichotomous? 1=Yes 2=No
multiple outcome categories	Does the outcome have more than 2 categories?
variance explained	What is the variance explained in the model?
type of coefficient	What is the coefficient type?
coefficient	What is the coefficient estimate?
standard error	What is the standard error of the coefficient estimate?
Standard deviation	Optional
Confidence Interval - lower bound	Optional
Confidence Interval - upper bound	Optional
t test	What is the t statistic associated with the focal predictor?
Wald test	What is the Wald statistic associated with the focal predictor?

1.4.1 Effectiveness

Studies included in the effectiveness domain of the review were critically appraised using a risk of bias tool originally formulated by Hombrados and Waddington (2012) and since adapted by systematic reviews by Oya et al. (2017) and Baird et al. (2014).

For papers included in the effectiveness, the 'Risk of Bias' was assessed the five key domains listed below. For each domain, papers were coded as 'Yes', 'No', 'Unclear' or 'No information', depending on their agreement with the description and scoring criteria listed.

Table A 1.10 Effectiveness 'Risk of Bias' categories

Category	Description	Further criteria for scoring
1) Selection bias and confounding	Selection bias and confounding: was the identification method free from any sources of bias or were sources of bias adequately corrected for with an appropriate method of analysis?	**Experimental approaches** (random allocation of the treatment): was the allocation free from any sources of bias or were sources of bias adequately corrected for with an appropriate method of analysis? i. Score "yes" if: a. A random component in the sequence generation process is described (e.g. Referring to a random number table) and if the unit of allocation is based on a sufficiently large sample size. b. The unit of allocation was by geographical/social unit, institution, team or professional and allocation was performed on all units at the start of the study; or if the unit of allocation was by beneficiary or group or episode of treatment and there was some form of centralised c. Randomisation scheme, an on-site computer system or sealed opaque envelopes were used. d. If the outcomes are objectively measurable. e. Baseline characteristics of the study and control/comparisons are reported and overall similar based on t-test or anova for equality of means across groups. f. if relevant (e.g. Cluster-rcts), authors control for external factors that might confound the impact of the programme (rain, infrastructure, community fixed effects, etc) through regression analysis or other techniques. g. The attrition and noncompliance rate is below 15%, or the study assesses whether drop-outs are random draws from the sample (e.g. By examining correlation with determinants of outcomes, in both treatment and comparison groups)? h. Score "unclear" if a) or b) not specified in the paper, c) scores "no" or if d) scores "no" but the authors controlled for the relevant differences through regression analysis. Score "no" otherwise. **Quasi-experimental approaches** (non-random allocation of the treatment): was the identification method free from any sources of bias or were sources of bias adequately corrected for with an appropriate method of analysis? I. Propensity score matching and combination of psm with panel models: i. Score "unclear" if :

		a. The study matched on either (1) baseline characteristics, (2) time-invariant characteristics or (3) endline variables not affected by participation in the programme. b. The variables used to match are relevant (e.g. Demographic and socio-economic factors) to explain a) participation and b) the outcome and thus there are not evident differences across groups in variables that explain outcomes. c. Except for kernel matching, the means of the individual covariates are equal for both the treatment and the control group after matching based on t-test for equality of means or anova. ii. score “no” otherwise. **Regression discontinuity design:** i. Score “yes” if: a. Allocation is made based on a pre-determined discontinuity blinded to participants or if not blinded, individuals cannot amend the assignment variable.the sample size immediately at both sides of the cut-off point is sufficiently large. b. The interval for selection of treatment and control group is reasonably small, or authors have weighted the matches on their distance to the cut-off point. c. the mean of the covariates of the individuals immediately at both sides of the cut-off point (selected sample of participants and non-participants) are overall not statistically different based on t test or anova for equality of means.. d. If relevant (e.g. Clustered studies) and although covariates are balanced, the authors include control for external factors through a regression analysis. i. Score “unclear” if a) or b is) not specified in the paper or d) scores “no” but authors control for covariate differences across participants and control individuals. ii. Score “no” otherwise.
2) Spill-overs, cross-overs and contamination	Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Score “yes” if the intervention is unlikely to spill-over to comparisons (e.g. Participants and non-participants are geographically and/or socially separated from one another and general equilibrium effects are not likely) and that the treatment and comparisons are isolated from other interventions which might explain changes in outcomes; or if the authors are able to account for contamination etc., for example through CA-ITT (contamination adjusted ITT). Score “no” if allocation was at the individual level and there are likely spill-overs within households and communities which are not controlled for, or other interventions likely to affect outcomes operating at the same time in either group. Score “unclear” if spill-overs and contamination are not addressed clearly.

3) Outcome reporting	Outcome reporting: was the study free from selective outcome reporting?	Score “yes” if there is no evidence that outcomes were selectively reported (e.g. All relevant outcomes in the methods section are reported in the results section). Score “no” if some important outcomes are subsequently omitted from the results or the significance and magnitude of important outcomes was not assessed. Score “unclear” if not specified in the paper.
4) Analysis reporting	Analysis reporting: was the study free from selective analysis reporting?	Score “yes” if authors use ‘common’ methods of estimation (i.e. Credible analysis method to deal with attribution given the data available). Score “no” if authors use uncommon or less rigorous estimation methods such as failure to conduct multivariate analysis for outcomes equations.
5) Unit of analysis	Is the study free from unit of analysis errors?	Unit of analysis (UoA) errors occur when the unit of analysis is different from the unit of randomisation (UoR) and authors do not correct for these unit of analysis differences in the standard error calculation. Score 'no' if randomisation was at the cluster level and this is not adjusted for in the analysis. Score 'Yes' if UoA = UoR OR if UoA != UoR and standard errors are clustered at the UoR level or if data is collapsed to the UoR level
6) Other Bias	Other risks of bias: Is the study free from other sources of bias? Including around measurement of the intervention	Score “yes” if the reported results do not suggest any other sources of bias Score “no” if other potential threats to validity are present and note these below (e.g. Unit of analysis error, coherence of results, data on the baseline collected retrospectively, information is collected using an inappropriate instrument or a different instrument/at different time/after different follow up period in the control and in the treatment group).

1.5 Synthesis

Our approach to synthesis varied across the three domains to reflect the different types of evidence included and research questions.

1.5.1 Prevalence and experience domains

For the prevalence domain, a ‘table of characteristics’ was created to summarise the evidence provided by the included studies. This was accompanied by a narrative synthesis of evidence on prevalence rates for different mental health conditions and behaviours by population of interest. The synthesis also included a discussion of the risk and protective factors associated with mental health conditions and behaviours. For evidence relating to the prevalence of mental health conditions and behaviours, rates were compared with the general population based on the Adult Psychiatric

Morbidity Survey (APMS) 2014 (also known as the National Study of Health and Wellbeing, McManus et al 2014)[4]. APMS findings were based on ICD-10 classifications[5], allowing the ease of comparison of data.

For the experience domain, the included studies were also captured in a table of characteristics providing details of the publication date, military population and design. We then synthesised the main findings of the papers in response to the reviews research questions on experience. Within each section, the narrative synthesis is structured by military population, to distinguish between serving and ex-Service personnel, as well by theme, to draw out the key findings in relation to the experiences of mental health problems and help-seeking.

Synthesis took place subsequent to data extraction. The data was analysed using the extraction sheets and grouped by thematic categories to reflect the research questions.

1.6 Synthesis: Effectiveness domain

For the studies that addressed the effectiveness of mental health interventions, it was decided that the most appropriate form of synthesis would be statistical meta-analysis, with any studies that could not be included in meta-analysis discussed narratively. See the main report Methodology Chapter (3) and Effectiveness Chapter (7) for an overview of our approach to the meta-analysis.

1.6.1 Statistical procedures and conventions

The following sections outlines the procedures and conventions that we used to carry out the quantitative analysis of results.

Effect Size Calculation

It was anticipated that the studies included in the meta-analyses would cover a diverse selection of interventions and consequently the meta-analyses would be analysed using random effects.

Where the necessary data could be extracted, standardised effect sizes were estimated. These are described as the “standard mean difference” (SMD) in the main report.

Continuous outcomes
For continuous outcomes, Hedge’s g sample-size corrected standardised mean differences (SMD) (Borenstein, et al., 2009), the variance, standard error and 95% confidence intervals were estimated using formulae by Bornstein and colleagues (2019, Chapter 4) as follows.
We estimated the sample estimate of the standardised mean difference (δ) (also known as Cohen’s d)

$$d = \frac{\bar{X}_t - \bar{X}_c}{S_{within}}$$

[4] The 2007 findings from APMS (McManus et al, 2007) were used to compare service personnel and the general population in the 2013 Mental Health Foundation review and as such, using the latest results from this survey will provide visibility of any changes to mental health provision in the general population.
[5] APMS 2014 study uses ICD-10, which pre-dates IDC-11 published this year.

Where $\bar{X}_t$ and $\bar{X}_c$ were the sample means in the two groups.

In the denominator S_{within} was the within-groups standard deviation, pooled across groups

$$S_{within} = \sqrt{\frac{(n_t - 1)S_t^2 + (n_c - 1)S_c^2}{n_t + n_c - 2}}$$

Where n_t and n_c were the sample sizes in the two groups, and S_t and S_c were the standard deviations in the two groups.

The variance of d was estimated as

$$V_d = \frac{n_t + n_c}{n_t n_c} + \frac{d^2}{2(n_t + n_c)}$$

The standard error of d was

$$SE_d = \sqrt{V_d}$$

The following correction factor J was applied to d to convert it to Hedges' g

$$J = 1 - \frac{3}{4df - 1}$$

Where df referred to degrees of freedom used to estimate S_{within} which for two independent groups was $n_t + n_c - 2$.

Then Hedges' g was

$$g = J \times d$$

The variance of g was

$$V_g = J^2 \times V_d,$$

and the standard error of g was

$$SE_g = \sqrt{V_g}.$$

Binary outcomes

Binary outcomes were estimated as Risk Ratios (RR), which were reported alongside the variance, standard error and 95% confidence intervals. These were arrived at using formulae by Borenstein and colleagues (2009, Chapter 5) as follows.

Table 1: Nomenclature for 2x2 table of outcome frequencies by treatment

	Success	Failure	Sample size N
Treatment group	A	B	n_t
Comparison group	C	D	n_c

The RR was the ratio of two risks as follow:

$$RR = \frac{A/n_t}{C/n_c}$$

For the purposes of the meta-analysis, we computed the log risk ratios and the standard errors of the log risk ratios and used these in the analysis

$$LogRR = \ln(RR)$$

With variance and standard error calculated as

$$V_{LogRR} = \frac{1}{A} - \frac{1}{n_t} + \frac{1}{C} - \frac{1}{n_c}$$

$$SE_{LogRR} = \sqrt{V_{LogRR}}$$

The resulting summary effect(s), confidence limits and so on were converted back into the original metric to facilitate interpretation of the findings using

$$RR = \exp(LogRR)$$

$$LL_{RR} = \exp(LL_{LogRR})$$

$$UL_{RR} = \exp(UL_{LogRR})$$

Where LL and UL referred to lower and upper confidence limits respectively.

An increase in the outcome for the intervention group, relative to the control group was indicated with an SMD of greater than zero for continuous outcomes, or an RR of greater than one for binary outcomes. Conversely, an SMD of less than zero, or an RR of less than one, indicated the outcome was lower for the intervention group, relative to the control group. An SMD that was not statistically significantly different to zero, or an RR that was not statistically significantly different to one, indicated no statistically significant impact. The interpretation of positive and negative impacts was dependent on the specific outcome considered.

Converting outcomes into a common metric

Where variables with different measurement scales measure the same outcome of interest for the same intervention, we converted binary outcomes to standardized mean differences using formulae by Borenstein and colleagues (2009, Chapter 7). By making effect sizes for continuous and dichotomous outcome variables comparable to each other, we ensured we can use studies with different measurement scales in the same analysis.

Using Table 1 as a reference for the nomenclature used, we calculated odds ratios, log odds ratios and their variance as follows:

$$OR = \frac{AD}{BC}$$

$$LogOR = \ln(OR)$$

$$V_{LogOR} = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}$$

We then converted the log odds ratio into the standardised mean difference (Cohen's d) using

$$d = LogOR \times \frac{\sqrt{3}}{\pi}$$

Where π was the mathematical constant, the variance of d was calculated using,

$$V_d = V_{LogOR} \times \frac{3}{\pi^2}$$

Partial effect sizes

In cases where studies use multivariate analysis to estimate effect sizes, we may need to extract and use partial effect-sizes in the meta-analysis (Keef & Roberts, 2004). Keef and Roberts considered the case of treatment effects based on continuous variables and proposed the use of a partial (or adjusted) standardised mean difference for this scenario. They examined the case of a two-group comparison analysed using an analysis of covariance model as follows,

$$Y_j = \alpha + \gamma D_j + \beta_2 X_{2j} + \cdots \beta_p X_{pj} + e_j$$

Where Y was an outcome score, D was a dummy variable representing a treatment or group effect, and X_2 to X_p were covariates. The errors e_j were assumed to have common variance σ_e^2.

Keef and Roberts proposed the following approach to calculating a partial index of treatment effects:

$$g_{adj} = \frac{\hat{\gamma}}{\hat{\sigma}_e}$$

Where $\hat{\gamma}$ represented an adjusted mean difference (accounting for all covariates in the model) and $\hat{\sigma}_e$ the residual variance (i.e., the variance of the Y scores, partialling out the effects of all predictors. In practice, we did not expect included studies to report $\hat{\sigma}_e^2$. Consequently, we proposed to collect the standard deviations of the outcome both before the adjustment (S_y) and after the adjustment ($\hat{\sigma}_e$). The decision on which standard deviation to standardising the partial effect sizes depended on what has been reported in most studies with common outcomes.

Partial effect sizes might be smaller or larger than zero-order (non-adjusted) effect sizes depending on the adjustments at play[6]. Nevertheless, for the sake of using valuable information provided by partial effect sizes in our synthesis, we combined both unadjusted and adjusted standardised mean differences from studies with common outcomes in our meta-analyses.

Should reporting of standard deviation information among included studies be limited, we proposed to calculate standardised mean difference d and variance of d using the following formulae instead:

$$d = \frac{2t}{\sqrt{n_t + n_c}}$$

[6] Adjusting for covariates can result in both an adjustment to the mean difference, as well as a reduction in the standard deviation. Consequently, g_{adj} can be either smaller or larger than the unadjusted effect size d. It can be larger than the unadjusted standardised mean difference if the adjusted and unadjusted mean differences do not differ much, but the adjusted standard deviation $\hat{\sigma}_e$ is much smaller than the unadjusted standard deviation (S_y).

$$V_d = \frac{2}{n_t + n_c} + \frac{d^2}{4(n_t + n_c)}$$

Where t referred to the t-statistic and n_t and $n_c$denoted the sample size of treatment and control group respectively.
For studies using multivariate analysis, we calculated the t-statistic (t) by dividing the coefficient by the standard error.

$$t = \frac{\hat{\gamma}}{\widehat{SE}_\gamma}$$

If the authors only reported confidence intervals and no standard error we calculated the standard error from the confidence intervals. If the study did not report the standard error but reports t, we extracted and used this as reported by the authors.
For studies reporting other data than coefficients and standard errors we used different formula to calculate d, as reported below:

Studies reporting mean differences ($\Delta\overline{X}$) between treatment (T) and control (C) and standard deviation (SD) at follow up (p+1):

$$d = \frac{\Delta\bar{X}_{p+1}}{SD_{p+1}} = \frac{\bar{X}_{Tp+1} - \bar{X}_{Cp+1}}{SD_{p+1}}$$

Studies reporting mean differences between treatment and control, standard error (SE) and sample size (n):

$$d = \frac{\Delta\bar{X}_{p+1}}{SE\sqrt{n}}$$

Studies reporting means and standard deviations for treatment and control groups at baseline (p) and follow up (p+1):

$$d = \frac{\Delta\bar{X}_p - \Delta\bar{X}_{p+1}}{SD_{p+1}}$$

where

$$SD_{p+1} = \sqrt{\frac{(n_{Tp+1} - 1)SD^2_{Tp+1} + (n_{Cp+1} - 1)SD^2_{Cp+1}}{n_{Tp+1} + n_{Cp+1} - 2}}$$

$$Vd = \frac{n_T + n_C}{n_T n_C} + \frac{d^2}{2(n_T + n_C)}$$

Finally, a few studies only provided exact p-values and sample sizes. For these studies we calculated d using the t-test p-value, unequal sample size formula provided in the Practical Meta-Analysis Effect Size Calculator (Wilson, n.d).

Where studies reported effect estimates as Cohen's d or Hedges' g and did not provide sufficient information for us to calculate these ourselves from other provided information, we used the effect sizes as reported by the authors.

Where statistical meta-analysis was not possible, results were reported narratively (see section 3.5).

Conversion of effect sizes to standardised percentages
Standardised percentages are presented alongside the standardised mean differences (effect sizes) to aid interpretation of the meta-analysis results. These are calculated using a "binomial effect size display (BESD)" (Randolph and Edmondson, 2006). The

standardised mean differences are first transformed into point-biserial correlation coefficients. They are then converted into success rates, where the base success rate is assumed to be 50%. The standardised percentage is the difference expected in an outcome expected as a result of the intervention.

The BESD is not equivalent to the translation of the standardised mean differences into the units presented in the individual studies themselves and are not the same as a percentage change in the raw data. The only purpose of the standardised percentages is to communicate the magnitude of the results more intuitively to the reader.

Unit of Analysis

If the unit of treatment assignment and the unit of analysis were different and clustering had not been accounted for in the analysis, the standard errors were adjusted using the formulas provided by Higgins and Green (2011), with the new effective sample size *Ne* for the standard error calculated as follows:

$$Ne = \frac{N}{1 + (m-1)c}$$

where *N* was total sample size, *m* was average number of observations per cluster, *c* was the intra-cluster correlation coefficient and Ne equals the revised sample size. We assumed a value of *c* equal to 0.15 unless an intra-cluster correlation coefficient reported in one of the other studies included in the meta-analysis was used as a guiding value.

After risk of bias assessment, we suspected two studies of unit of analysis errors, which we corrected for in our effect size calculations. These were Raskind et al. (2018) and Greenberg et al. (2010).

Missing Data

Where studies provided insufficient data to estimate the effect size, this information was sought directly from authors in the first instance. If the authors could not be contacted, or they did not provide the required data, where possible the effect size was extracted or imputed based on commonly reported statistics (Lipsey & Wilson, 2001).

Presentation of Effect Sizes

The heterogeneity of effect sizes was presented graphically using forest plots. Formal tests were also conducted to estimate the Q-statistic for heterogeneity and the I^2 and Tau^2 statistics were estimated to assess the variability in the distribution of the true effect sizes (Borenstein, et al., 2009).

Sensitivity Analysis

Sensitivity analysis was conducted according to risk of bias category and follow-up period, where data allowed.

Statistical Software

The meta-analysis was conducted using either Stata's *metan* command (Harris et al., 2008) in Stata 14.1 SE.

Publication Bias

We undertook tests to explore for the presence of publication bias through). The tandem procedure (Ferguson and Brannick, 2012) which comprised three criteria: Orwin's fail-safe N test; rank order correlation and Egger's regression tests; and Trim and Fill test.

1.6.2 Criteria for determination of independent findings

In order to facilitate the analysis and avoid double-counting of evidence, we endeavoured to link multiple papers that evaluate the same intervention. For the purposes of this protocol, the data related to a specific research project was referred to as a 'study'. Multiple 'papers' could report on different aspects of that study, for example by using different methods to analyse the same dataset relating to the same project or by exploring different outcomes, or by considering different time points.

Dependent Effect Sizes

Meta-analysis relies on the assumption of the independence of each included effect estimate (Gleser & Olkin, 2007). Dependent effect sizes may result from a variety of circumstances. For example, when there are multiple papers using a single dataset and reporting on the same outcome, when multiple results are reported for the same outcome, when outcomes of interest are measured at multiple points in time, or where there are multiple treatment arms compared to a common comparison group.

To address these concerns, we applied the following rules. We included a single effect estimate for each outcome from a given study in a meta-analysis. Where multiple papers report on the same intervention and outcome, we included the most recent publication. Where we identify studies that report multiple outcome constructs for a single outcome of interest, we selected the construct that was most similar to estimates for that outcome from other studies to be included in the meta-analysis. If we identified multiple papers that analyse the same intervention but explore outcomes for different sub-groups, we included all such estimates, provided they were measured in comparison with different control or comparison groups. For studies with multiple treatment arms and only a single control or comparison group, we either chose the estimate for the treatment arm most similar to other interventions included in the meta-analysis or created a 'synthetic effect'. Synthetic effect estimates were created using the sample-weighted average using the following formulae to recalculate variances (as per Borenstein 2009, chapter 24).

Where outcome measures at multiple time points were available for a single outcome, we identified the most common follow-up period for meta-analysis and reported other follow-ups narratively.

2 Critical Appraisal

The study team critically appraised the quality of all studies included in the review. The following sections summarise the critical appraisal tools used to assess studies, then an overview of the results of the critical appraisal, structured by domain of interest: Prevalence of mental health problems; Experiences of mental health problems; and Effectiveness of interventions to address mental health problems. Finally, we present study-by-study critical appraisal assessments.

2.1 Critical appraisal tools

2.1.1 Prevalence and Experience

We assessed the quality of studies included in the Prevalence and Experience domains using an adapted version of the Critical Appraisal Skills Programme (CASP, 2019) checklist for primary studies and an adapted version of the Specialist Unit for Review Evidence (SURE, 2018) checklist for evidence reviews and meta-analyses. The tools used are provided below. For each domain, papers were coded as 'Yes', 'No', 'Unclear' or 'No information' according to the degree to which it met the category description.

Primary Studies
The domain descriptions for primary studies included for prevalence and experience are listed below.

Table A 2.1 Prevalence Primary studies critical appraisal categories

Category	Description
1) **Research aim clearly stated**	Is the research aim clearly stated?
2) **Appropriate sampling strategy**	Was the sampling strategy appropriate to the aims of the research? Have the researchers explained how the participants were selected? Have the researchers explained why the participants they selected were the most appropriate to provide access to the type of knowledge sought by the study?
3) **Sample characteristics sufficiently reported**	Are sample characteristics sufficiently reported? (E.g. sample size, location, and at least one additional characteristic)
4) **Clear data collection**	Is it clear how the data were collected? (E.g. for interviews, is there an indication of how interviews were conducted? Are methods of data recording also reported?)
5) **Link to relevant literature/theory**	Is there a clear link to relevant literature/theoretical framework?

6) Sufficiently rigorous analysis	Was the data analysis sufficiently rigorous? Is there a detailed description of the analysis process? Do the data support the findings? To what extent are contradictory data taken into account? If the findings are based on quantitative analysis of survey data, are multivariate techniques used to control for potential confounding variables?

Table A 2.2 Experience Primary studies critical appraisal categories

Category	Description
1) Research aim clearly stated	Is the research aim clearly stated?
2) Sample characteristics sufficiently reported	Are sample characteristics sufficiently reported? (E.g. sample size, location, and at least one additional characteristic)
3) Clear data collection	Is it clear how the data were collected? (E.g. for interviews, is there an indication of how interviews were conducted? Are methods of data recording also reported?)
4) Link to relevant literature/theory	Is there a clear link to relevant literature/theoretical framework?
5) Sufficiently rigorous analysis	Was the data analysis sufficiently rigorous? Is there a detailed description of the analysis process? Do the data support the findings? To what extent are contradictory data taken into account? If the findings are based on quantitative analysis of survey data, are multivariate techniques used to control for potential confounding variables?

The category 'Appropriate sampling strategy' was part of the critical appraisal for the prevalence domain as most studies included for prevalence are large surveys.

Evidence reviews and meta-analyses

The domain descriptions for reviews included for prevalence and experience are listed below.

Table A 2.3 Prevalence and Experience reviews critical appraisal categories

Category	Description
1) Research aim clearly stated	Is the research aim clearly stated?

2) Inclusion criteria set out	Inclusion criteria is a list of criteria that evidence will be screened against to determine eligibility for being included in the review. Does it state and define the topic areas, populations, publication type/date, type of study design of interest to the research?
3) Search strategy clearly described	Is the way in which the researchers located the relevant literature explained? (including stating the search terms)
4) Clear which resources were searched	Does it state the databases and websites that were searched?
5) Quality check of included studies	A quality assessment is an appraisal of the methodological quality/rigour of the studies included in the review e.g. sample size, design. Does the study include a quality assessment?

2.2 Results of critical appraisal

2.2.1 Prevalence

Figure A 2.1 summarises the critical appraisal for primary studies included in the prevalence domain

Figure A 2.1 Prevalence Quality Appraisal for Primary studies

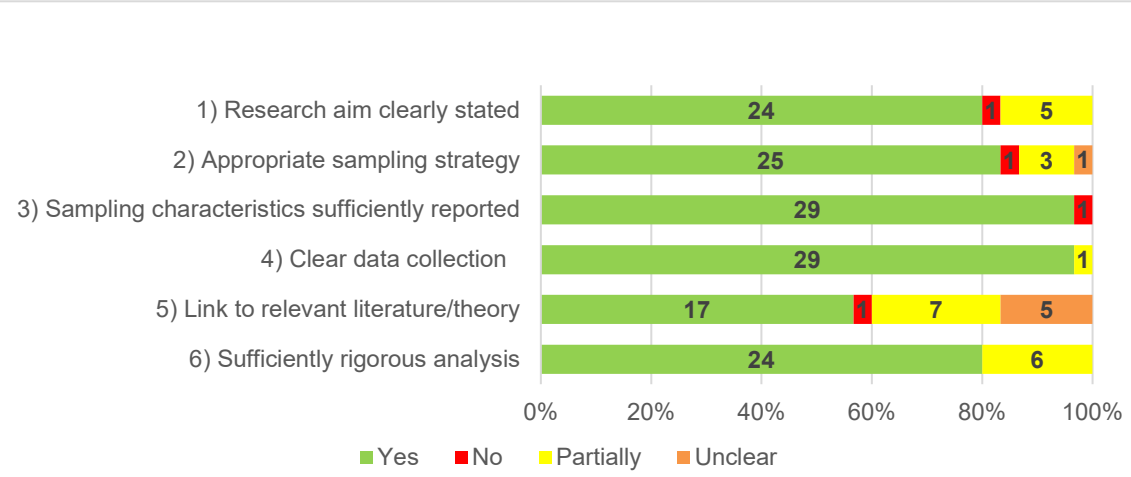
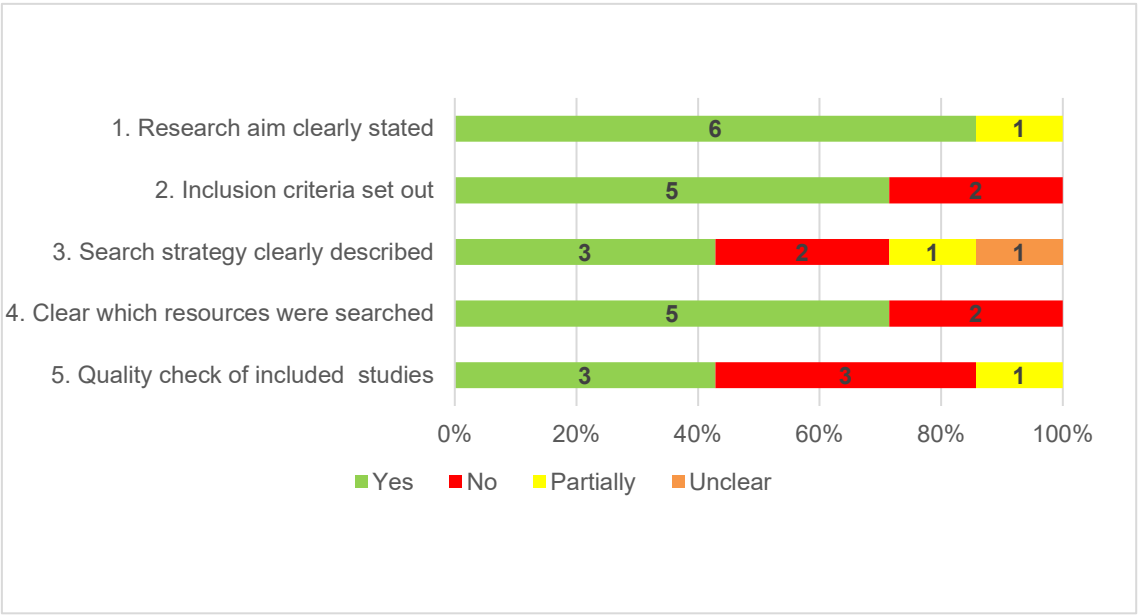


Figure A 2.2 summarises the critical appraisal for reviews included in the prevalence domain.

Figure A 2.2 Prevalence Quality Appraisal for Reviews



2.2.2 Experience

Figure A 2.3 summarises the critical appraisal for primary studies included in the experience domain.

Figure A 2.3 Experience Quality Appraisal for Primary studies

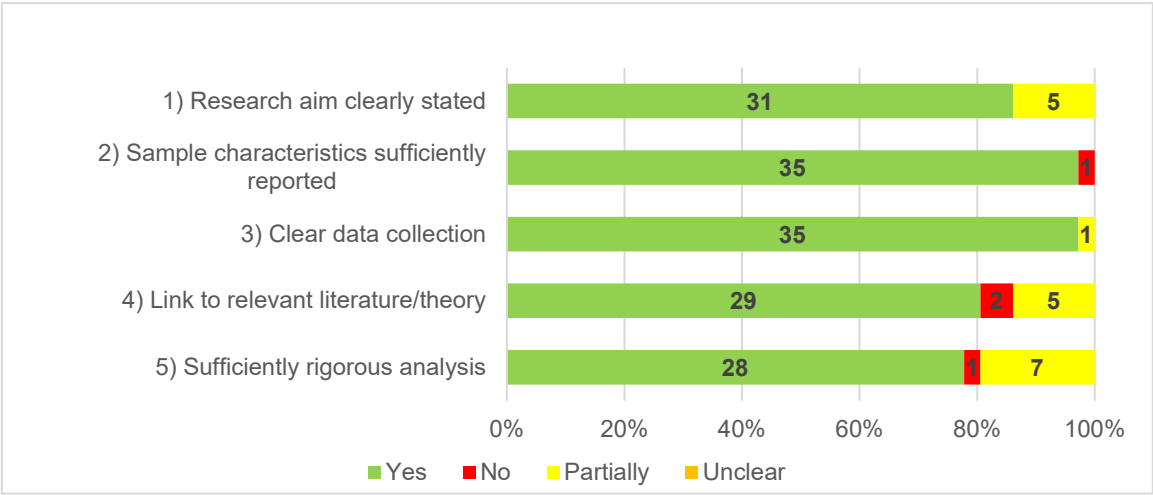
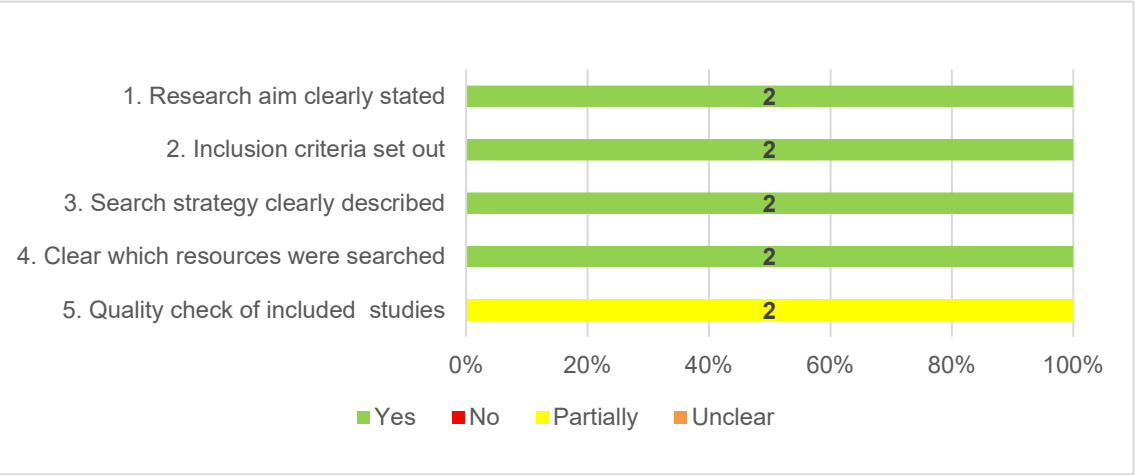


Figure A 2.4 summarises the critical appraisal for reviews included in the experience domain.

Figure A 2.4 Experience Quality Appraisal for Reviews



2.2.3 Effectiveness

Studies included in the effectiveness domain of the review were critically appraised using a risk of bias tool originally formulated by Hombrados and Waddington (2012) and since adapted by systematic reviews by Oya et al. (2017) and Baird et al. (2014).

Risk of bias assessment was undertaken at the paper level[7] by a single reviewer. A second reviewer checked overall risk of bias assessments.

For each domain, papers were coded as 'Yes', 'No', 'Unclear' or 'No information'.

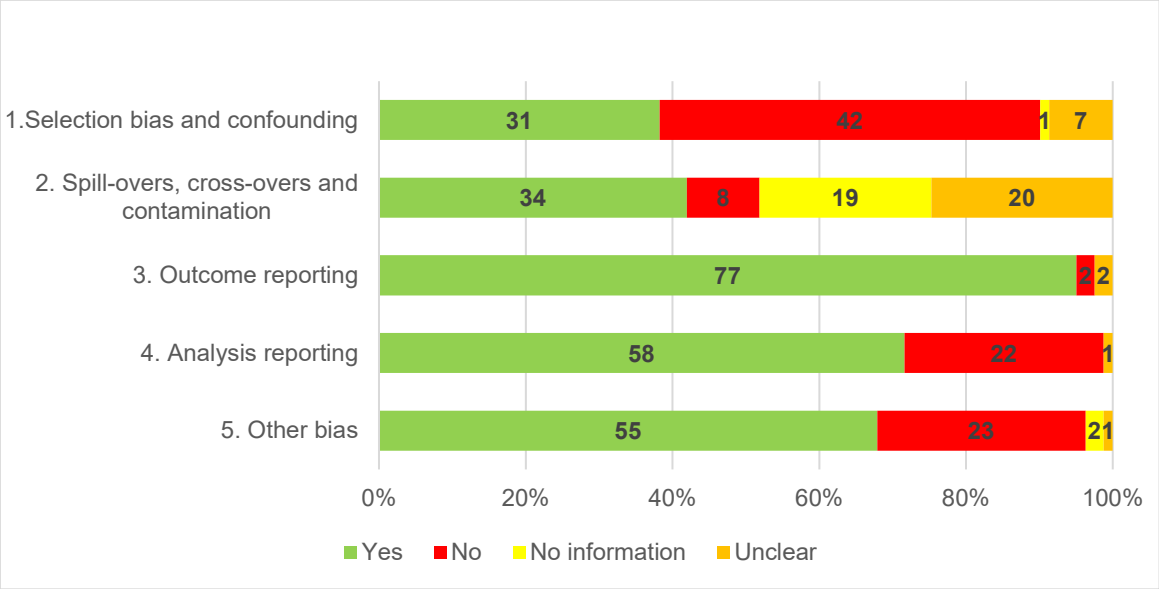
Overall risk of bias for each paper were determined as follows using a decision rule adapted from Baird et al. (2014):

Low risk of bias	'Yes' for four or five categories where a 'yes' indicates a particular bias has been adequately addressed
Medium risk of bias	'Yes' for three categories.
High risk of bias	'Yes' for two categories or fewer.

Figure A 2.5 summarises the risk of bias assessment for papers included in the effectiveness domain.

[7] This is because one paper reported on two separate studies.

Figure A 2.5 Effectiveness Risk of Bias appraisal



2.3 Study-by-study critical appraisal: Prevalence

The following section presents a study-by-study overview of the critical appraisal of studies included in the prevalence domain. Table A 2.4 is for included primary studies and Table A 2.5 for systematic reviews.

Table A 2.4 Prevalence primary studies - individual critical appraisals

Reference	1. Research aim clearly stated	2. Appropriate sampling	3. Sampling characteristics	4. Clear data collection	5. Link to literature/ theory	6. Sufficiently rigorous analysis
Aguirre et al., 2014a	Yes	Yes	Yes	Yes	Yes	Partially
Aguirre et al., 2014b	Partially	No	No	Partially	Partially	Partially
Ashwick et al., 2017	Yes	Yes	Yes	Yes	Unclear	Yes
Bennett, 2017	Yes	Yes	Yes	Yes	Yes	Yes
Bergman et al, 2017	Yes	Yes	Yes	Yes	Unclear	Yes
Cawkill et al., 2015	Yes	Yes	Yes	Yes	Unclear	Yes
Dighton et al, 2018	Yes	Yes	Yes	Yes	Yes	Yes
Goodwin et al., 2015	Yes	Yes	Yes	Yes	Yes	Yes

Reference	1. Research aim clearly stated	2. Appropriate sampling	3. Sampling characteristics	4. Clear data collection	5. Link to literature/ theory	6. Sufficiently rigorous analysis
Goodwin et al., 2017	Yes	Yes	Yes	Yes	Unclear	Yes
Harden and Murphy, 2018	Yes	Yes	Yes	Yes	Yes	Yes
Head et al., 2016	Yes	Yes	Yes	Yes	Yes	Yes
Hines et al., 2014a	Yes	Yes	Yes	Yes	Yes	Yes
Kwan et al., 2017	Yes	Yes	Yes	Yes	Yes	Yes
Kwan, 2016	Yes	Yes	Yes	Yes	Yes	Yes
MacManus, 2013	Yes	Yes	Yes	Yes	Partially	Yes
Ministry of Defence, 2016	Partially	Yes	Yes	Yes	Unclear	Yes
Ministry of Defence, 2018a	Partially	No	No	Yes	Partially	Yes
Ministry of Defence, 2018b	No	Unclear	Yes	Yes	No	Yes
Murphy and Turgoose, 2018b	Yes	Yes	Yes	Yes	Yes	Yes
Murphy et al., 2015	Partially	Partially	Yes	Yes	Yes	Partially
Murphy et al., 2016	Yes	Yes	Yes	Yes	Yes	Yes
RBL,2014	Partially	Yes	Yes	Yes	Yes	Yes
Short et al., 2018	Yes	Partially	Yes	Yes	Yes	Yes
Stevelink et al., 2015	Yes	Partially	Yes	Yes	Partially	Partially
Stevelink et al., 2018	Yes	Yes	Yes	Yes	Partially	Yes
Thandi et al., 2015a	Yes	Yes	Yes	Yes	Partially	Yes
Thandi et al., 2015b	Yes	Yes	Yes	Yes	Partially	Partially
Turgoose and Murphy, 2018b	Yes	Yes	Yes	Yes	Yes	Yes

Reference	1. Research aim clearly stated	2. Appropriate sampling	3. Sampling characteristics	4. Clear data collection	5. Link to literature/ theory	6. Sufficiently rigorous analysis
Whybrow et al., 2015	Yes	Yes	Yes	Yes	Yes	Partially
Woodhead, 2013	Yes	Yes	Yes	Yes	Yes	Yes

Table A 2.5 Prevalence reviews - individual critical appraisals

Reference	1. Research aim clearly stated	2. Inclusion criteria set out	3. Search strategy clearly described	4. Clear which resources searched	5. Quality check of included studies
Diehle et al., 2017	Yes	Yes	Yes	Yes	Yes
MacManus et al., 2015	Yes	Yes	Yes	Yes	Partially
Rona et al., 2016	Yes	Yes	Yes	Yes	Yes
Hines et al., 2014b	Yes	Yes	Yes	Yes	Yes
MacManus et al., 2014	Yes	Yes	Unclear	Yes	No
Murphy and Turgoose, 2018a	Yes	No	No	No	No
Turgoose and Murphy, 2018a	Yes	No	No	No	No
Williamson et al., 2018	Partially	Yes	Partially	Yes	Yes

2.4 Study-by-study critical appraisal: Experience

The following section presents a study-by-study overview of the critical appraisal of studies included in the prevalence domain. Table A 2.6 is for included primary studies and Table A 2.7 for systematic reviews.

Table A 2.6 Experience primary studies - individual critical appraisals

Reference	1. Research aim clearly state	2. Sampling characteristics	3. Clear data collection	4. Link to literature/ theory	5. Sufficiently rigorous analysis
Bull et al., 2015	Yes	Yes	Yes	Partially	Yes
Brian Parry Associates, 2015	Yes	Yes	Yes	Yes	Partially

Reference	1. Research aim clearly state	2. Sampling characteristics	3. Clear data collection	4. Link to literature/ theory	5. Sufficiently rigorous analysis
Caddick et al., 2015	Yes	Yes	Partially	Yes	Yes
De Rond,and Lok 2016	Yes	Yes	Yes	Yes	Yes
Dighton et al., 2018	Yes	Yes	Yes	Yes	Yes
Doncaster et al. 2015	Yes	Yes	Yes	Yes	Yes
Farrand et al., 2018	Yes	Yes	Yes	Yes	Yes
Fertout et al., 2015	Partially	Yes	Yes	Yes	Partially
Hallett, 2012	Yes	Yes	Yes	Yes	Partially
Hatch et al., 2013	Yes	Yes	Yes	Yes	Yes
Hatton, 2016	Yes	Yes	Yes	Yes	Yes
Hines et al., 2014	Yes	Yes	Yes	Yes	Yes
Hunt, 2016	Partially	Yes	Yes	Yes	Yes
Jones and Coetzee, 2018	Yes	Yes	Yes	Yes	Yes
Jones et al., 2013	Yes	Yes	Yes	Yes	Yes
Jones et al., 2015	Yes	Yes	Yes	Yes	Yes
Jones et al., 2018	Yes	Yes	Yes	Yes	Yes
Keeling et al., 2017	Yes	Yes	Yes	Partially	Yes
Kiernan et al., 2018	Yes	Yes	Yes	Yes	Yes
Lovatt, 2017	Yes	Yes	Yes	Yes	Yes
Mellotte et al., 2017	Yes	Yes	Yes	Yes	Yes
Murphy and Palmer et al., 2017	Yes	Yes	Yes	Partially	Partially
Murphy et al., 2014	Yes	Yes	Yes	Yes	Yes
Murphy et al., 2016	Yes	Yes	Yes	Yes	Yes
Murphy et al., 2017	Yes	Yes	Yes	Yes	Partially

Reference	1. Research aim clearly state	2. Sampling characteristics	3. Clear data collection	4. Link to literature/ theory	5. Sufficiently rigorous analysis
NHS, 2016	Partially	Partially	Yes	No	Partially
Northern Hub for Veterans and Military Families, 2017	Yes	Yes	Yes	Yes	Yes
Palmer et al., 2016	Yes	Yes	Yes	Yes	Partially
Rafferty et al., 2017	Yes	Yes	Yes	Yes	Yes
Rowe et al., 2013	Yes	Yes	Yes	Partially	Yes
Rowe et al., 2014	Yes	Yes	Yes	Yes	Yes
Sharp, 2015	Yes	Yes	Yes	Yes	Yes
Stevelink and Fear, 2016	Yes	Yes	Yes	No	No
Turgoose et al., 2018	Partially	Yes	Yes	Partially	Yes
Wainwright et al., 2016	Partially	Yes	Yes	Yes	Yes
Woodhead, 2013	Yes	Yes	Yes	Yes	Yes

Table A 2.7 Experience reviews - individual critical appraisals

Reference	1) Research aim clearly stated	2) Inclusion criteria set out	3) Search strategy clearly described	4) Clear which resources were searched	5) Quality check of included studies
Kantor et al., 2017	Yes	Yes	Yes	Yes	Partially
Ramchand, 2015	Yes	Yes	Yes	Yes	Partially

2.5 Study-by-study critical appraisal: Effectiveness

Table A 2.8 presents a study-by-study overview of the risk of bias of primary studies included in the effectiveness domain.

Table A 2.8 Effectiveness individual Risk of Bias assessments

References	1. Selection bias and confounding	2. Spill-overs, cross-overs, contamination	3. Outcome reporting	4. Analysis reporting	5. Other bias
Acierno et al., 2017	Yes	No information	Yes	Yes	Yes
Badura-Brack et al., 2015	No	No information	Yes	Yes	Yes
Batki et al., 2014	Yes	No	Yes	Yes	No
Bormann et al., 2013	Yes	No information	Yes	Yes	No
Bourque et al., 2015	Unclear	Yes	Yes	Yes	Yes
Bremner et al., 2017	Unclear	Yes	Yes	No	Yes
Brief et al., 2013	Yes	Yes	Yes	Yes	No information
Brown, 2013	Unclear	Unclear	Yes	No	Yes
Buffington et al., 2016	No	No	Yes	No	No
Carter et al., 2013	Unclear	No information	Yes	Yes	No
Castillo et al., 2016	No	Yes	Yes	No	Yes
Chinman et al., 2013	Yes	Yes	Yes	Yes	No
Church et al., 2013	Yes	Yes	Yes	Yes	Yes
Davis et al., 2018	Yes	Yes	Yes	Unclear	Yes
Dretsch et al., 2014	No	Yes	Yes	Yes	Yes
Egede et al., 2015	No	Yes	Yes	Yes	No
Eisen et al., 2012	No	Yes	No	No	Unclear
Engel et al., 2014	Yes	Yes	Yes	Yes	Yes
Engel et al., 2016	Yes	Yes	Yes	Yes	Yes

Foa et al., 2018	Yes	No information	Yes	Yes	No
Fortney et al., 2015	Yes	Yes	Yes	Yes	Yes
Gelkopf et al., 2013	No	Yes	Yes	Yes	No
Geronilla et al., 2014	No	Unclear	Yes	Yes	Yes
Gmel et al., 2013	No	Unclear	Yes	No	Yes
Gray et al., 2017	No	Yes	Yes	No	Yes
Greenberg et al., 2010	No	Yes	Yes	Yes	No
Harris et al., 2015	Unclear	Unclear	Yes	Yes	Yes
Hobfoll et al., 2016	Yes	Yes	Yes	Yes	Yes
Johnson et al., 2018	No	Yes	Yes	No	Yes
Jones et al., 2013	Unclear	Unclear	Yes	Yes	Yes
Kahn et al., 2016	Yes	Unclear	Yes	Yes	Yes
Kearney et al., 2013	Yes	Unclear	Yes	No	Yes
Kearney et al., 2016	No	No	Yes	Yes	Yes
Kilbourne et al., 2014	Yes	Yes	Yes	Yes	No information
Kip et al., 2013	No	Unclear	Yes	Yes	Yes
Krupnick et al., 2017	No	Unclear	Yes	No	Yes
Kuckertz et al., 2014	No	Unclear	Yes	No	Yes
LaCroix et al., 2018	No	No information	Yes	Yes	Yes
Luxton et al., 2016	No	No	Yes	No	Yes
Mack, 2013	No	Unclear	Yes	Yes	Yes
Mackintosh et al., 2017	Yes	Unclear	Yes	Yes	Yes
Maguen et al., 2017	Yes	Yes	Yes	Yes	No
Maieritsch et al.,2016	No	No	Yes	No	Yes

Margolies et al., 2013	No	No information	Yes	Yes	Yes
Martens et al., 2015	No	Unclear	Yes	No	Yes
McDevitt-Murphy et al., 2014	Yes	No information	Yes	Yes	Yes
McLay et al., 2017	No	Yes	Yes	Yes	Yes
Mithoefer et al., 2018	No	No	Yes	Yes	No
Moriarty et al., 2016	Yes	Yes	Yes	Yes	No
Morland et al., 2014	No	Unclear	Yes	Yes	Yes
Mulligan et al., 2011	Yes	Yes	Yes	Yes	Yes
Nacasch et al., 2015	No	No information	Yes	Yes	Yes
Oman and Bormann, 2015	Yes	Unclear	Yes	Yes	No
Oslin et al., 2014	Yes	No information	Yes	Yes	Yes
Pedersen et al., 2017	Yes	Yes	Yes	Yes	Yes
Polusny et al., 2015	No	Unclear	Yes	No	Yes
Possemato et al., 2016	Yes	Unclear	Yes	Yes	No
Raskind et al., 2018	No	No	Yes	No	No
Rauch et al., 2015	No	Unclear	Yes	No	Yes
Reger et al., 2016	Yes	Yes	Yes	Yes	No
Resick et al., 2015	Yes	Yes	Yes	Yes	Yes
Resick et al., 2017	No	Yes	Yes	Yes	Yes
Rona et al., 2017	No	Yes	Yes	Yes	No
Rosen et al., 2013	Yes	No information	Yes	Yes	Yes
Rosen et al., 2017	Unclear	No information	Yes	Yes	Yes
Sautter et al., 2015	Yes	No information	Yes	Yes	No
Sayer et al., 2015	No	No information	Yes	Yes	Yes
Seppälä et al., 2014	Yes	No information	Yes	Yes	Yes

Shea et al., 2013	No	No information	Yes	Yes	No
Shipherdet al., 2016	No	Unclear	Yes	Yes	Yes
Surís et al., 2013	Yes	Yes	Yes	Yes	Yes
Tuerk et al., 2018	No	Yes	Yes	Yes	Yes
Tylee et al., 2017	No	No	Yes	No	Yes
Valenstein et al., 2015	Yes	No information	Yes	Yes	No
Verduin et al., 2013	No	No information	Yes	No	Yes
Wahbeh, 2017	No information	Unclear	Yes	Yes	Yes
Wesemann, 2016	No	Yes	No	No	No
Wolf, 2016	No	No information	Yes	Yes	Yes
Yehuda et al., 2014	No	Yes	Yes	No	Yes
Yuen et al., 2015	No	Yes	Unclear	Yes	No
Ziemba, 2014	No	Yes	Unclear	No	No

3 Evidence Map Methodology

3.1 Search and inclusion

Studies included in the evidence map were sourced via the review's main search and screening process. In addition to the main review search, we also undertook a 'horizon scanning' exercise to search for any relevant ongoing research. We searched a range of search registries such as ClinicalTrials.gov, NIHR, the ISRCTN and WHO, in addition to Google and Google Scholar. We searched these using the following terms: "mental health" "protocol" ("armed forces" OR military OR army OR navy OR "air force" OR airforce OR veterans). The horizon scanning exercise took place during April 2019. The inclusion criteria for the review's Effectiveness domain (see Table A 1.4 of this Technical Annex) were used to determine whether studies would be included in the map.

3.2 Map framework

The map's framework of interventions and outcomes categories were drawn up using a combination of the following approaches:

- A review of the literature including previous systematic reviews and key sources such as the ICD-11 categorisation and the UK Council for Psychotherapy (UKCP) website;
- An examination of the interventions and outcomes described in studies we found via our search strategy;
- A consultation with experts in the field, including the review advisory group.

3.3 Map visualisation

The map includes completed primary studies, ongoing primary studies, and evidence reviews and meta-analyses. The primary studies were mapped onto this framework based on the type of intervention they evaluate and the outcomes they report on (or plan to report on in the case of ongoing studies). If a single study reports on multiple outcomes, it may be represented at multiple map intersections. A similar process was undertaken for evidence reviews and meta-analyses. Synthesis studies were mapped according to their reported scope, so if they were designed to report on multiple interventions and outcomes, they were represented at all such map intersections.

The final map was created using Adobe Design and Adobe Illustrator.

4 Stakeholder Interview Methodology

Interviews with key stakeholders were conducted after completion of the systematic review, with the aim of drawing on participants' expert knowledge of the area, gaining additional perspectives on the extent to which the research has addressed the important issues in each area, and contributing to the prioritisation of areas for future research. Interviewees were provided with a summary of some of the draft findings and were given the opportunity to review the evidence map prior to the interviews to facilitate an informed discussion.

4.1 Stakeholder interviews

4.1.1 Overview

We conducted depth interviews with 15 key stakeholders from a range of relevant fields, allowing us to incorporate the views of a variety of individuals with relevant expertise in military mental health. The interviews lasted for approximately one hour and were conducted over the telephone to ensure that stakeholders from across the UK were able to take part around their existing schedules and commitments.

Participants were provided with a summary of the emerging findings of the review in advance of the interviews, which allowed us to explore their views on the evidence base, such as breadth of coverage, quality, comprehensiveness and the key gaps. Due to the scope of the review, it was not feasible to cover all three domains (prevalence, experience and effectiveness) in each interview, whilst also retaining depth of discussion. The focus the interviews was therefore on either the prevalence and experience findings (10 interviews) or the evidence map (5 interviews).

4.1.2 Focus of the interviews

The main aim of the stakeholder interviews was to draw on the expert knowledge of people working in the field, whether based in health service provision and policy or academic research. The secondary aims of the interviews were as follows:

- To sense-check the findings from the systematic review, outlining breadth of coverage, depth and quality of evidence, comprehensiveness of the literature covered.
- To gauge key perspectives on issues pertaining to mental health service provision for service personnel
    - Views on current service provision
    - Gaps in current provision and suggestions for improvement
- To discuss existing ways of data collection on mental health of service personnel
    - How data is currently collected
    - Views on the current data gaps
    - How data collection can be improved
- To identify priority area for future research, building on the topics outlined by stakeholders in the previous review published in 2013

The focus of the 'sense check' differed according to the topic being discussed. For the interviews on the findings of the prevalence and experience domains, the discussion focused on their views on the evidence, in relation to:

- How the findings compare with their knowledge;
- Their thoughts on the factors that can explain the findings; and
- Anything the review does not cover that they consider important.

The interviews on the effectiveness domain explored the patterns of evidence displayed in the map, in relation to interventions and outcomes, as well as the spread of evidence by population and country. Stakeholders were also asked to comment on how the map might be used in their field, and how useful it might be in policy and research.

All of the interviews touched on the pressing issues of mental health service provision, as this is an important part of the overall picture of military mental health. This included stakeholder views on what is working well and less well in current service provision; the direction that services are moving in at present and views on what improvements are needed. All interviews also explored stakeholder views on current research on the mental health needs of serving and ex-Service personnel and their families, such as data collection, methodology and coverage of different topics.

4.2 Sampling and recruitment

The stakeholder group included members of the Review Advisory Group as well as representatives from a range of relevant organisations, such as the MOD, NHS, military charities and academic institutions. Participants were recruited through the networks of FiMT, Advisory Group members and Professor David Denney. A purposive sampling approach was used to achieve a diverse sample of representatives from a range of organisations and with different research interests.

4.3 Analysis

The interviews were audio recorded with permission from participants, to ensure an accurate record of the discussion. Audio recordings were used to take detailed fieldnotes that captured the key ideas that emerged from the interviews. This approach to data management and analysis ensured a balance between capturing the full depth and breadth of the data, whilst being able to collect and analyse the interview responses over a short period of time.

The notes were then analysed in a framework that reflected the aims of the interview and the wider research questions of the review. The analytical framework was drawn up by the research team after completion of the interviews to ensure it reflected the key issues emerging from the interviews.

The process of 'charting' the interview data using the analytical framework enabled researchers to analyse the data both between cases (looking at what different stakeholders said on the same issue) and within cases (looking at how an individual's or group's opinions on one topic relate to their views on another). Through reviewing the charts, the full range of views were mapped, and the accounts of different participants, or groups, compared and contrasted.

5 Effectiveness findings: additional reporting

Terminology:

This glossary is provided as a short summary of some technical terminology used in the effectiveness chapter. Full details on methodology are in chapter 1 of this Technical Annex.

Confidence intervals (CIs): For a given statistic calculated for a sample of observations (such as the standardised mean differences in our forest plots), confidence intervals represent the range of values around that statistic that are understood to contain, with a certain probability, the true value of that statistic.

Forest plot: The forest plots we present in this chapter are graphical representations of the average effect of an intervention on a given outcome, calculated from a meta-analysis of a number of independent studies evaluating the same type of intervention.

Meta-analysis: Statistical meta-analysis is the combination of data from a number of independent studies of the same topic, in order to synthesise their results and determine overall trends.

Propensity Score Matching (PSM): This is a form of evaluation intended to estimate the causal impact of an intervention. A simple comparison between a group receiving an intervention and another group who do not receive it may be biased as the two groups may have different characteristics. PSM aims to reduce this bias by accounting for the characteristics that might predict receiving the intervention.

Publication bias: If the publication or non-publication of studies depends on the nature and direction of the results, this can bias the results of the meta-analysis.

Randomised Controlled Trial (RCT): An RCT is a type of experiment that involves randomly assigning subjects to two or more groups. One group receives an intervention, while the other(s) receive an alternative or no intervention. Outcomes are then measured for each group and compared (Higgins and Green, 2011).

Statistical significance: This is measure of the probability of making a false positive error – i.e. the probability of falsely concluding that an intervention has had an effect when there is no real effect and the results have occurred by chance. Where results are reported as being 'not statistically significant', this means that statistical evidence does not meet the threshold set by the evaluator to conclude that the true impact is non-zero.

Standardised mean difference (SMD): This is a measure of the effect of an intervention of interest. SMDs are used in meta-analysis when a set of studies all assess the same outcome but measure it in a variety of ways.

Standardised percentages: Throughout this chapter we convert SMDs into percentage changes to aid interpretation of the results of the meta-analysis. These percentage changes are statistical constructs that rely on various assumptions. They

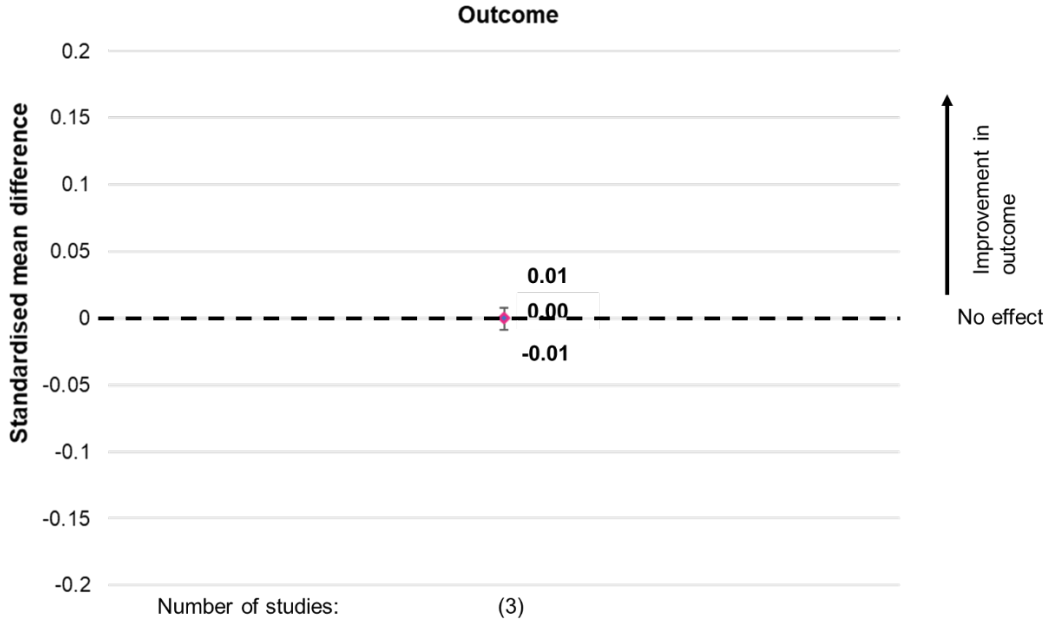
are presented only to convey a more intuitive measure of the size of reported effects and should be interpreted cautiously.

Treatment-as-usual: In studies with a 'treatment-as-usual' control group, this group receives an alternative intervention (typically current best practice), rather than no treatment. This means that the reported results are an assessment of the relative effectiveness of the intervention against another intervention.

5.1 The effectiveness of interventions to address mental health issues: tables and pooled forest plots

This section provides the underlying tables from the meta-analysis explored in Chapter 7 of the main report. This includes the standardised mean differences additional statistics and forest plots.

Figure A 5.1 Effectiveness of awareness, screening and prevention interventions



Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

How to interpret forest plots

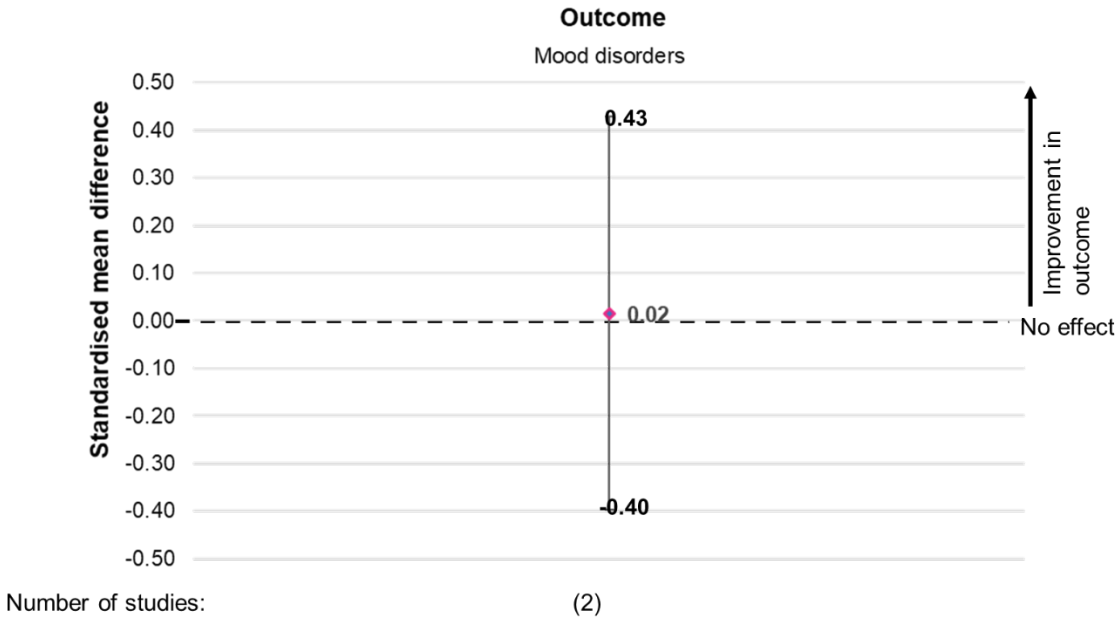
The forest plots in this chapter are graphical representations of the average effect of an intervention on a given outcome, calculated from a meta-analysis of several independent studies evaluating the same type of intervention. Each point estimate (depicted as a diamond) indicates the results of one meta-analysis, showing the average pooled effect (or standardised mean difference, SMD) of an intervention on a specific outcome. The line passing through the diamond indicates the 95% confidence intervals associated with each average pooled effect. Where vertical lines do *not* intersect with the dotted line of no effect, results are statistically significant.

Table A 5.1 The effectiveness of awareness, screening and prevention interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Stress and associated disorders including PTSD	-0.19 [-0.65, 0.27]	2	4.49 [0.034]	77.7	0.090	4.7

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.2 Effectiveness of Behavioural Activation Therapy



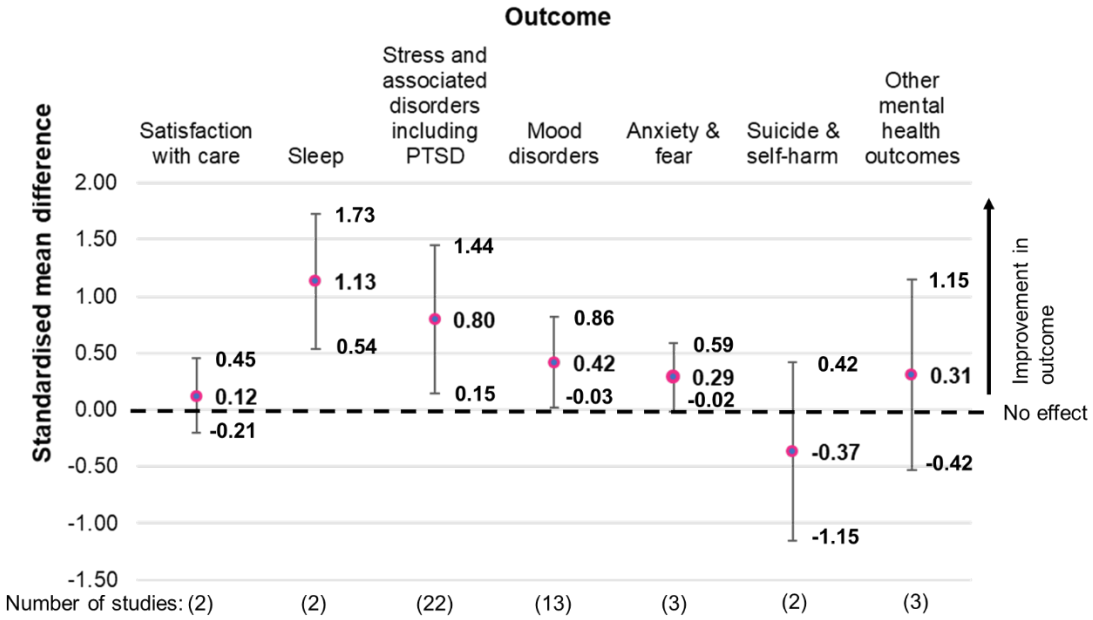
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.2 The effectiveness of Behavioural Activation Therapy

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Mood disorders	0.02 [-0.40, 0.43]	2	0.00 [0.97]	0.0	0.000	0.4

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.3 Effectiveness of CBT interventions



Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

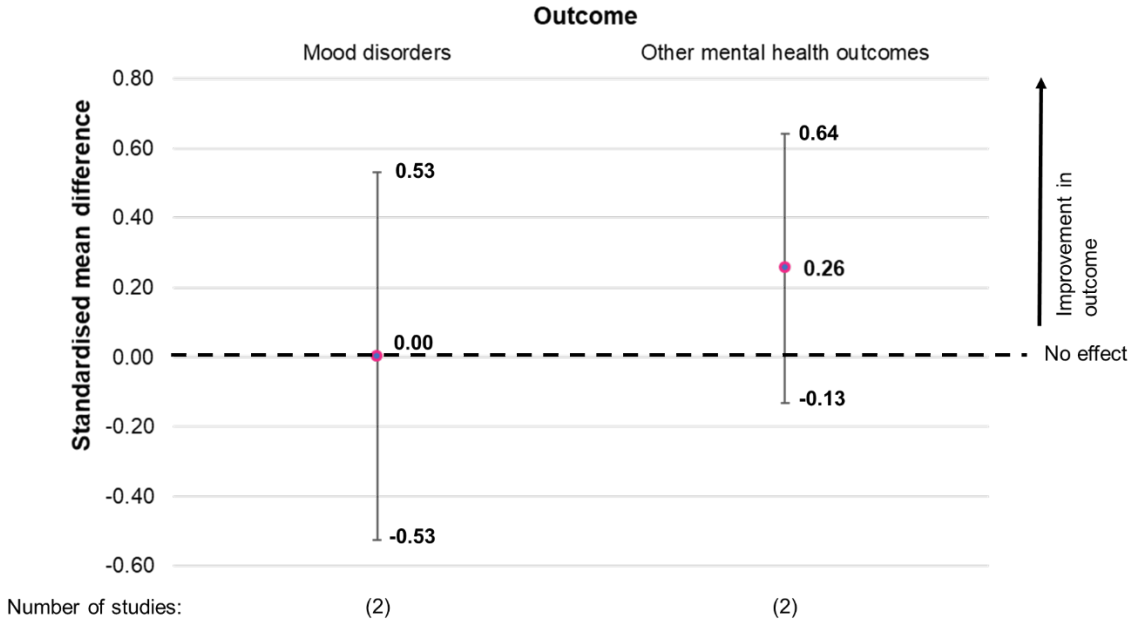
Table A 5.3 The effectiveness of CBT interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Satisfaction with care	0.12 [-0.21, 0.45]	2	1.25 [0.263]	20.3	0.013	3.0
Physical health & wellbeing (sleep)	1.13 [0.54, 1.73]	2	0.20 [0.653]	0.0	0.000	24.6
Stress and associated disorders including PTSD	0.80 [0.15, 1.44]	23	299.27 [0.000]	92.6	1.704	18.5
Mood disorders	0.42 [-0.03, 0.86]	13	48.22 [0.000]	75.1	0.364	10.3
Anxiety and fear	0.29 [-0.02, 0.59]	3	0.44 [0.804]	0.0	0.000	7.1

Suicide and self-harm	-0.37 [-0.42, 1.15]	2	1.99 [0.158]	49.9	0.194	9.1
Other mental health outcomes	0.31 [0.53, 1.15]	3	7.97 [0.019]	74.9	0.019	7.6

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.4 Effectiveness of family therapy interventions



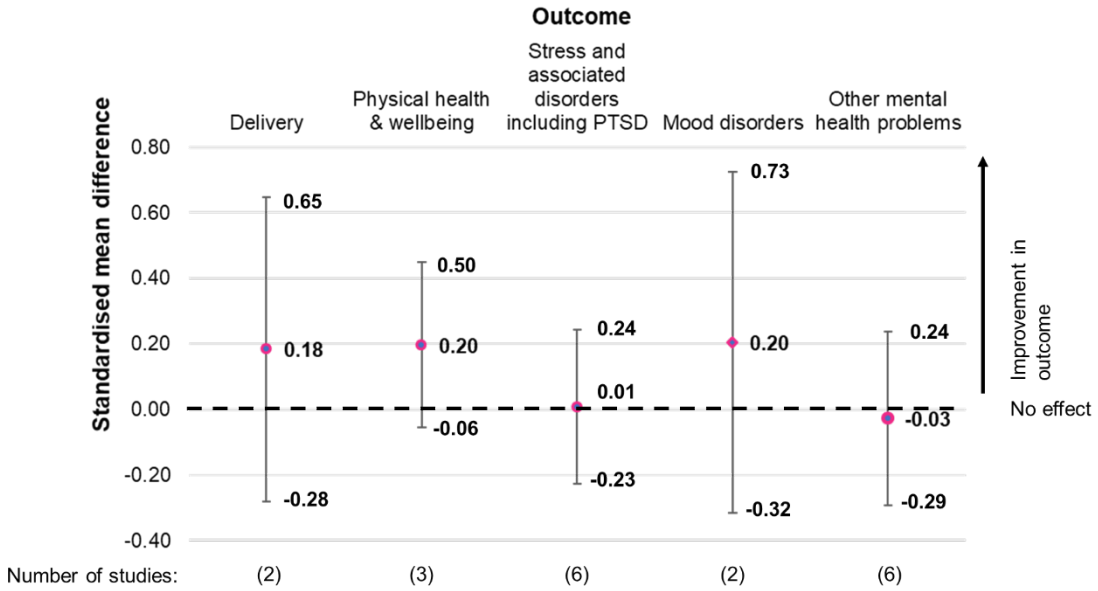
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.4 The effectiveness of Family Therapy interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Mood disorders	0.00 [-0.53, 0.53]	2	1.82 [0.178]	45.0	0.066	0.0
Other mental health outcomes	0.26 [-0.13, 0.64]	2	0.13 [0.715]	0.0	0.000	6.4

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.5 Effectiveness of wellbeing interventions



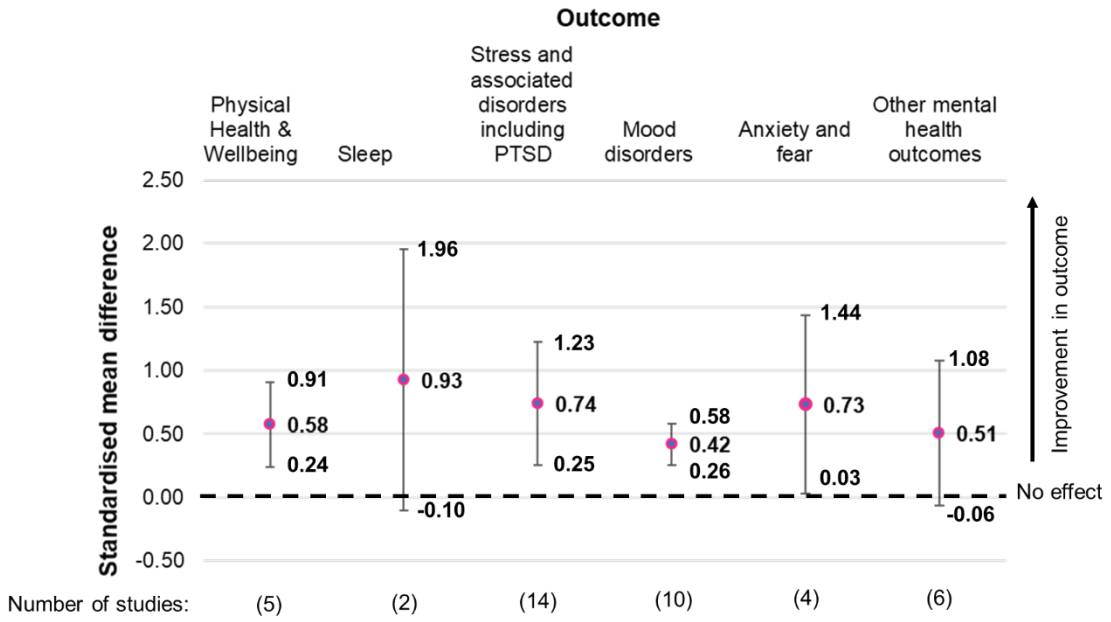
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.5 The effectiveness of wellbeing interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Delivery	0.18 [-0.28, 0.65]	2	0.26 [0.612]	0.0	0.000	4.6
Physical health and wellbeing	0.20 [-0.06, 0.50]	3	6.80 [0.033]	70.6	0.031	4.9
Stress and associated disorders including PTSD	0.01 [-0.23, 0.24]	6	12.25 [0.031]	59.2	0.038	0.2
Mood disorders	0.20 [-0.32, 0.73]	2	1.67 [0.196]	40.1	0.057	5.1
Other mental health outcomes	-0.03 [-0.29, 0.24]	6	15.45 [0.001]	67.6	0.055	-0.7

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.6 Effectiveness of meditation and mindfulness interventions



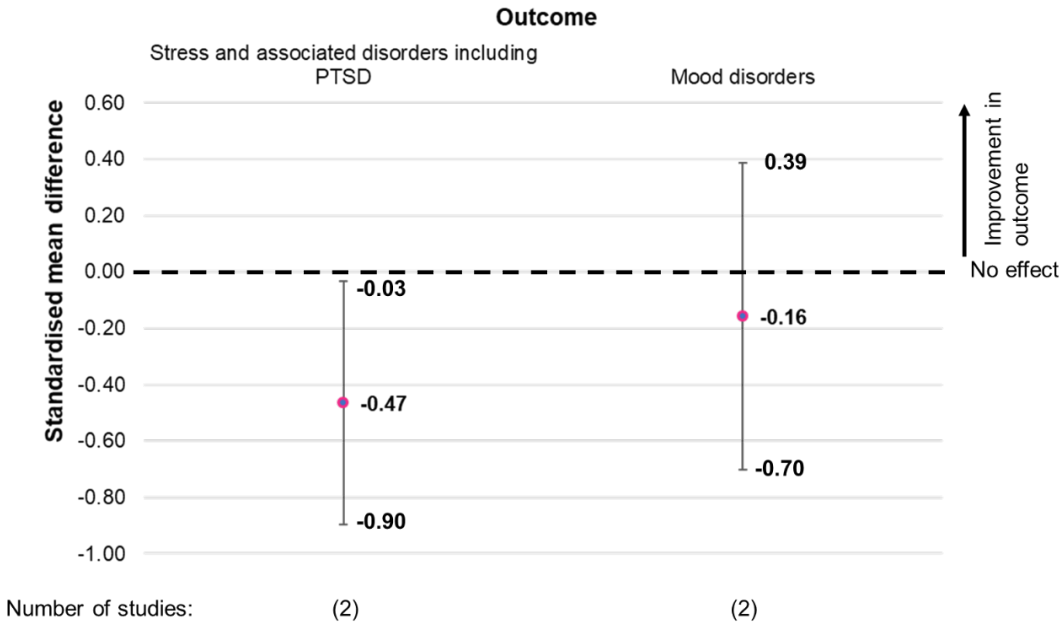
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.6 The effectiveness of meditation and mindfulness interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Physical health & wellbeing	0.58 [0.24, 0.91]	5	6.58 [0.160]	39.2	0.056	13.8
Sleep	0.93 [-0.10, 1.96]	2	4.97 [0.026]	79.9	0.441	21.0
Stress and associated disorders including PTSD	0.74 [0.25, 1.23]	14	126.58 [0.000]	89.7	0.627	17.4
Mood disorders	0.42 [0.26, 0.58]	10	6.17 [0.723]	0.0	0.000	10.3
Anxiety and fear	0.73 [0.03, 1.44]	4	18.36 [0.000]	83.7	0.418	17.2
Other mental health outcomes	0.51 [0.06, 1.08]	6	25.76 [0.000]	80.6	0.407	12.3

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.7 Effectiveness of other therapeutic wellbeing interventions



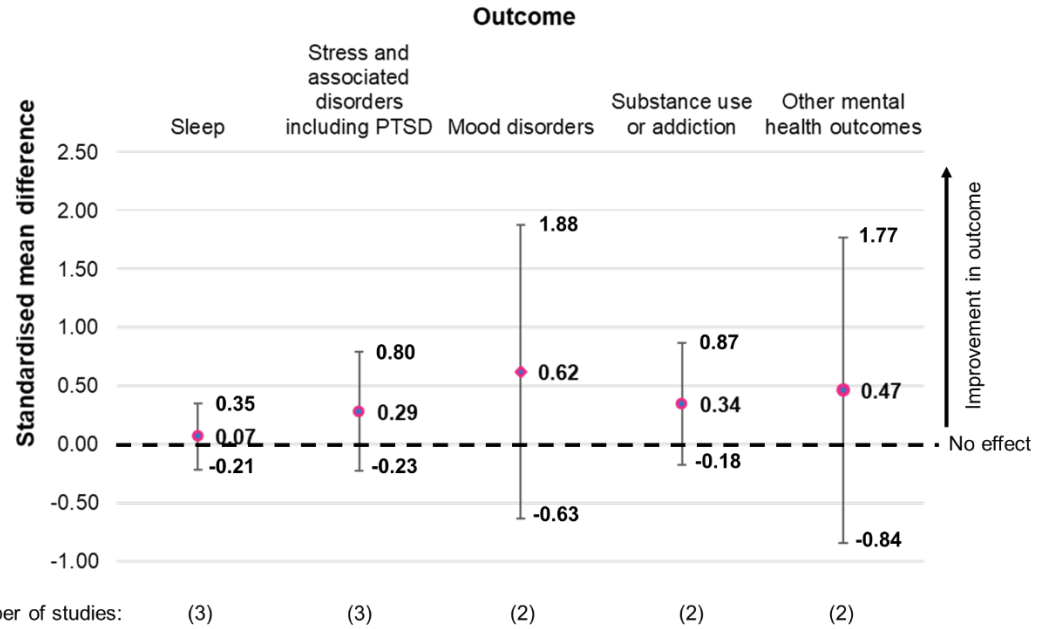
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.7 The effectiveness of other therapeutic wellbeing interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Stress and associated disorders including PTSD	-0.47 [-0.90, -0.03]	2	1.18 [0.278]	14.9	0.018	-11.3
Mood disorders	-0.16 [-0.70, 0.39]	2	2.30 [0.130]	56.5	0.104	-3.9

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.8 Effectiveness of medication interventions



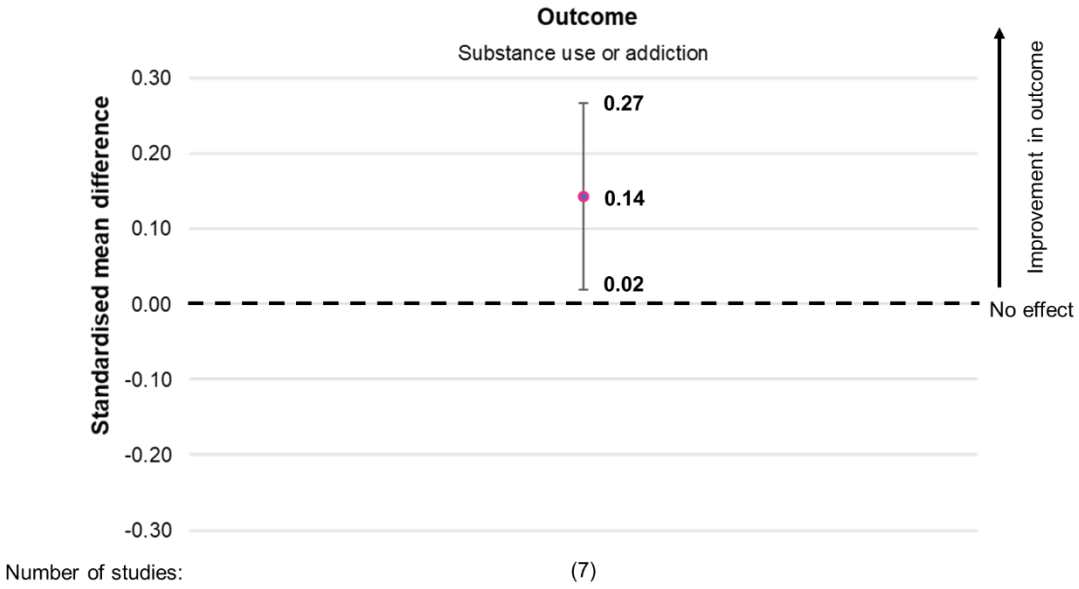
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.8 The effectiveness of medication interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Sleep	0.07 [-0.21, 0.35]	3	1.17 [0.558]	0.0	0.000	1.7
Stress and associated disorders including PTSD	0.29 [-0.23, 0.80]	3	3.31 [0.191]	39.6	0.086	7.1
Mood disorders	0.62 [-0.63, 1.88]	2	4.23 [0.040]	76.3	0.652	14.9
Substance use or addiction	0.34 [-0.18, 0.87]	2	1.62 [0.204]	38.1	0.063	8.5
Other mental health outcomes	0.47 [-0.84, 1.77]	2	5.12 [0.024]	80.5	0.731	11.4

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.9 Effectiveness of interventions for substance misuse and gambling interventions



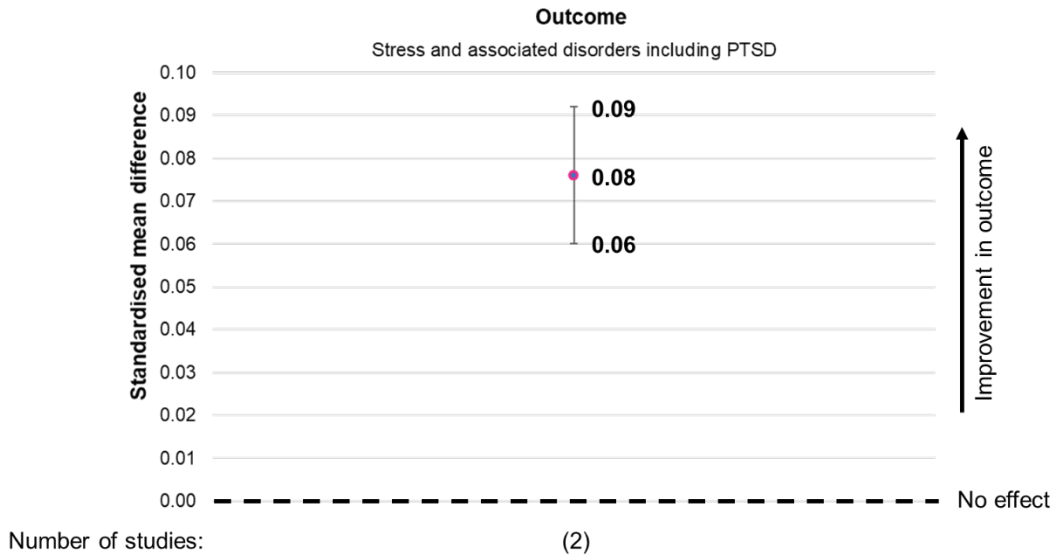
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.9 The effectiveness of substance misuse and gambling interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Substance use or addiction	0.14 [0.02. 0.27]	7	9.60 [0.143]	37.5	0.009	3.6

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.10 Effectiveness of resettlement interventions



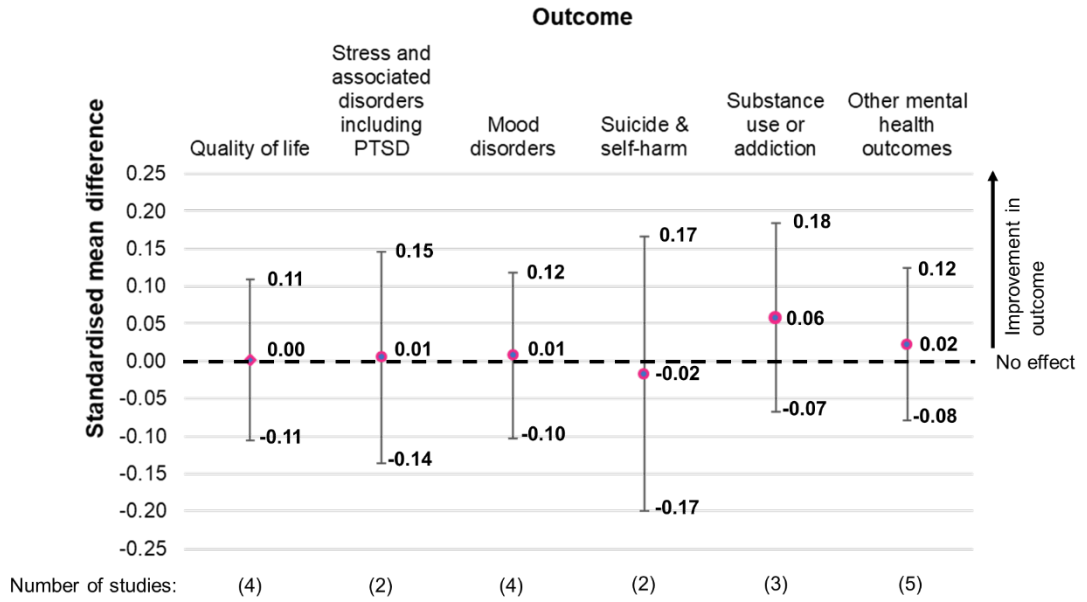
Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.10 The effectiveness of resettlement interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Stress and associated disorders including PTSD	0.08 [0.06, 0.09]	2	0.49 [0.486]	0.0	0.000	1.9

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Figure A 5.11 Effectiveness of advice and support interventions



Note that the sign of standardised mean differences may be switched in forest plots for comparability across outcomes.

Table A 5.11 The effectiveness of advice and support interventions

Outcome	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Quality of life	0.00 [-0.11, 0.11]	4	0.72 [0.868]	0.0	0.000	0.0
Stress and associated disorders including PTSD	0.01 [-0.14, 0.15]	2	0.84 [0.360]	0.0	0.000	0.1
Mood disorders	0.01 [-0.10, 0.12]	4	2.01 [0.570]	0.0	0.000	0.2
Suicide and self-harm	-0.02 [-0.20, 0.17]	2	0.02 [0.897]	0.0	0.000	-0.4
Substance use or addiction	0.06 [-0.07, 0.18]	3	0.00 [0.998]	0.0	0.000	1.4
Other mental health outcomes	0.02 [-0.08, 0.12]	5	3.04 [0.551]	0.0	0.000	0.5

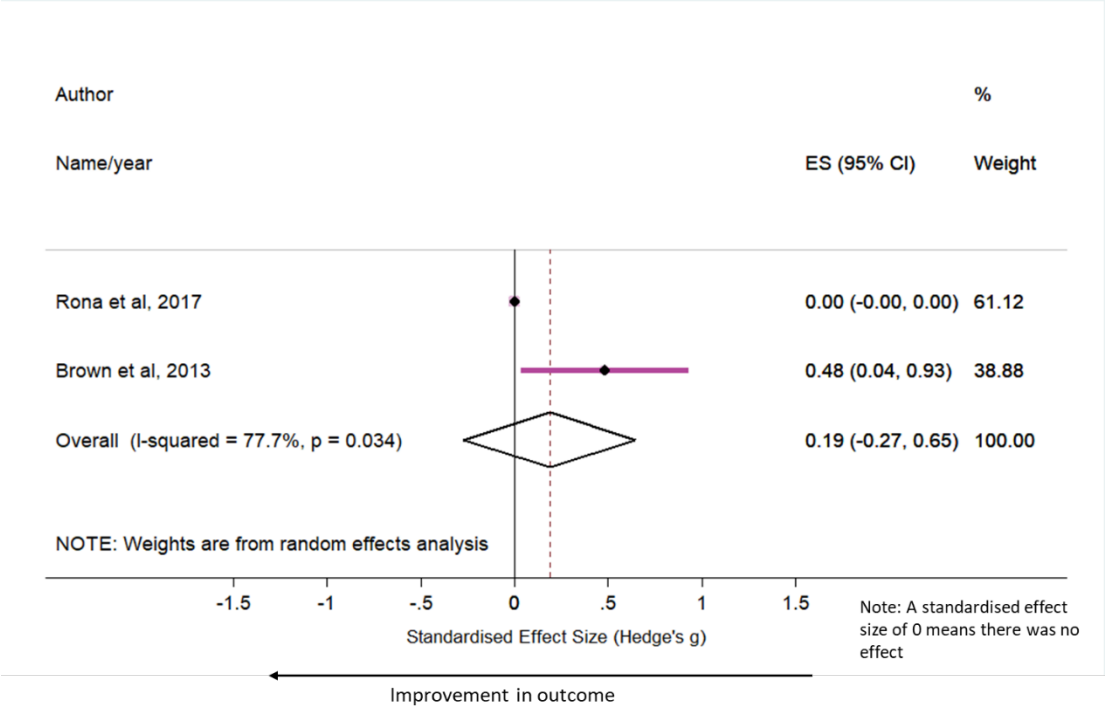
Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

5.2 The effectiveness of interventions to address mental health issues: forest plots

In the following section, we present forest plots from our meta-analyses. The forest plots are presented by intervention, and then by outcome, as in the main report. Where it has been possible to conduct sub-group and sensitivity analyses, these are presented immediately after the primary analysis for that intervention and outcome.

5.2.1 Awareness, screening and prevention

Figure A 5.12 The effect of awareness, screening and prevention on stress and associated disorders, including PTSD

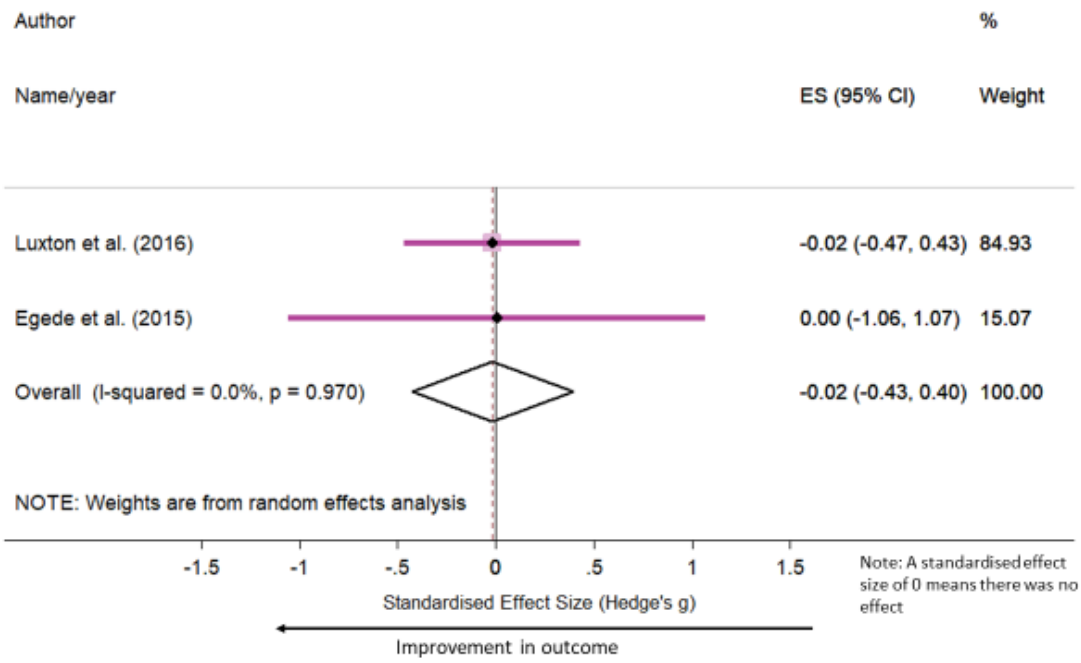


How to interpret forest plots

The forest plots in this chapter are graphical representations of the average effect of an intervention on a given outcome for each study. Each horizontal line indicates the results of one study, showing the average pooled effect (SMD) and the associated 95% confidence intervals. Where vertical lines do *not* intersect with the line of no effect, results are statistically significant. The diamond indicates the pooled SMD, whilst the shaded boxes represent the total weight placed on each study.

5.2.2 Behavioural Activation Therapy

Figure A 5.13 The effect of behavioural activation therapy on mood disorders



5.2.3 Cognitive behavioural therapy

Figure A 5.14 The effect of CBT on satisfaction with care

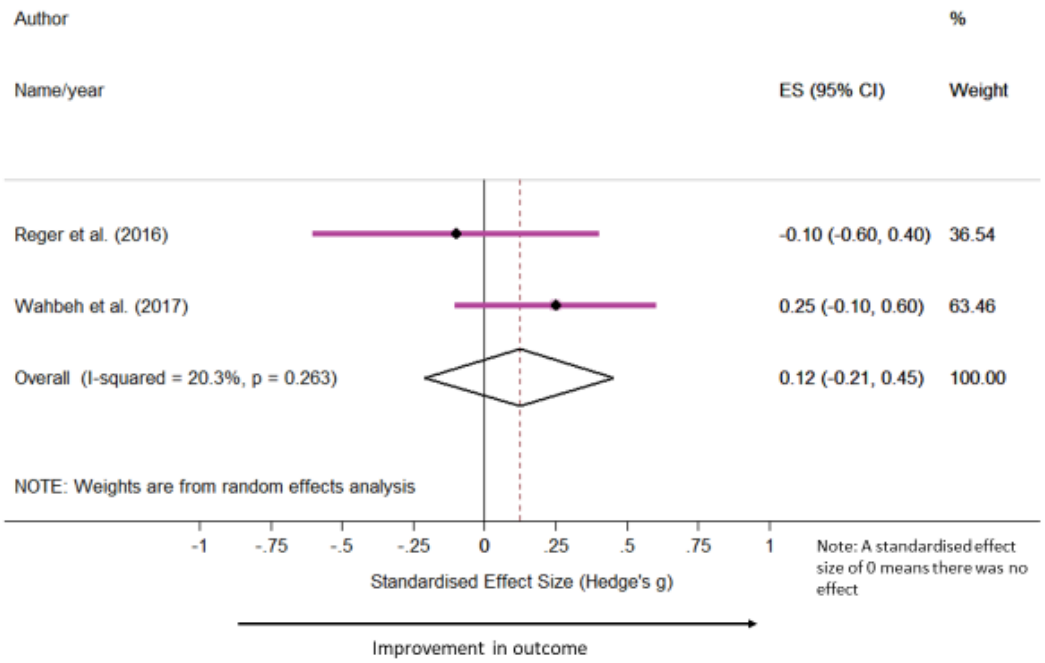


Figure A 5.15 The effect of CBT on physical health and wellbeing

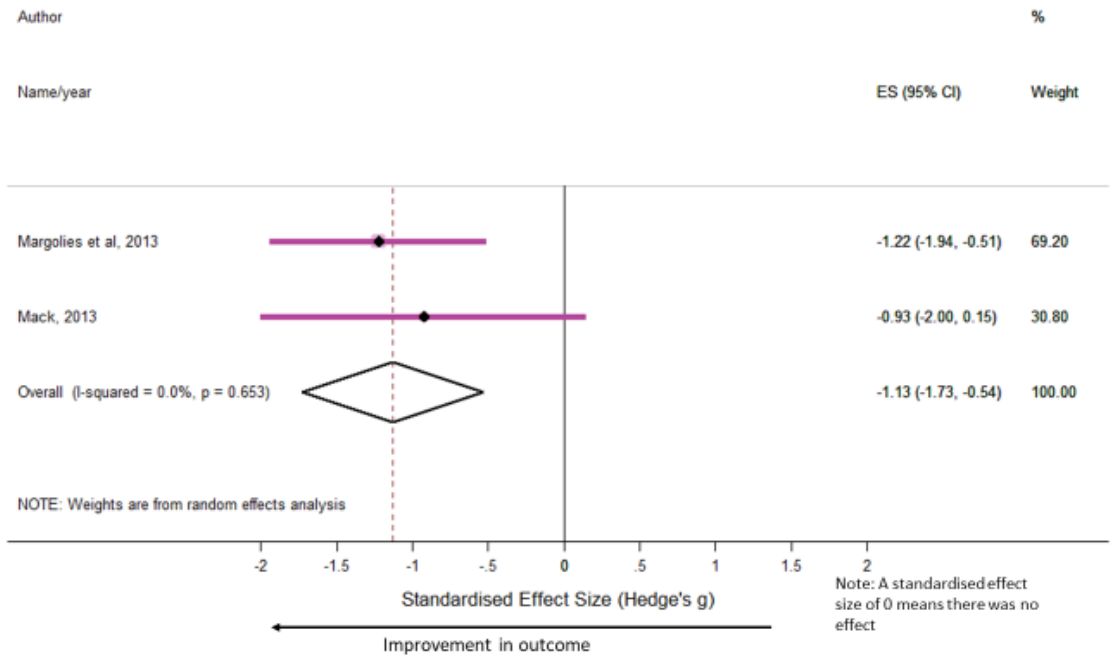


Figure A 5.16 The effect of CBT on stress and disorders associated with stress, including PTSD

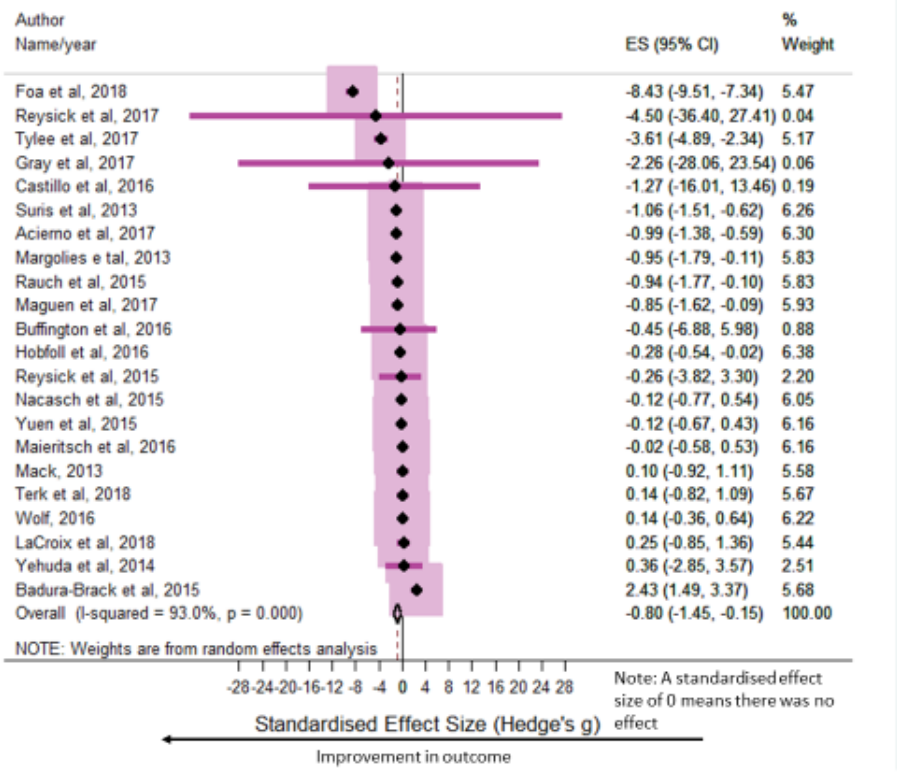


Figure A 5.17 The effect of CBT on stress and disorders associated with stress, including PTSD. Sub-group analysis : serving regulars

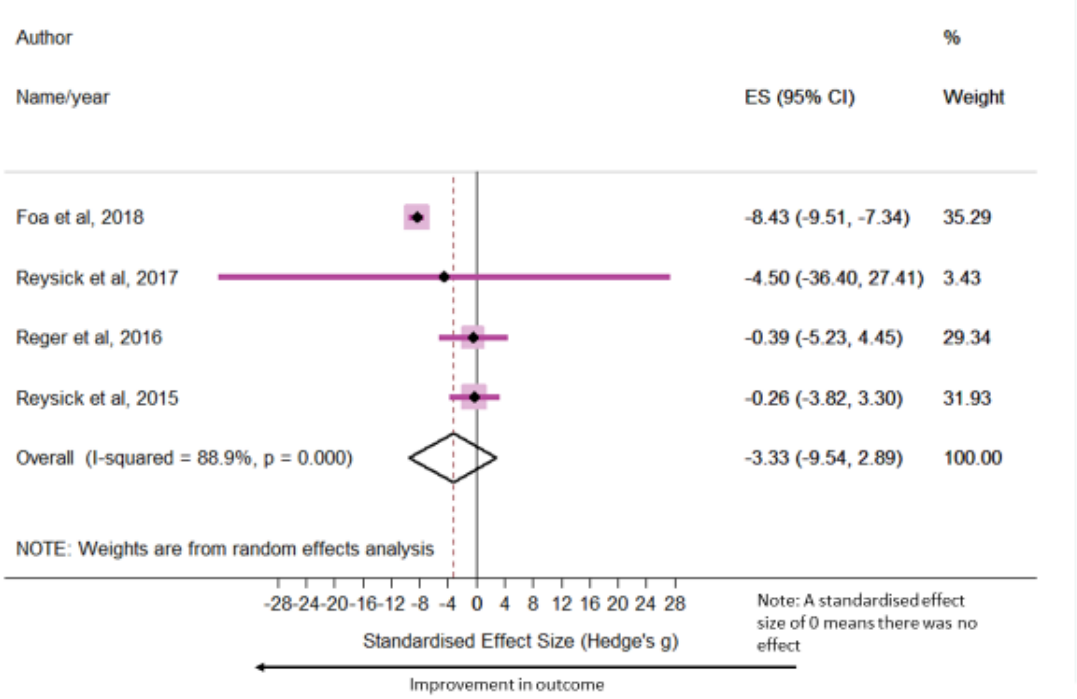


Figure A 5.18 The effect of CBT on stress and disorders associated with stress, including PTSD. Sub-group analysis : veterans

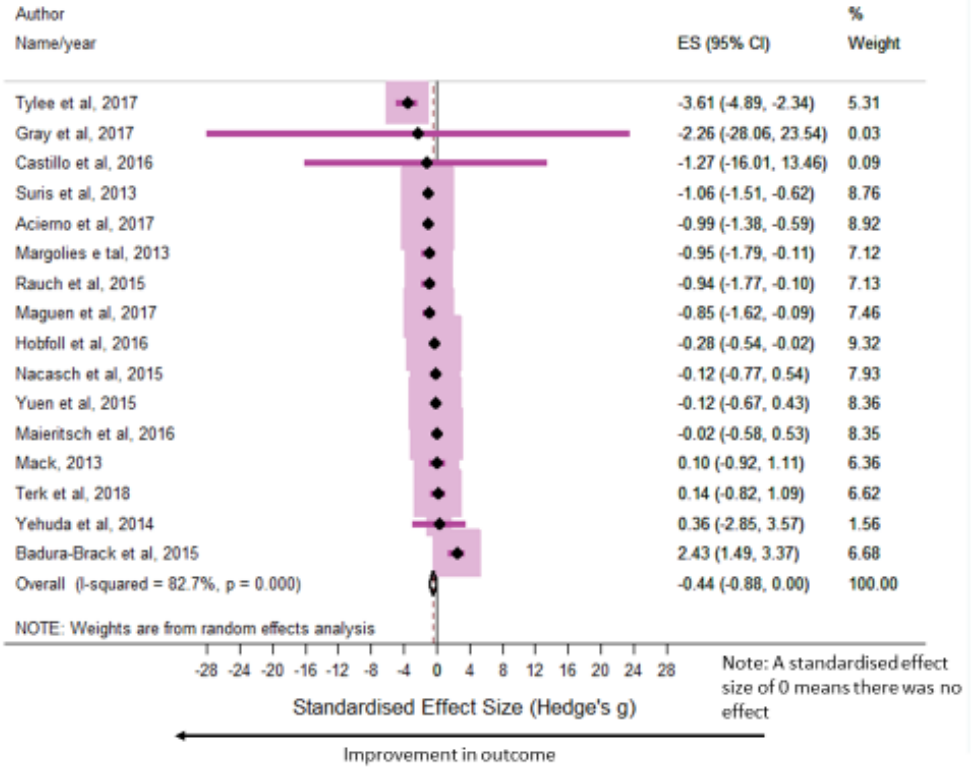


Figure A 5.19 The effect of CBT on stress and disorders associated with stress, including PTSD. Sensitivity analysis : low risk of bias

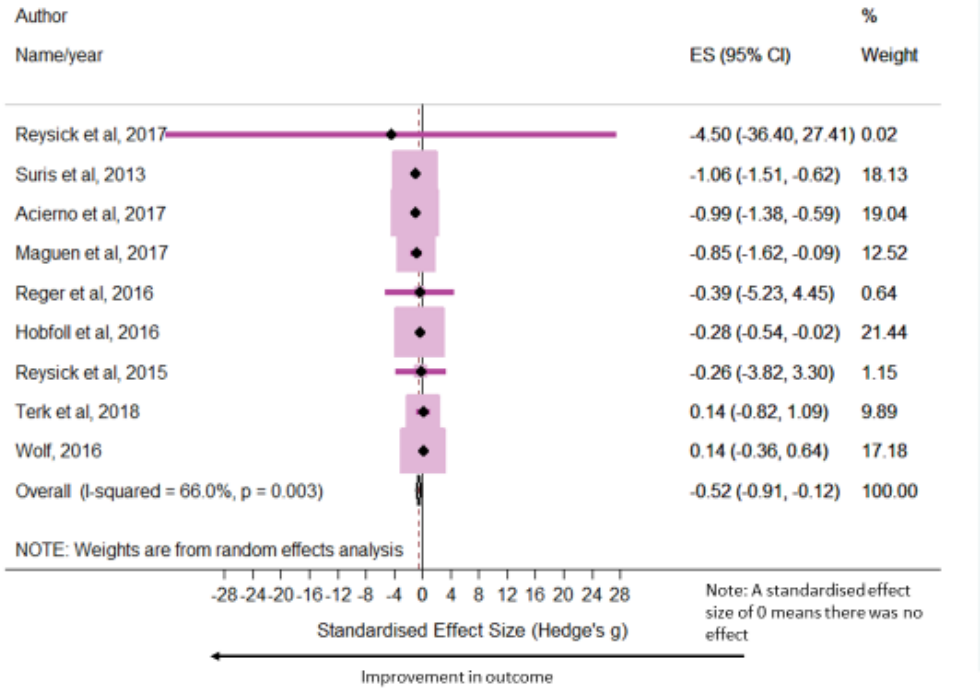


Figure A 5.20 The effect of CBT on stress and disorders associated with stress, including PTSD. Sensitivity analysis : six-month follow-up period

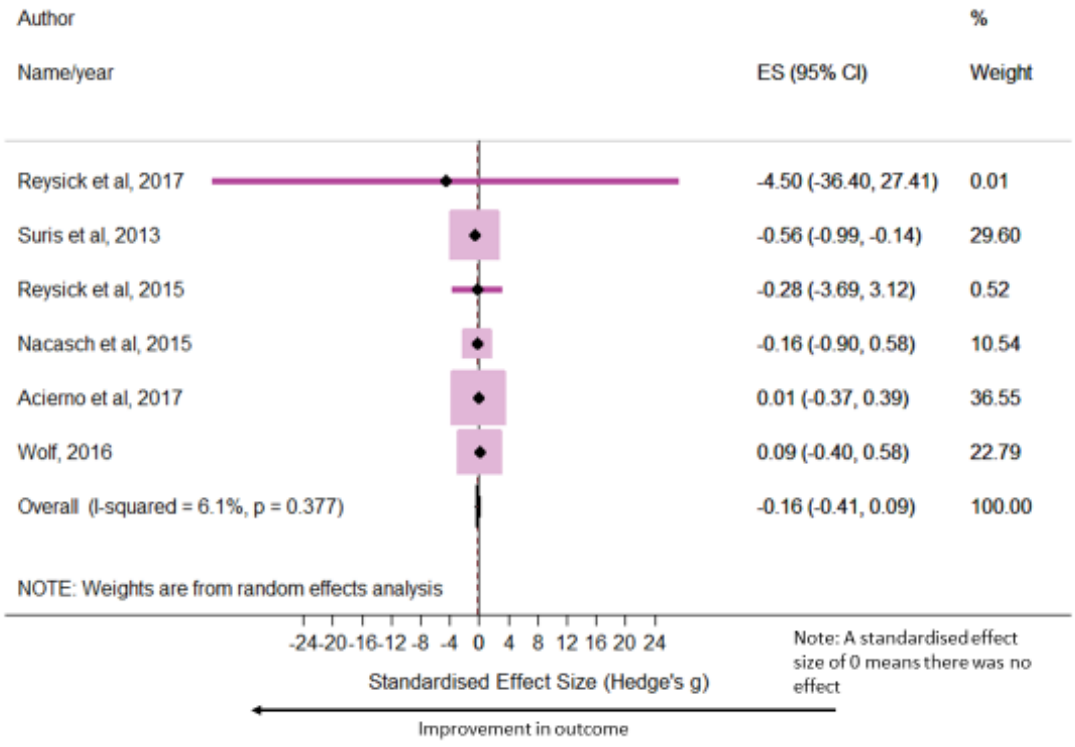


Figure A 5.21 The effect of CBT on mood disorders

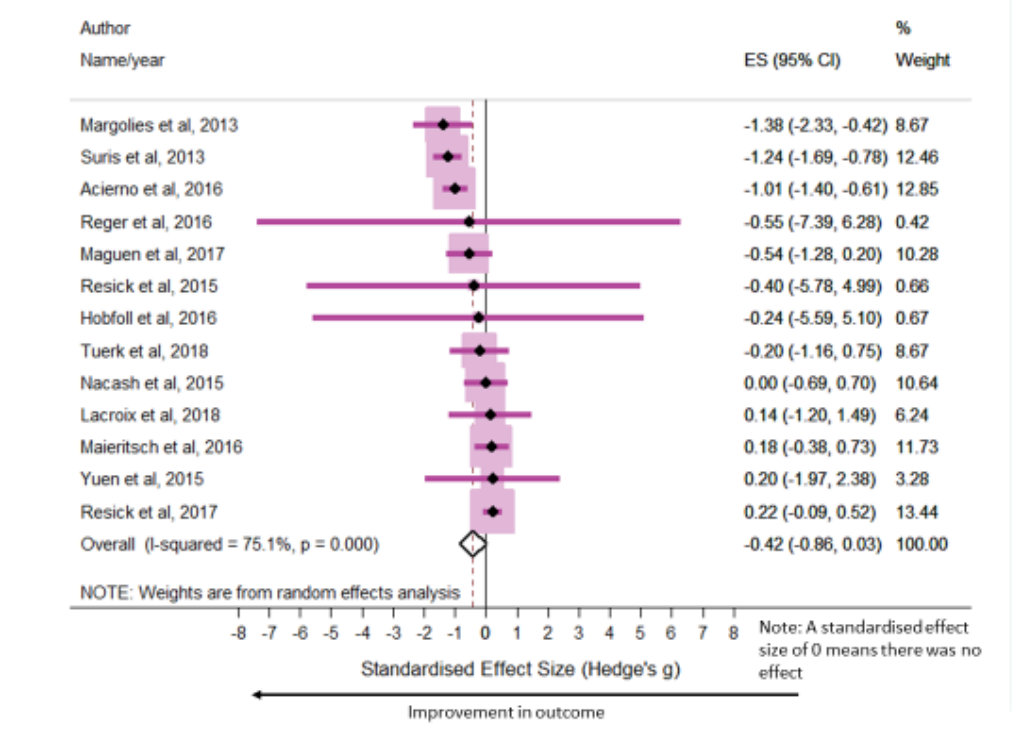


Figure A 5.22 The effect of CBT on mood disorders. Sub-group analysis : serving regulars

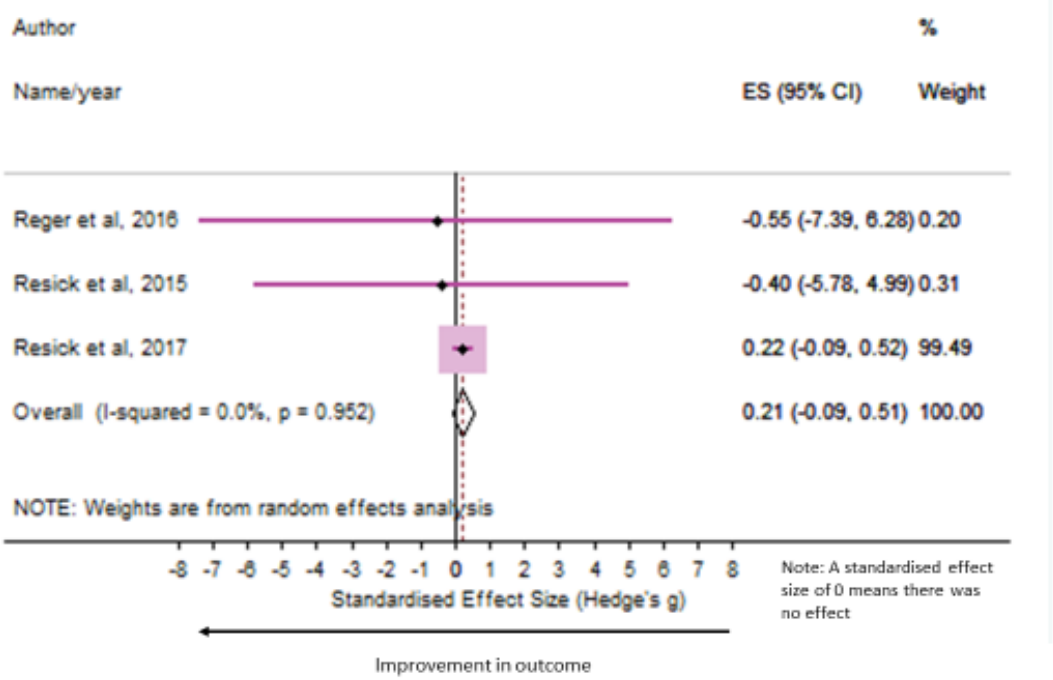


Figure A 5.23 The effect of CBT on mood disorders. Sub-group analysis : veterans

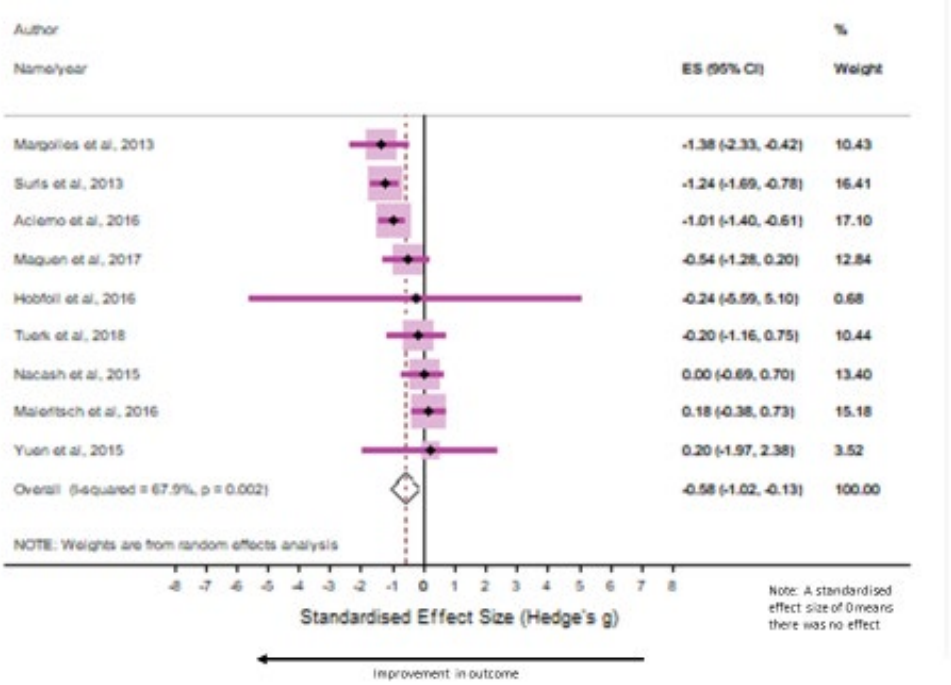


Figure A 5.24 The effect of CBT on mood disorders. Sensitivity analysis : six-month follow-up period

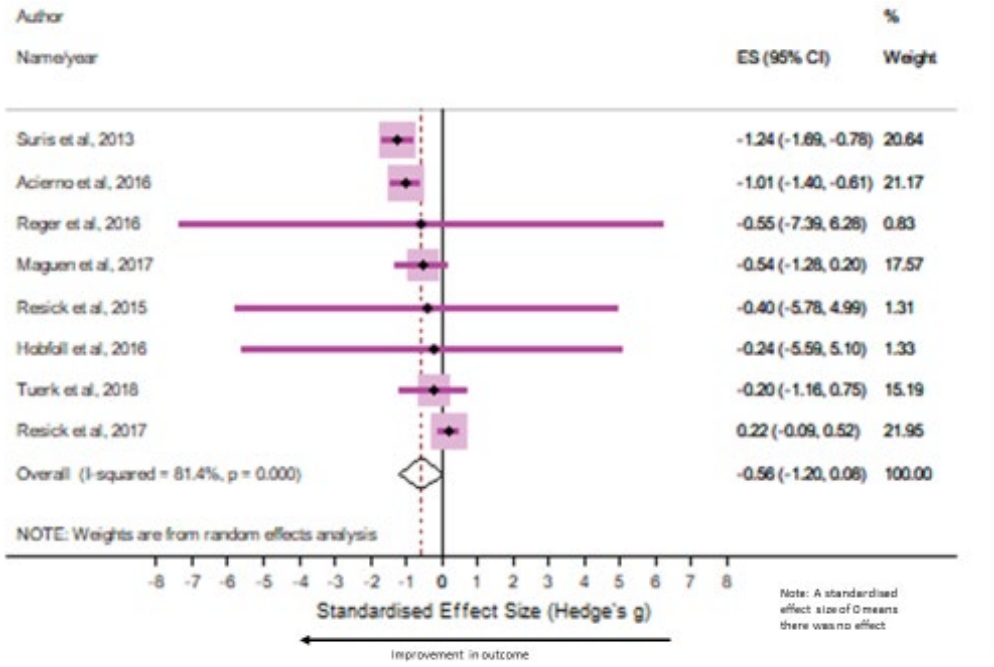


Figure A 5.25 The effect of CBT on anxiety and fear

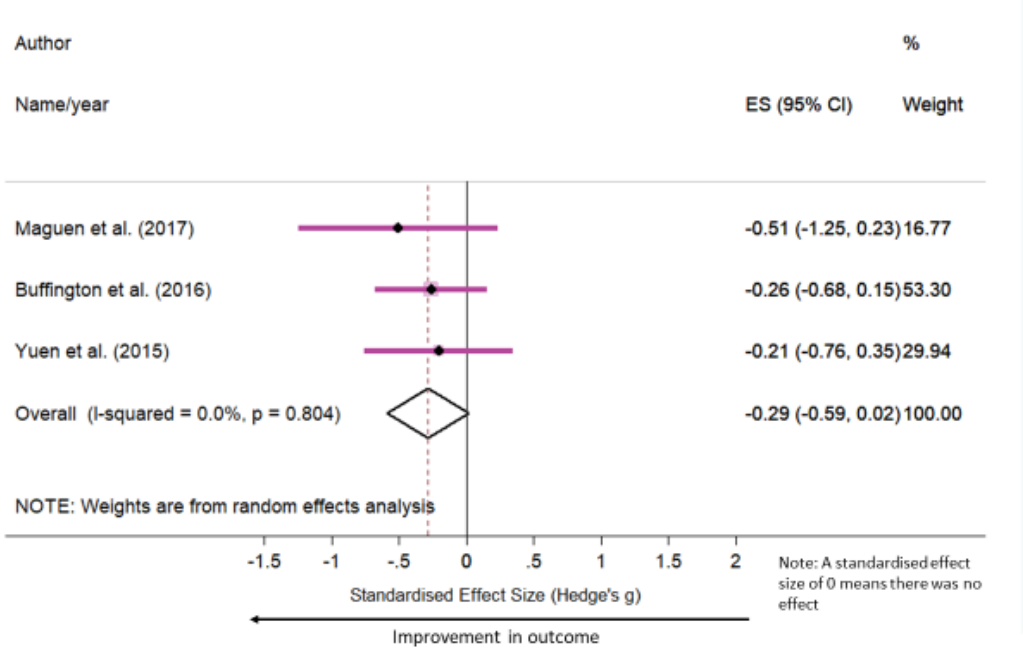


Figure A 5.26 The effect of CBT on anxiety and fear. Sub-group analysis : veterans

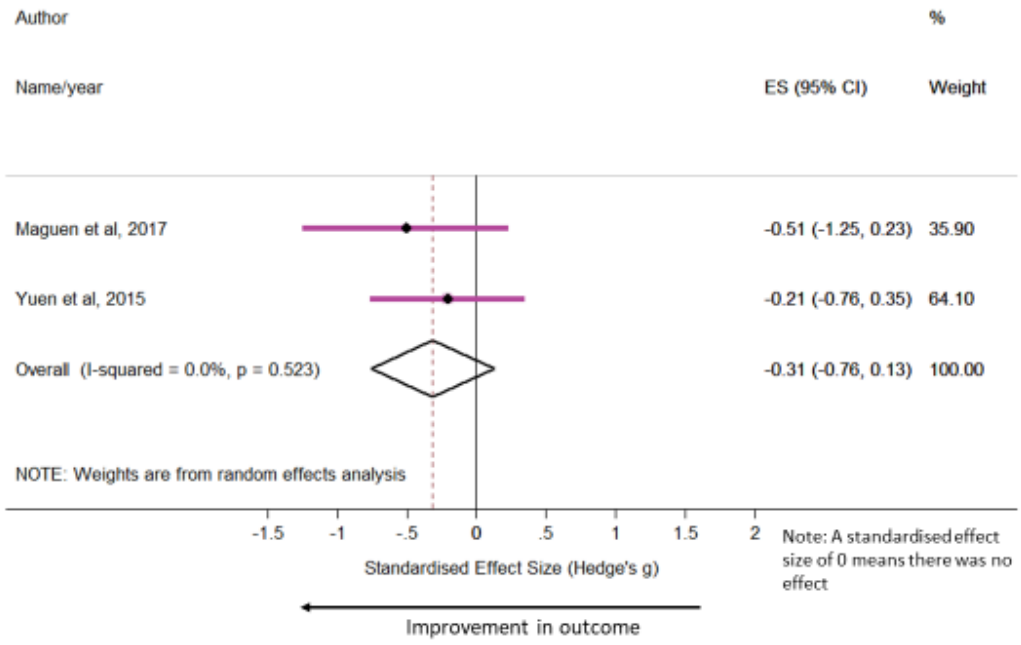


Figure A 5.27 The effect of CBT on suicide and self-harm

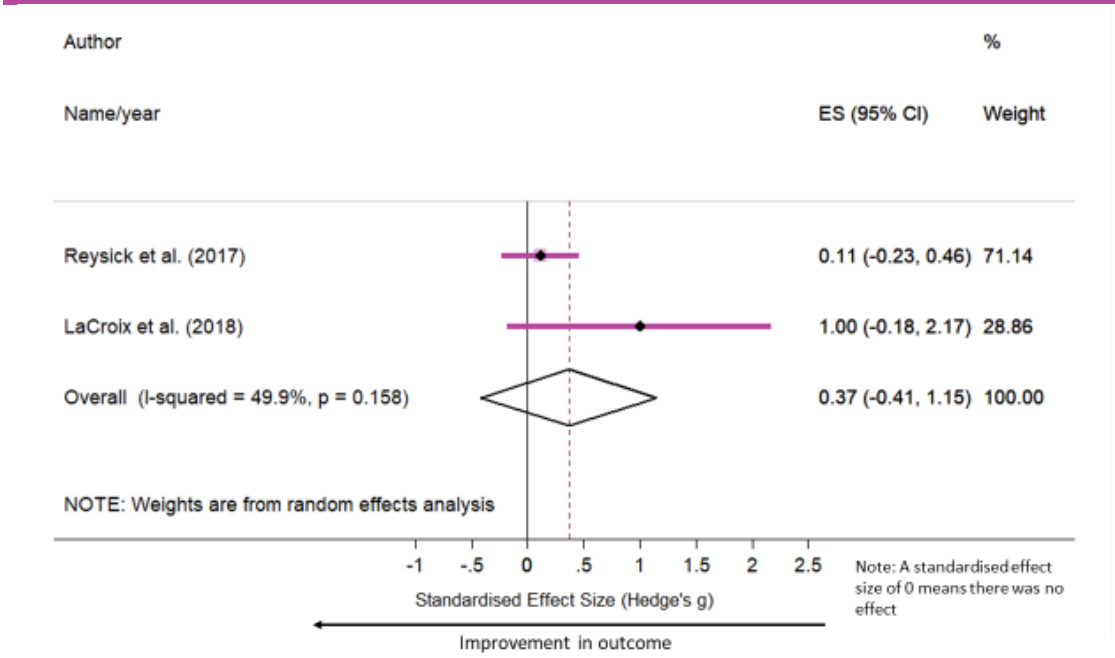


Figure A 5.28 The effect of CBT on other mental health issues

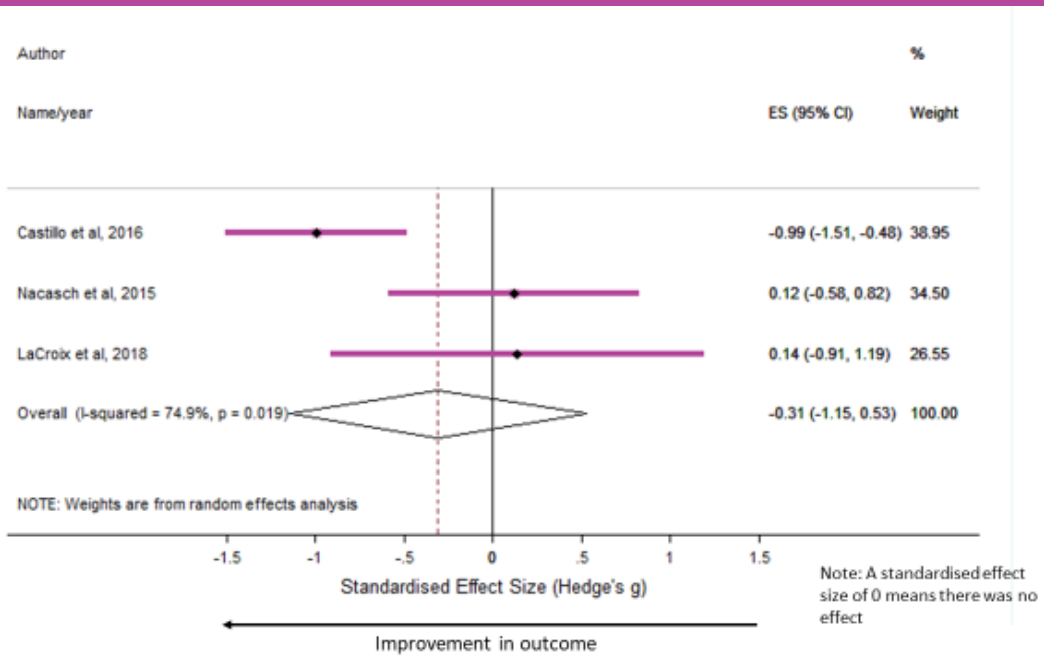
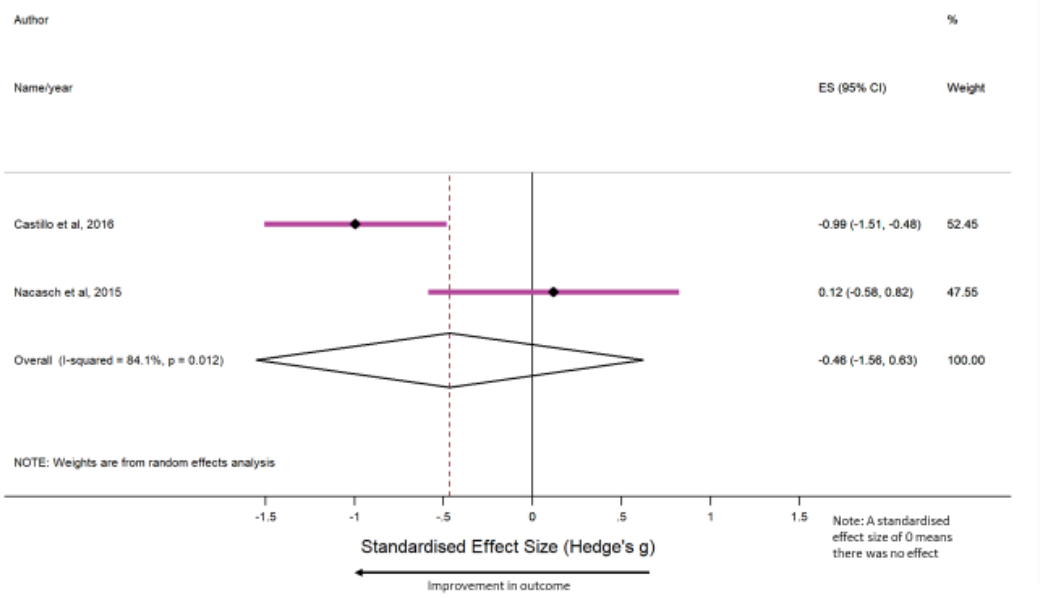


Figure A 5.29 The effect of CBT on other mental health issues. Sub-group analysis : veterans



5.2.4 Family therapy interventions

Figure A 5.30 The effect of family therapy interventions on mood disorders

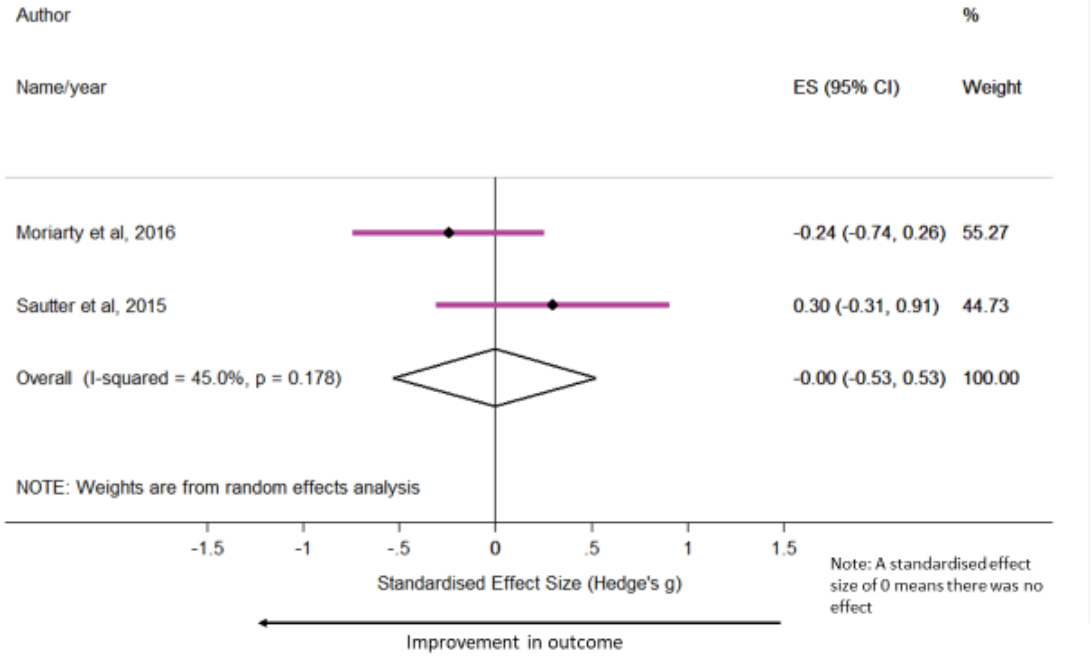
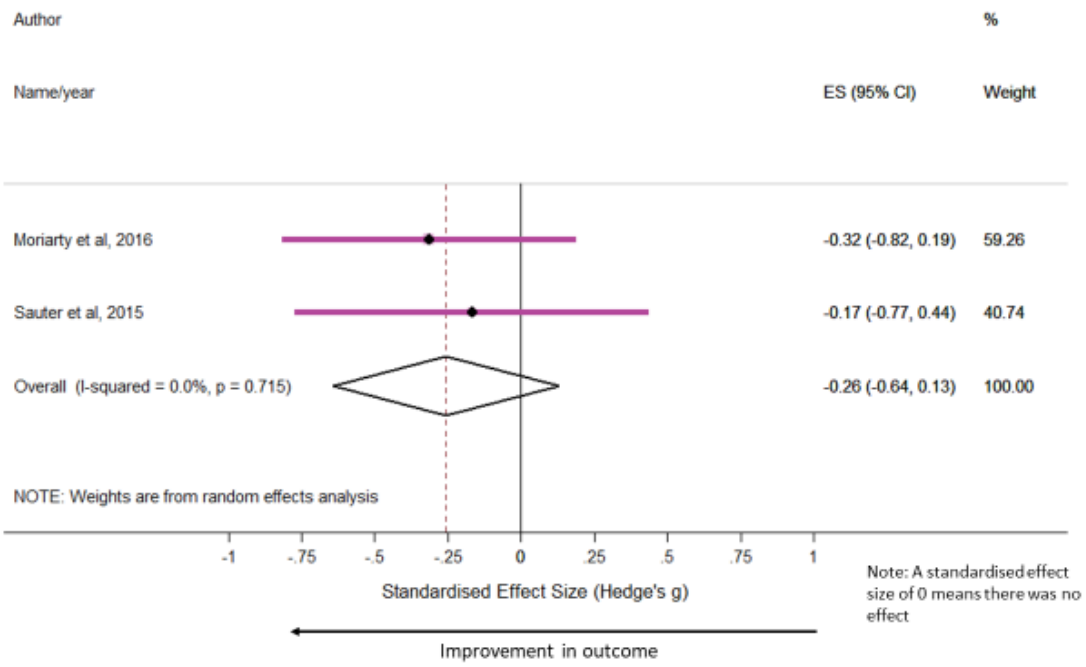


Figure A 5.31 The effect of family therapy interventions on other mental health issues



5.2.5 Wellbeing interventions

Figure A 5.32 The effect of wellbeing interventions on delivery

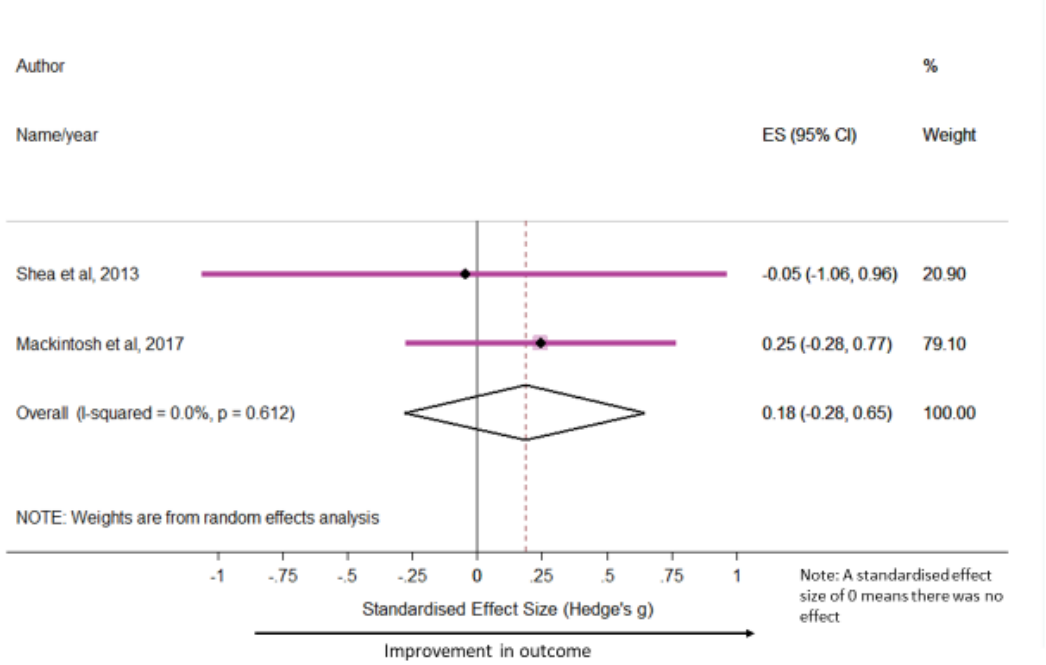


Figure A 5.33 The effect of wellbeing interventions on physical health and wellbeing

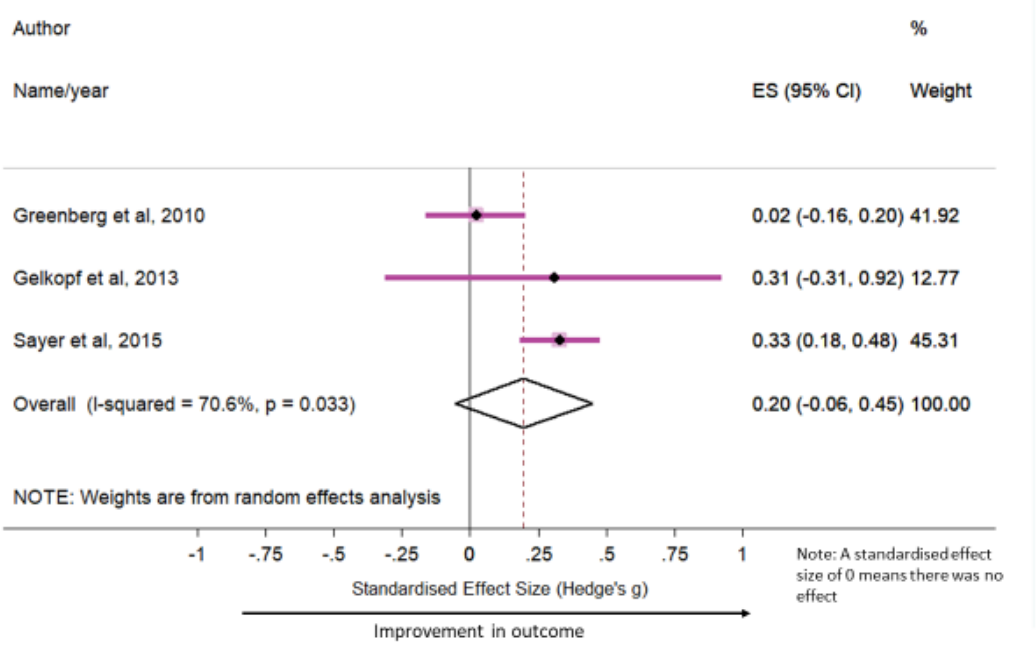


Figure A 5.34 The effect of wellbeing interventions on physical health and wellbeing. Sub-group analysis : veterans

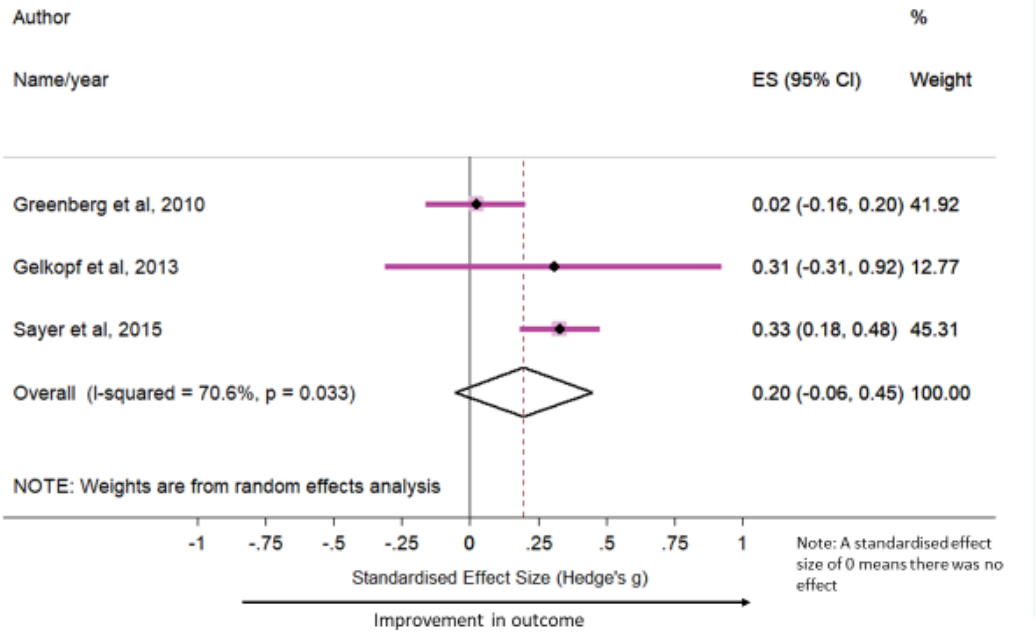


Figure A 5.35 The effect of wellbeing interventions on stress and disorders associated with stress

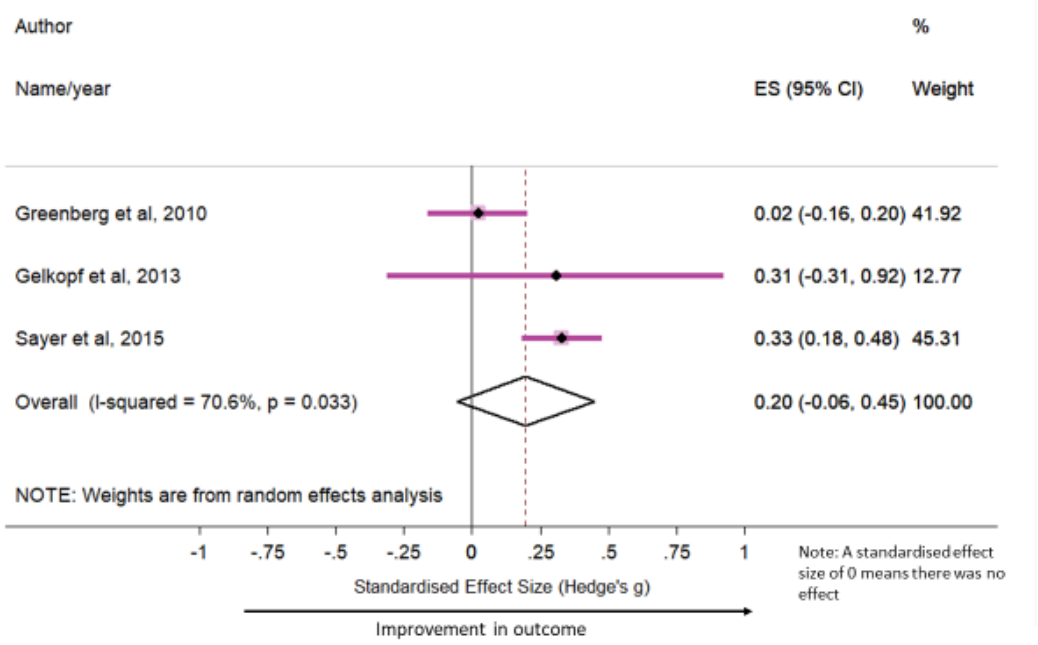


Figure A 5.36 The effect of wellbeing interventions on stress and disorders associated with stress. Sub-group analysis : veterans

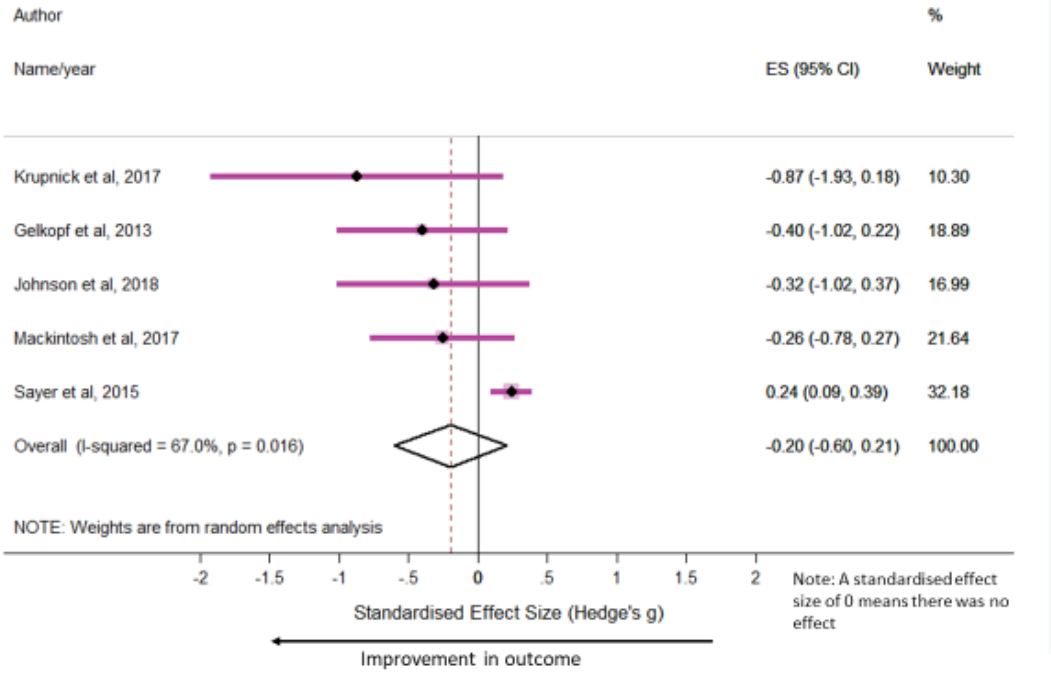


Figure A 5.37 The effect of wellbeing interventions on mood disorders

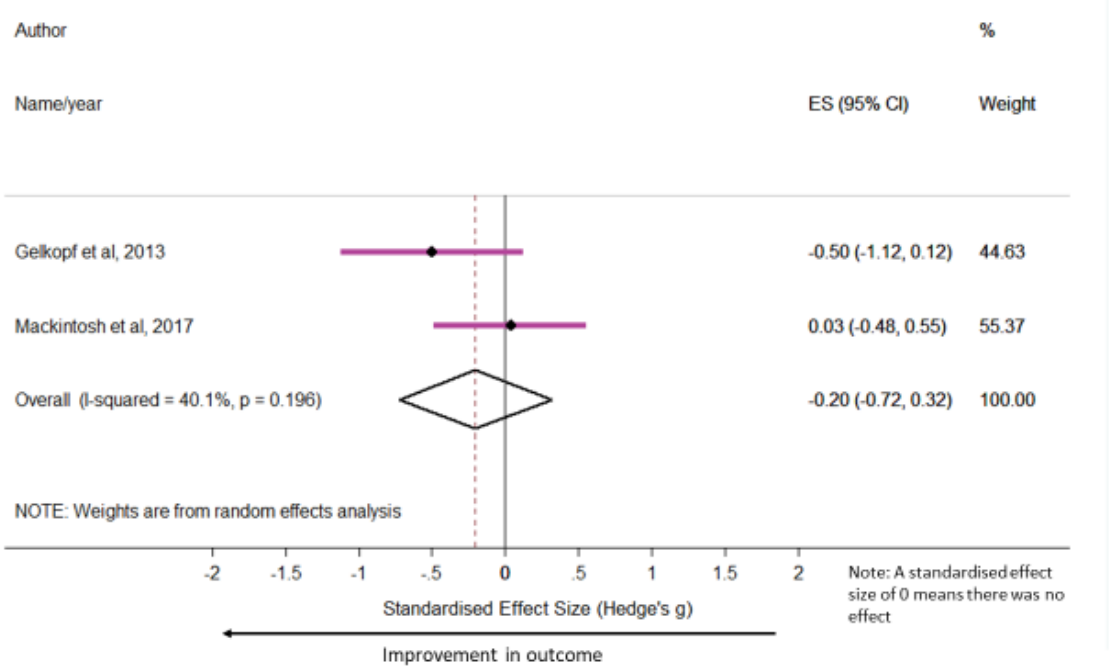


Figure A 5.38 The effect of wellbeing interventions on other mental health issues

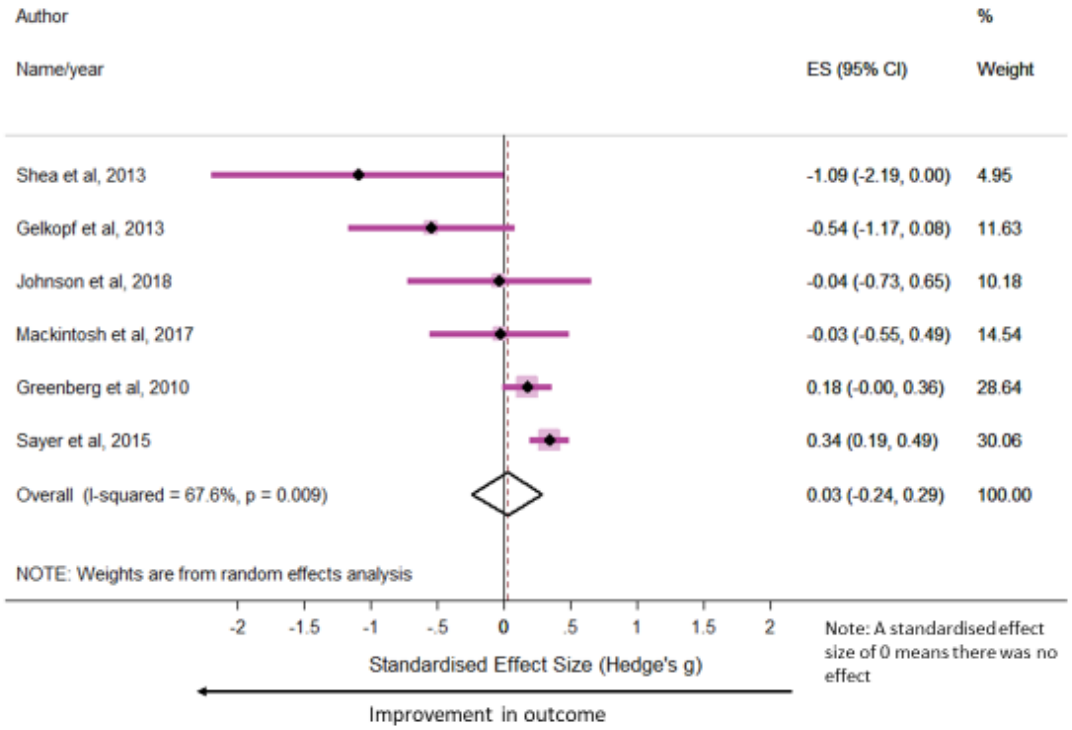


Figure A 5.39 The effect of wellbeing interventions on other mental health issues. Sub-group analysis : veterans

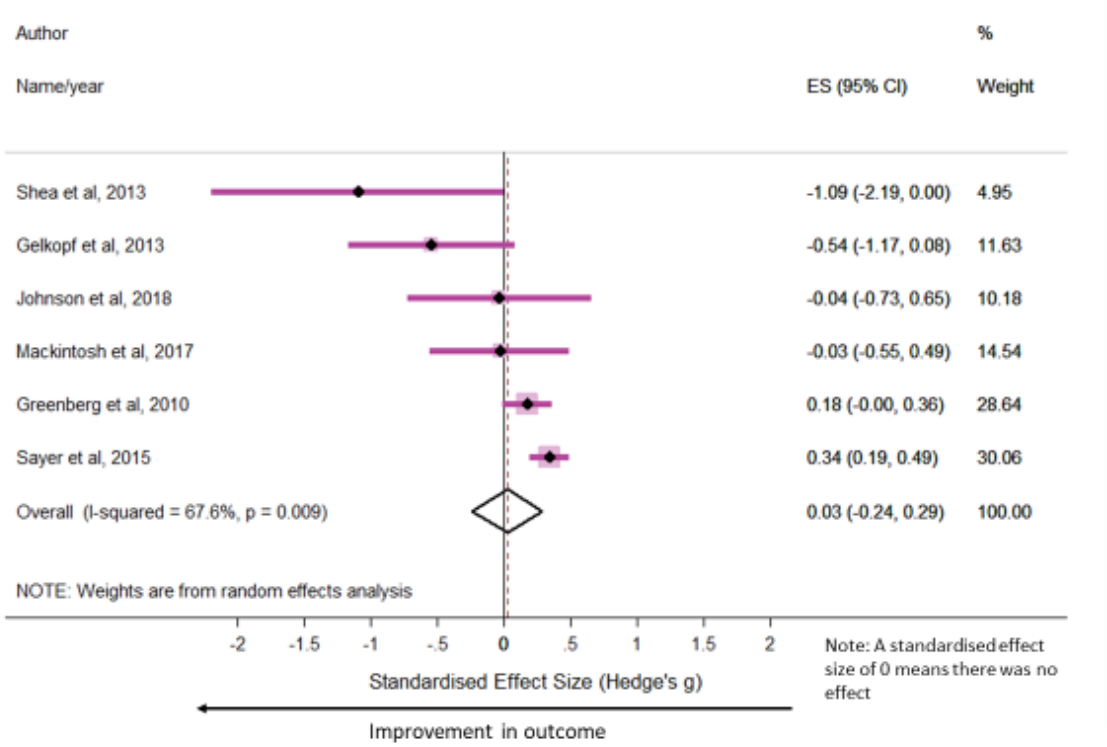
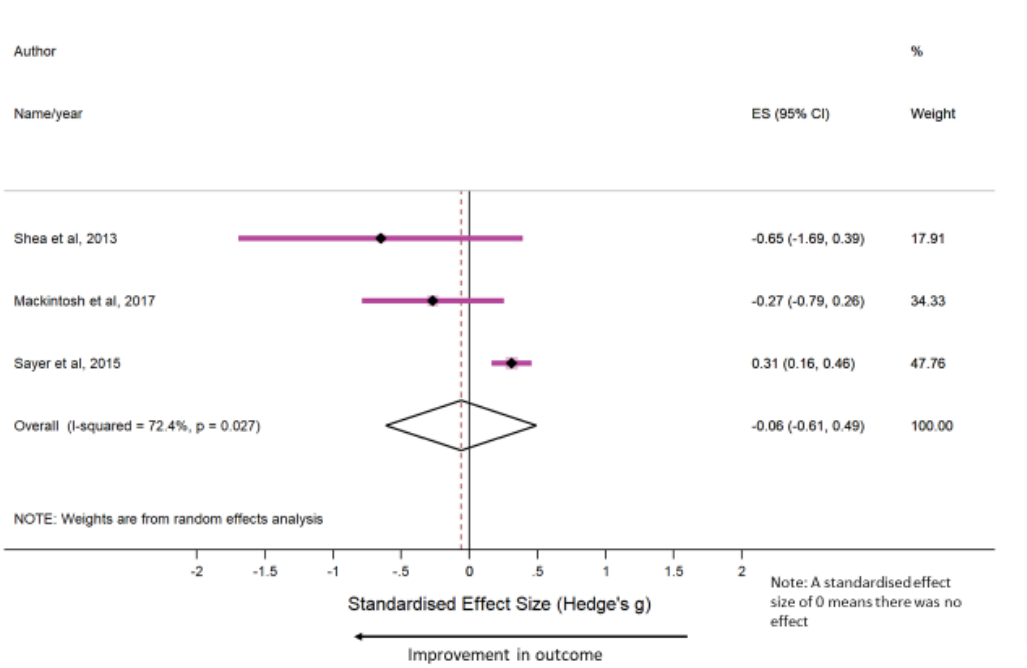


Figure A 5.40 The effect of wellbeing interventions on other mental health issues. Sensitivity analysis : three-month follow-up period



5.2.6 Meditation and mindfulness interventions

Figure A 5.41 The effect of meditation and mindfulness interventions on physical health and wellbeing (excluding sleep)

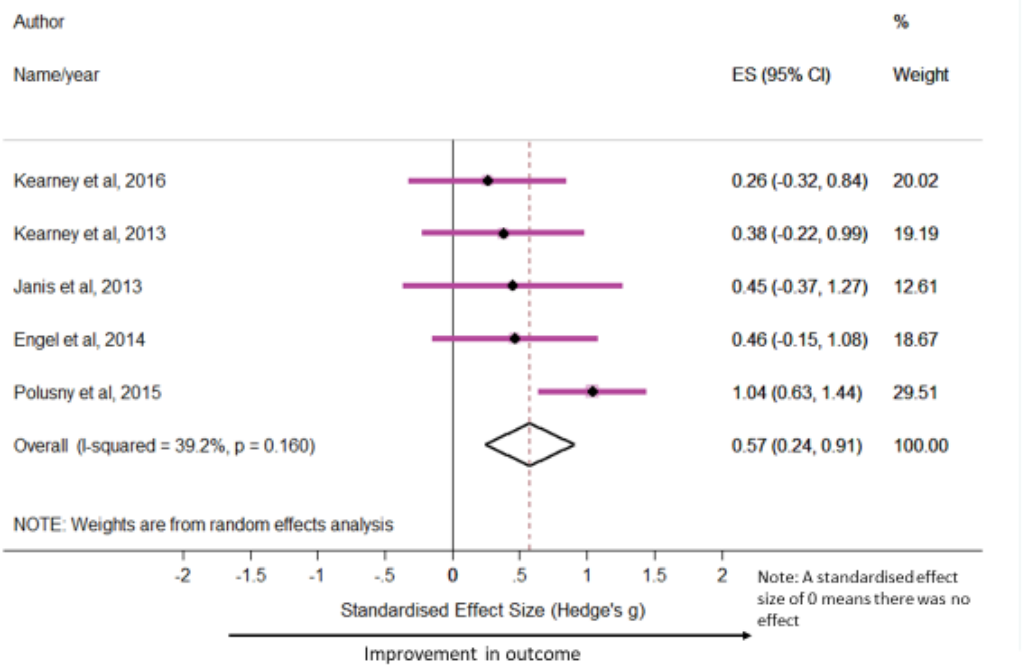


Figure A 5.42 The effect of meditation and mindfulness interventions on physical health and wellbeing (excluding sleep). Sub-group analysis : veterans

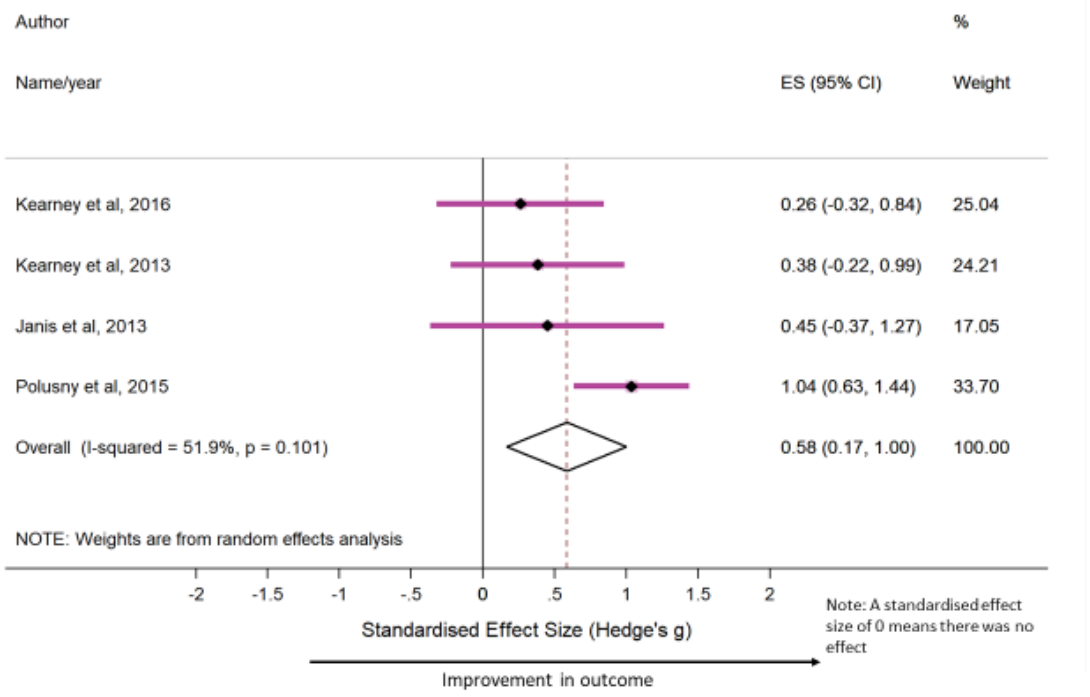


Figure A 5.43 The effect of meditation and mindfulness interventions on physical health and wellbeing (sleep)

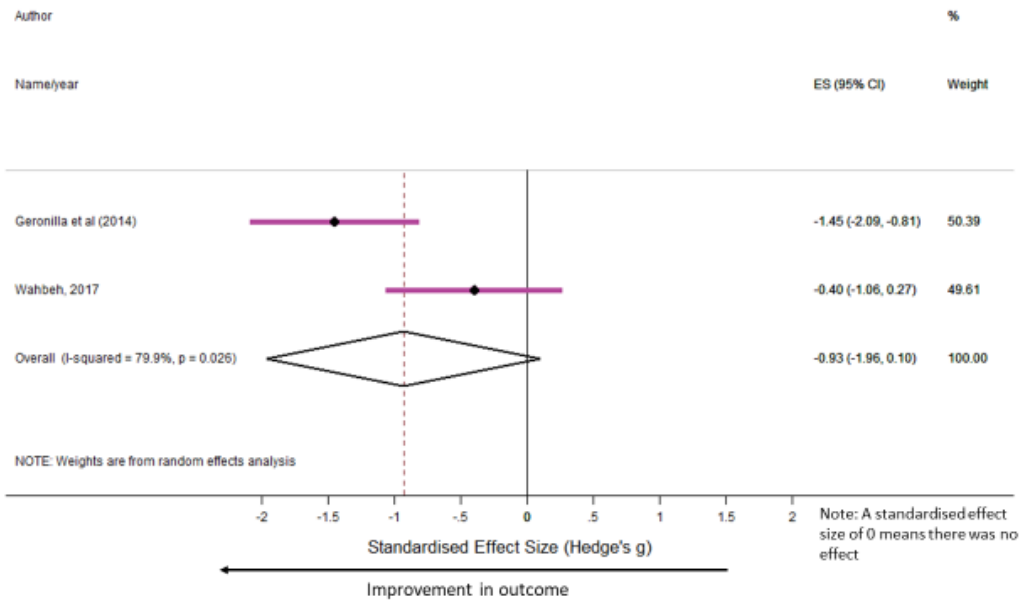


Figure A 5.44 The effect of meditation and mindfulness interventions on stress and disorders associated with stress, including PTSD

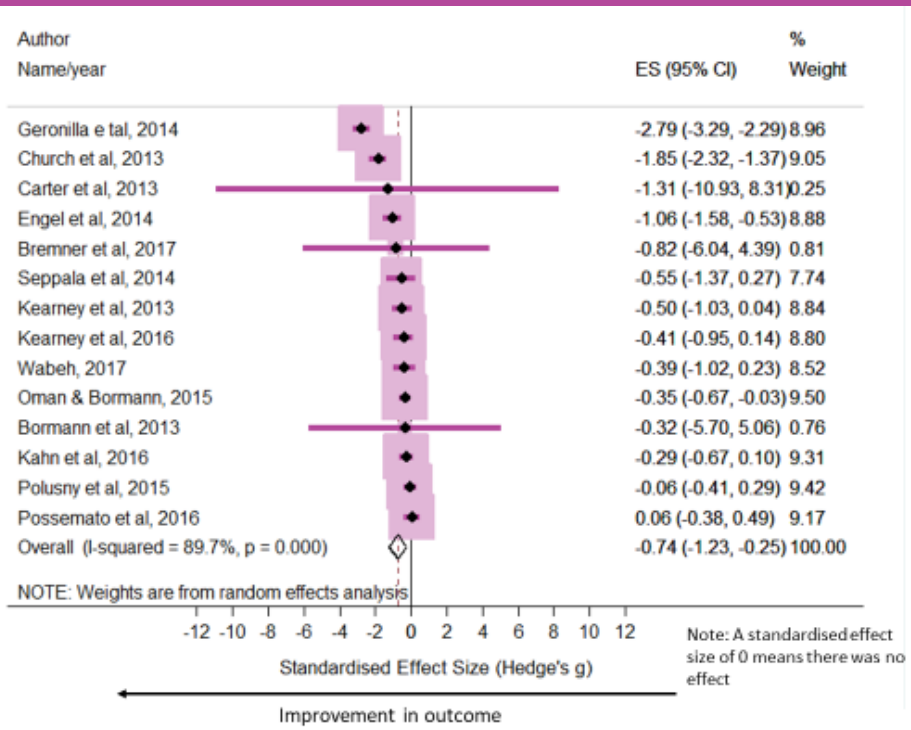


Figure A 5.45 The effect of meditation and mindfulness interventions on stress and disorders associated with stress, including PTSD. Sub-group analysis : veterans

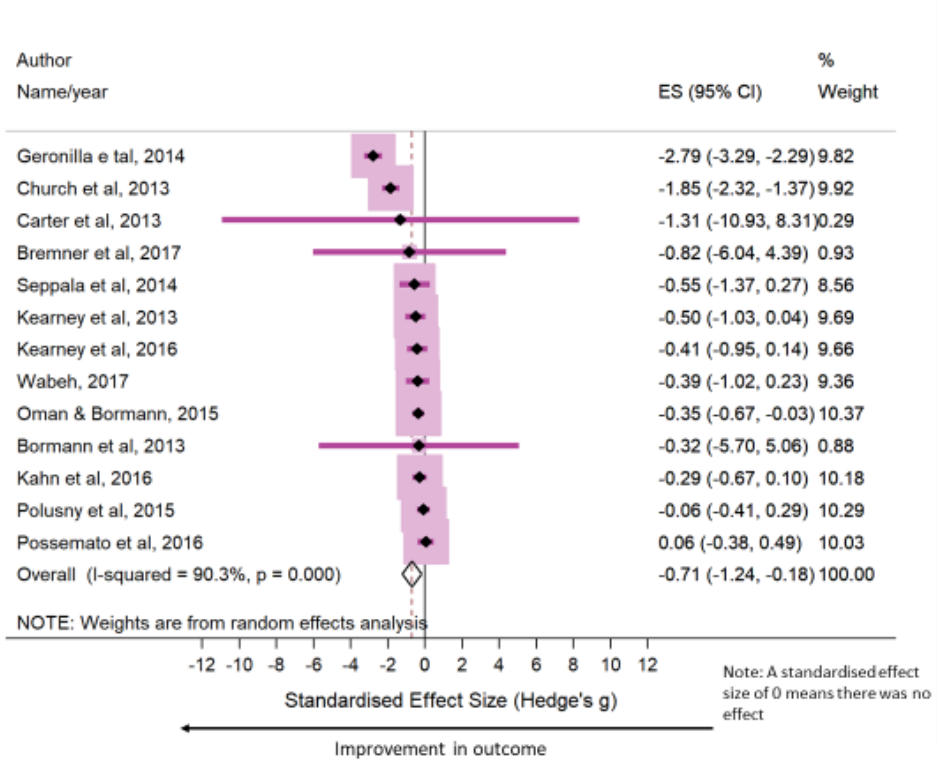


Figure A 5.46 The effect of meditation and mindfulness interventions on stress and disorders associated with stress, including PTSD. Sensitivity analysis : low risk of bias

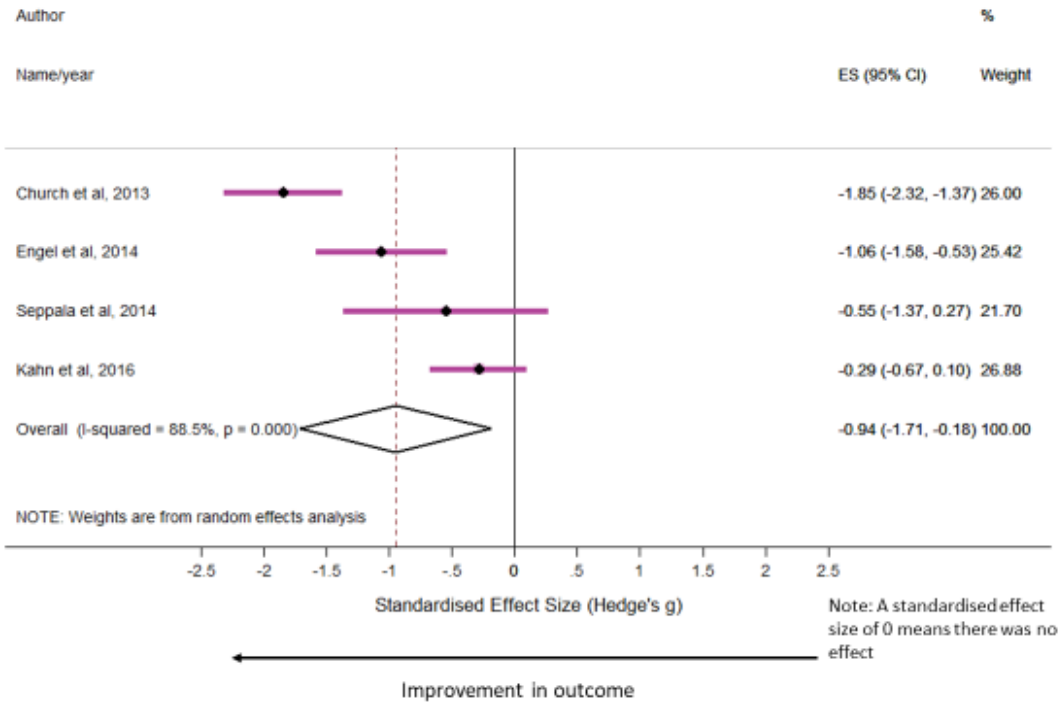


Figure A 5.47 The effect of meditation and mindfulness interventions on stress and disorders associated with stress, including PTSD. Sensitivity analysis : two-months follow-up period

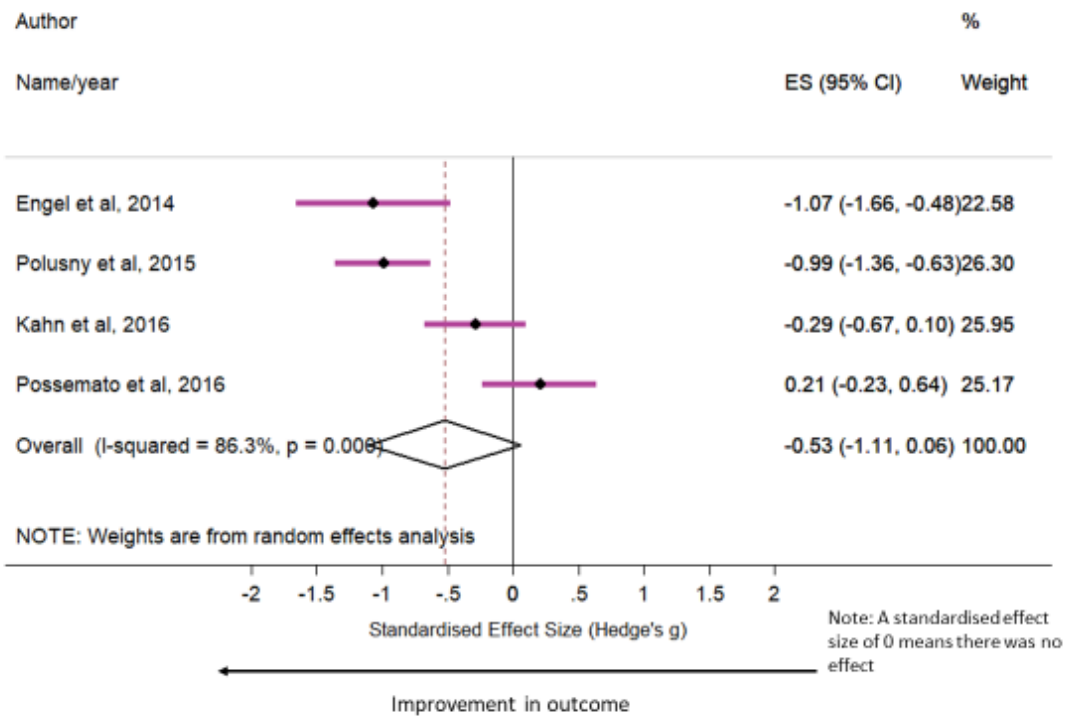


Figure A 5.48 The effect of meditation and mindfulness interventions on mood disorders

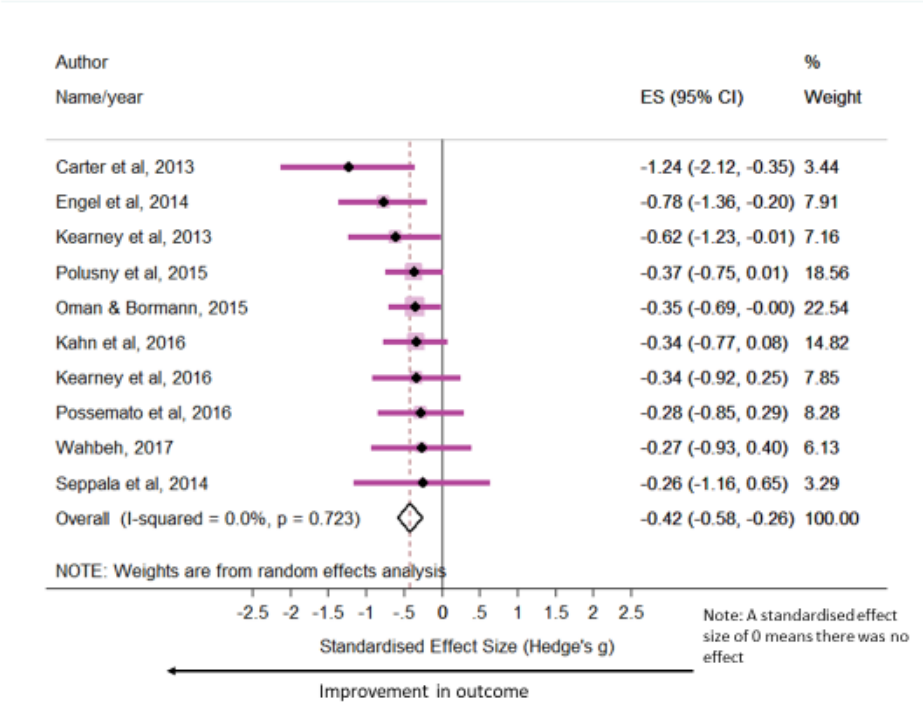


Figure A 5.49 The effect of meditation and mindfulness interventions on mood disorders. Sub-group analysis : veterans

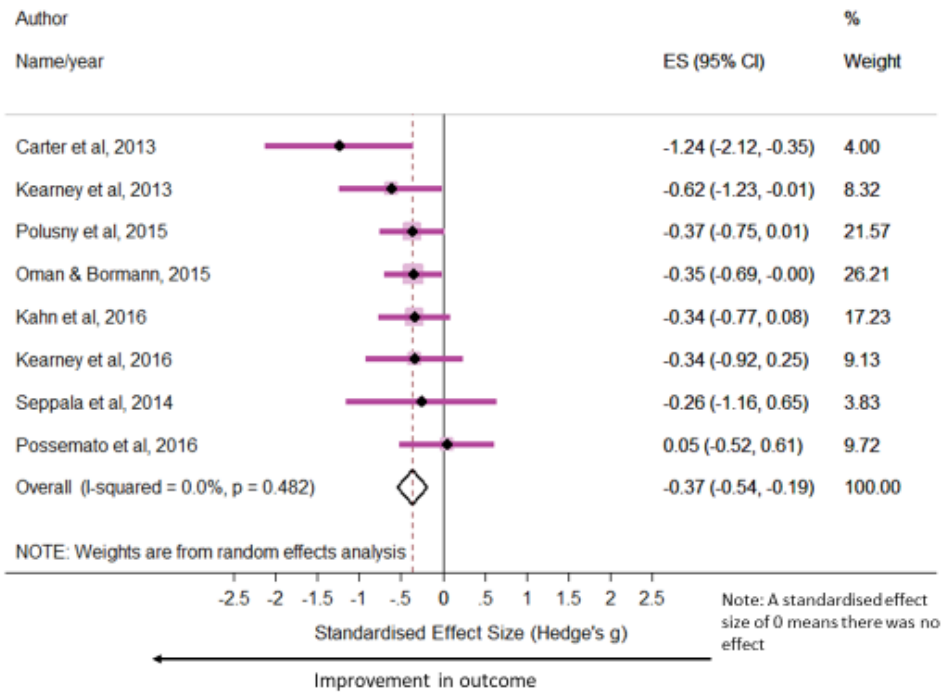


Figure A 5.50 The effect of meditation and mindfulness interventions on mood disorders. Sensitivity analysis : low risk of bias

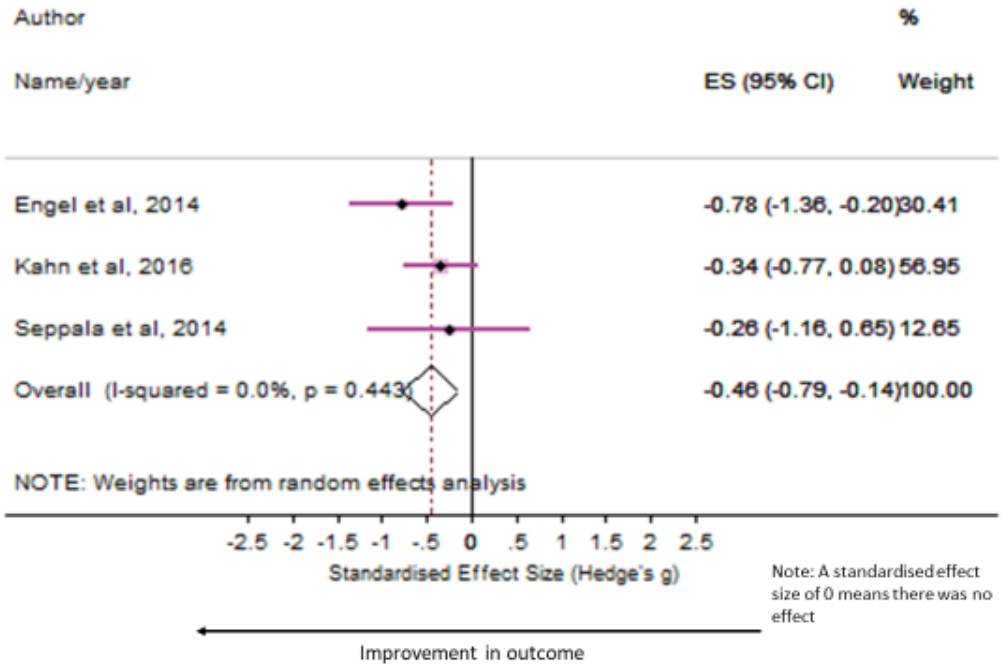


Figure A 5.51 The effect of meditation and mindfulness interventions on mood disorders. Sensitivity analysis : two-month follow-up period

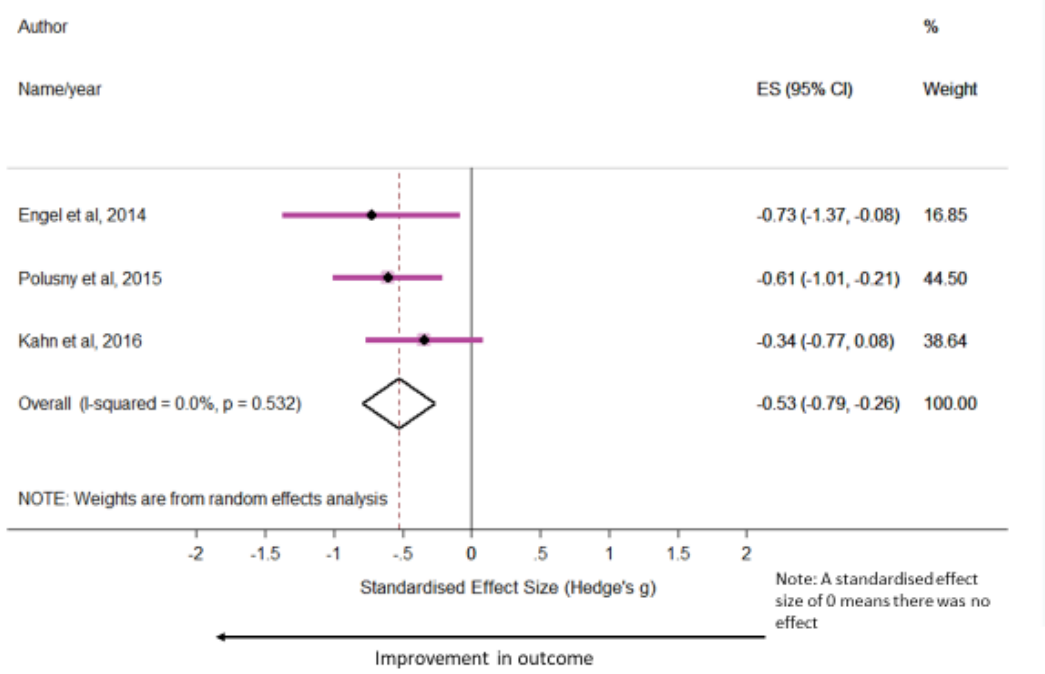


Figure A 5.52 The effect of meditation and mindfulness interventions on anxiety and fear

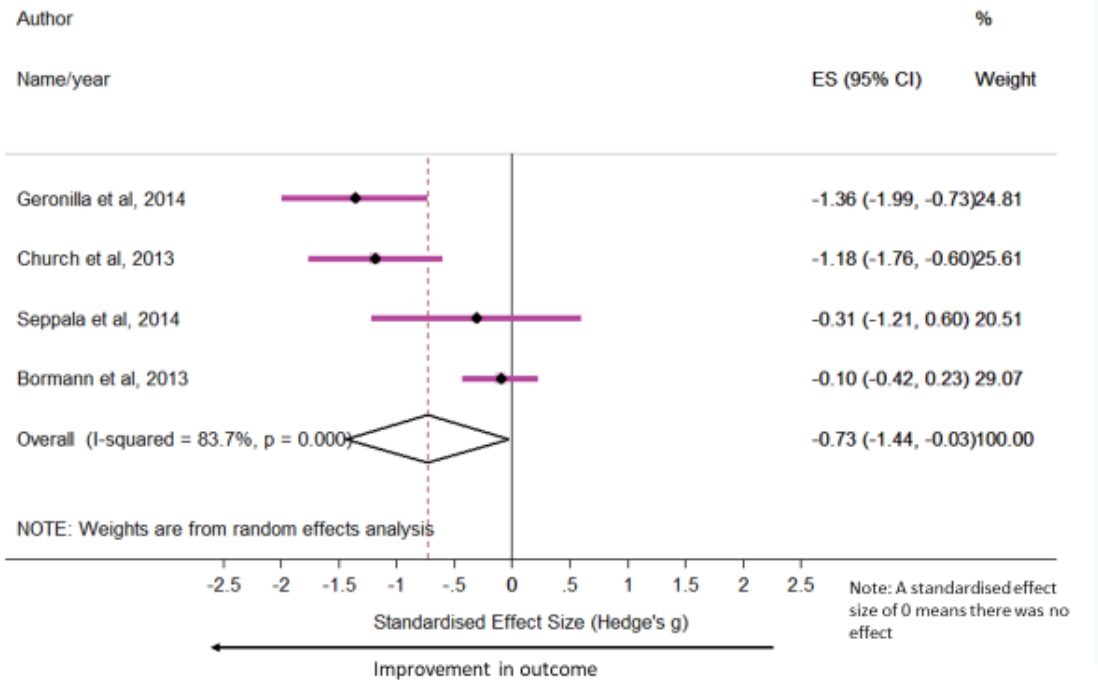


Figure A 5.53 The effect of meditation and mindfulness interventions on anxiety and fear. Sensitivity analysis : low risk of bias

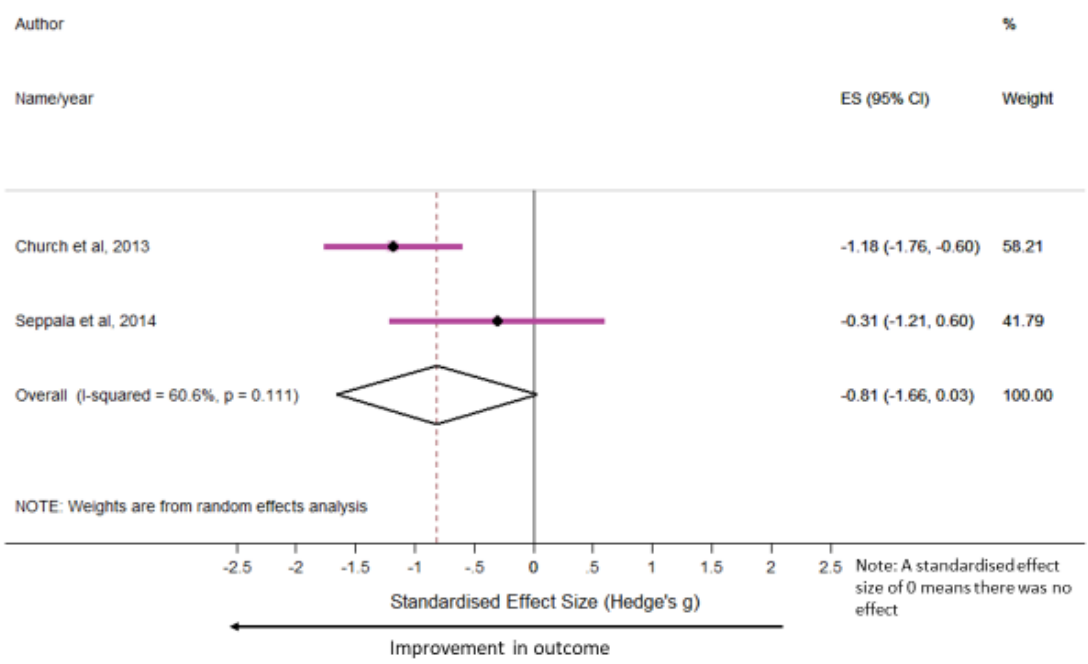


Figure A 5.54 The effect of meditation and mindfulness interventions on other mental health issues

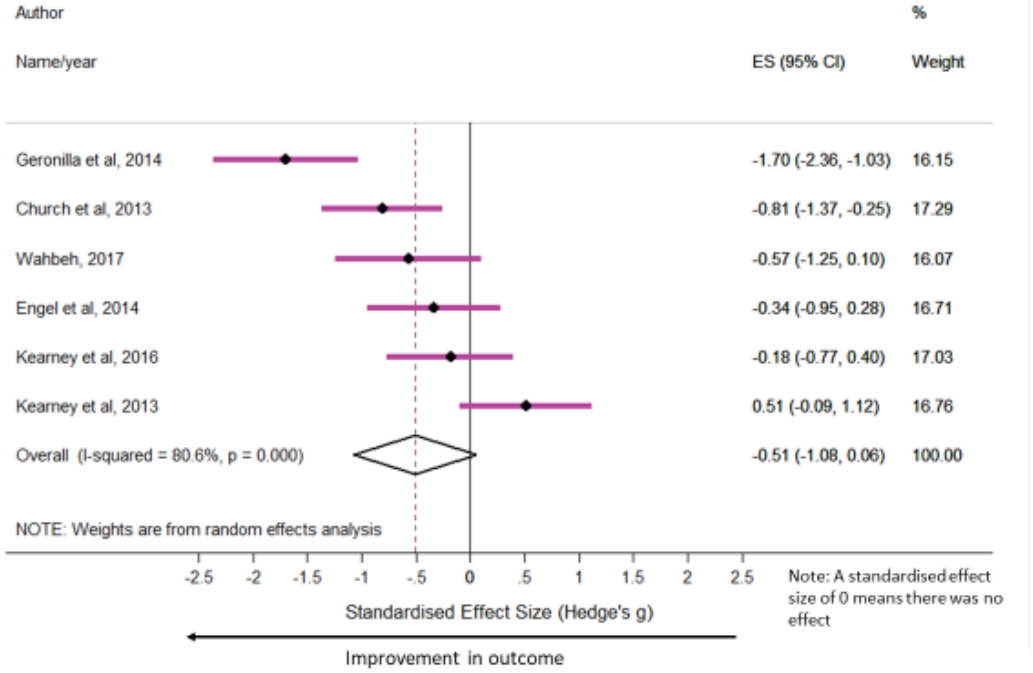


Figure A 5.55 The effect of meditation and mindfulness interventions on other mental health issues. Sub-group analysis : veterans

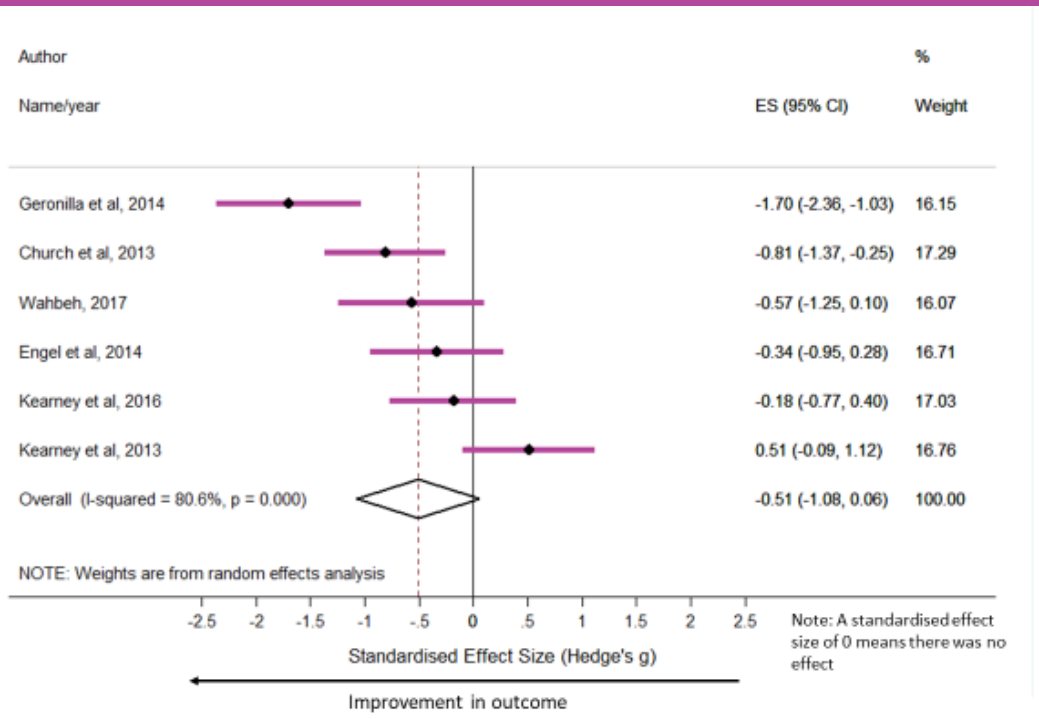
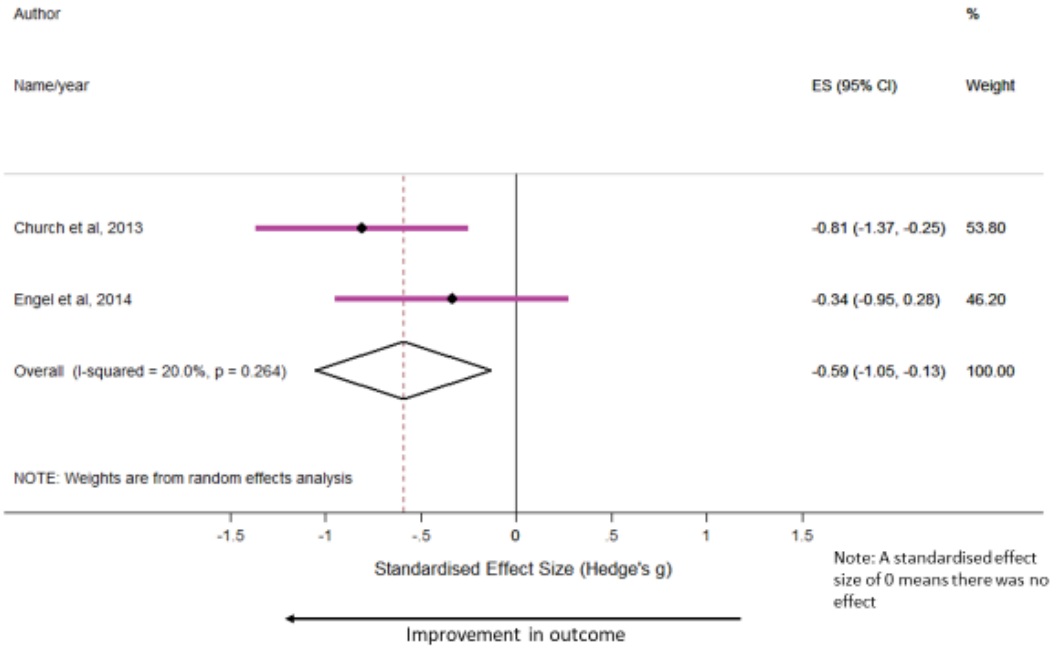


Figure A 5.56 The effect of meditation and mindfulness interventions on other mental health issues. Sensitivity analysis : low risk of bias



5.2.7 Other therapeutic wellbeing interventions

Figure A 5.57 The effect of other therapeutic wellbeing interventions on stress and disorders associated with stress, including PTSD

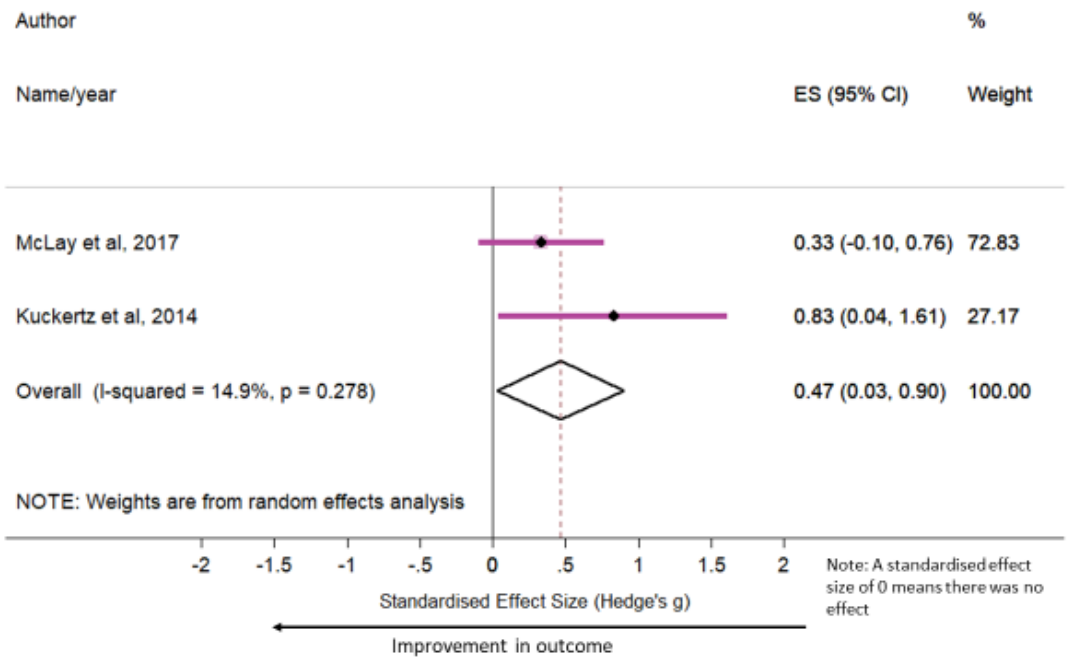
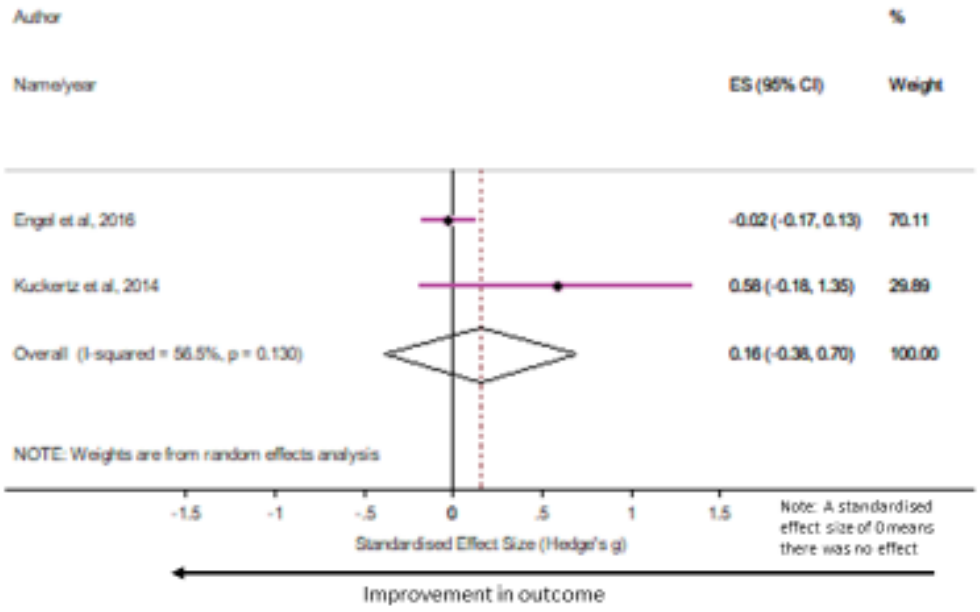


Figure A 5.58 The effect of other therapeutic wellbeing interventions on mood disorders



5.2.8 Medication

Figure A 5.59 The effect of medication on physical health and wellbeing

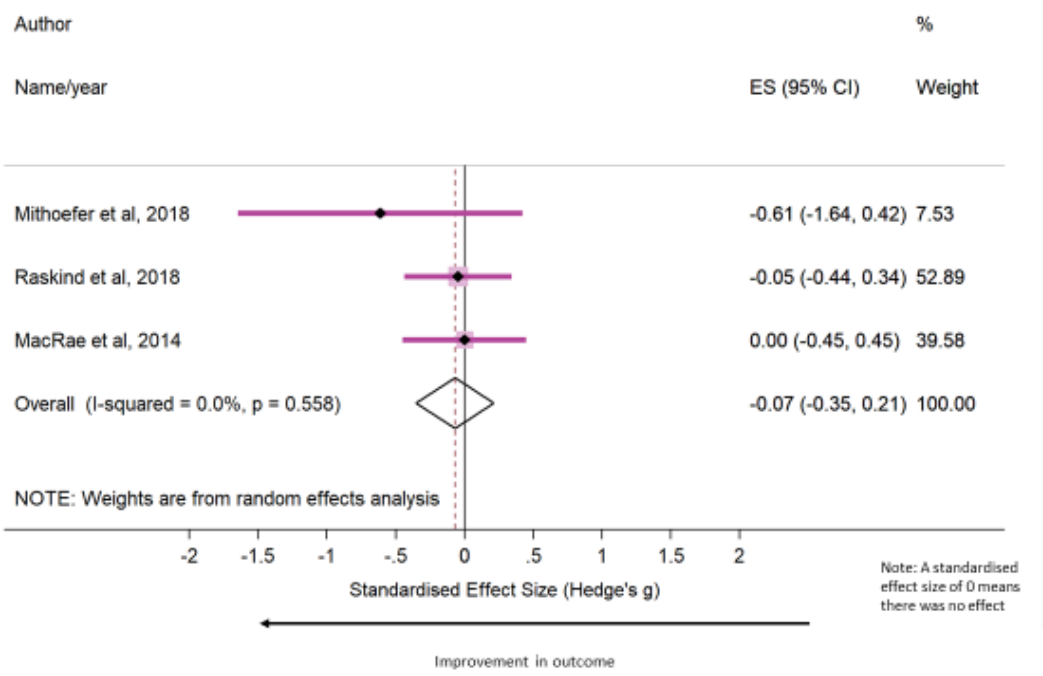


Figure A 5.60 The effect of medication on stress and disorders associated with stress including PTSD

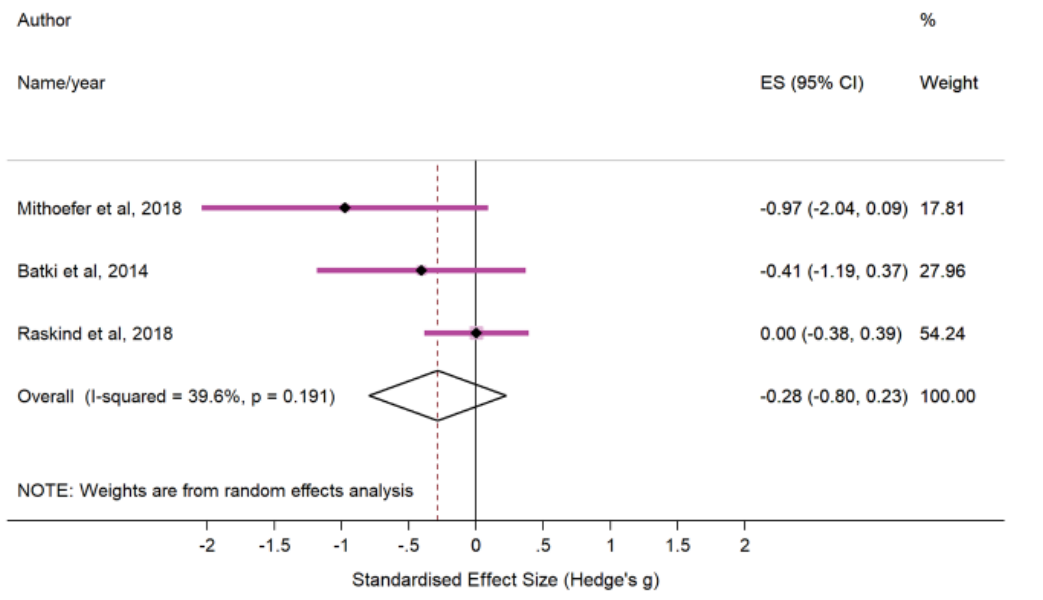


Figure A 5.61 The effect of medication on stress and disorders associated with stress including PTSD : veterans

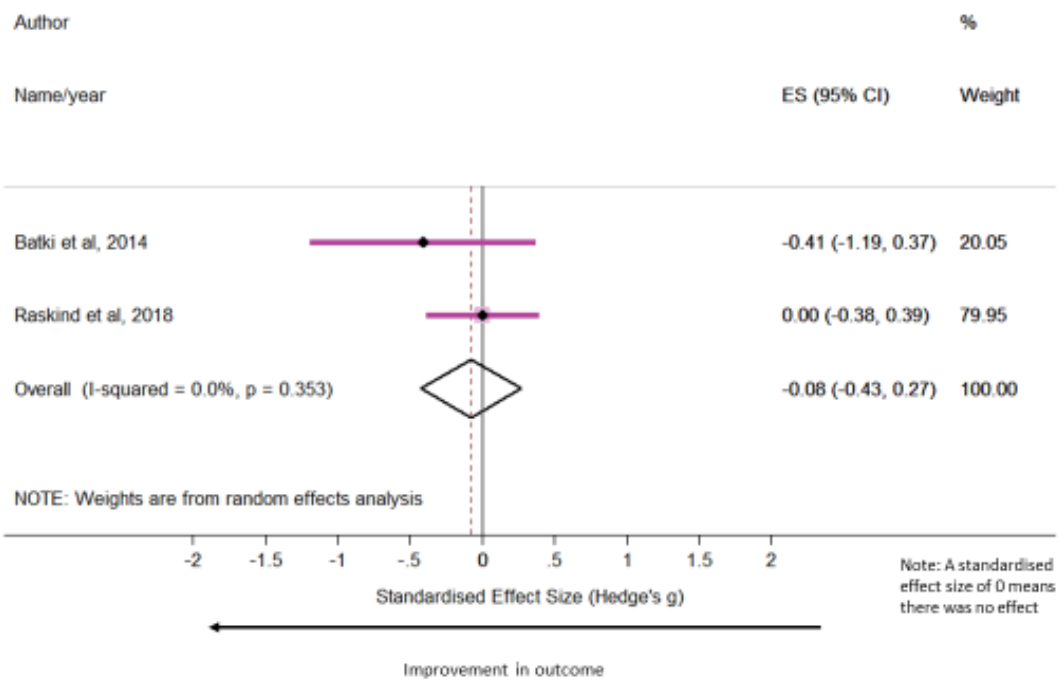


Figure A 5.62 The effect of medication on mood disorders

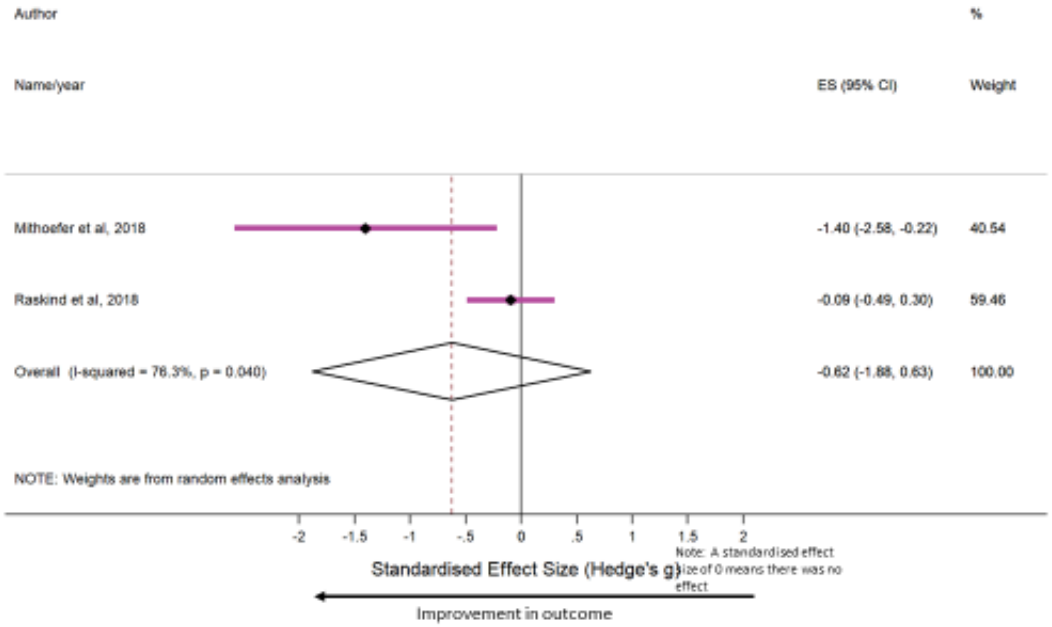


Figure A 5.63 The effect of medication on substance misuse

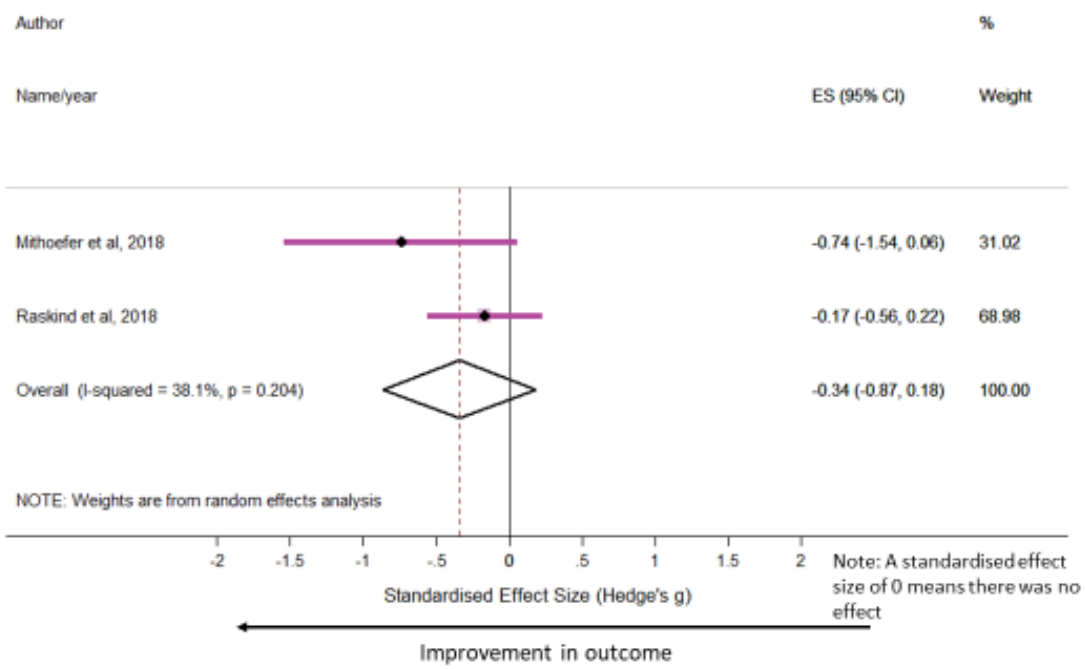
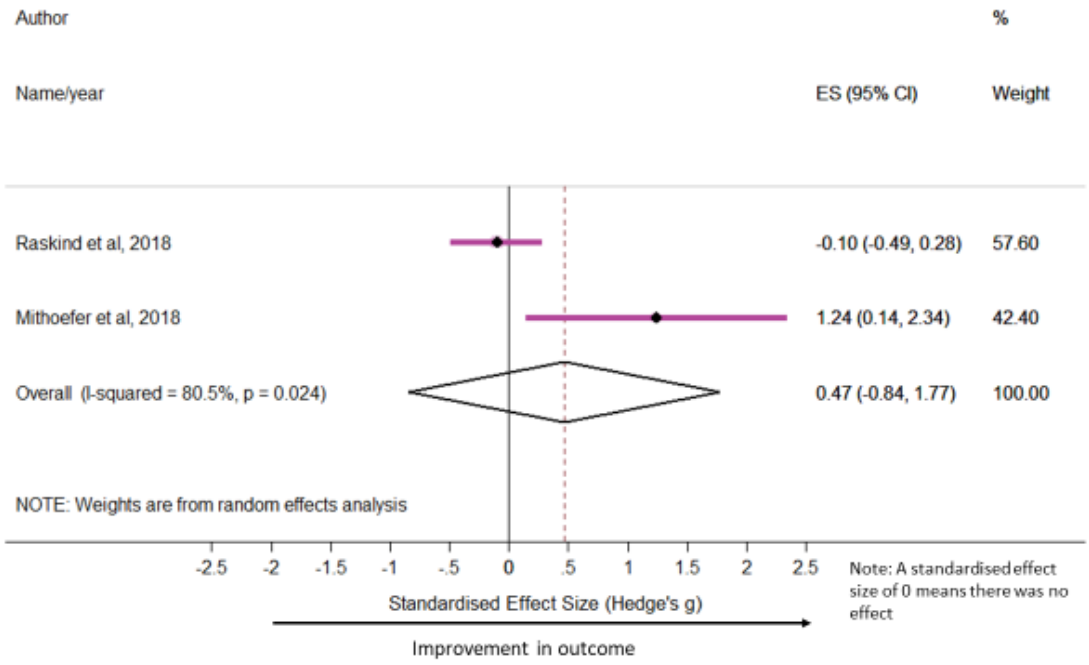


Figure A 5.64 The effect of medication on other mental health issues



5.2.9 Interventions addressing substance misuse and gambling

Figure A 5.65 The effect of interventions addressing substance misuse on disorders due to substance misuse or addictive behaviours

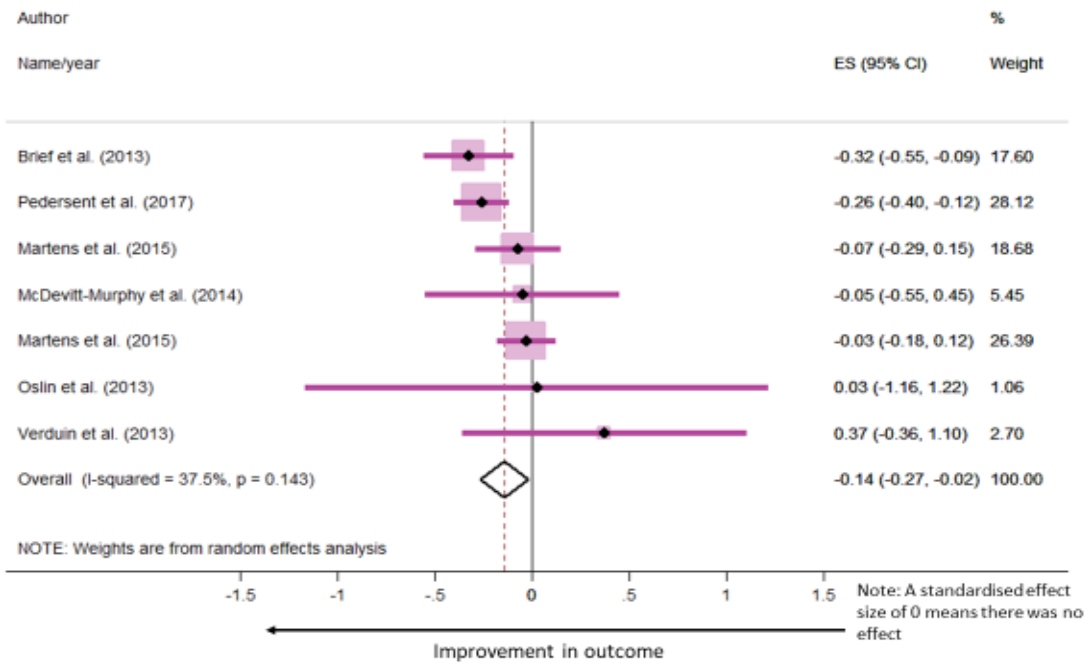
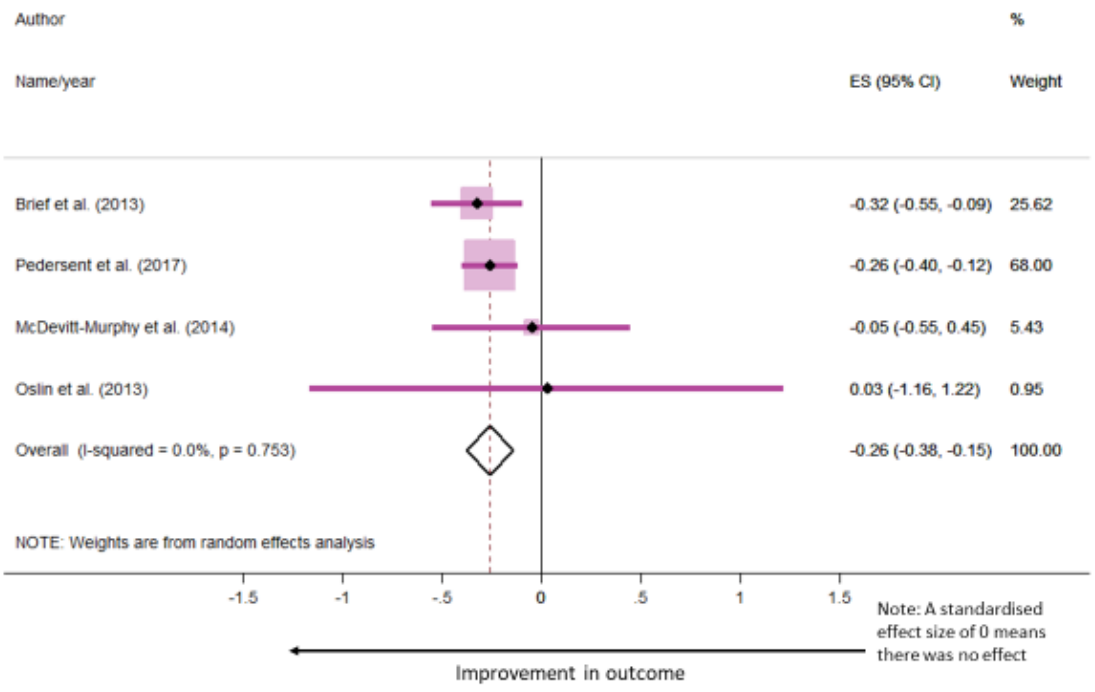
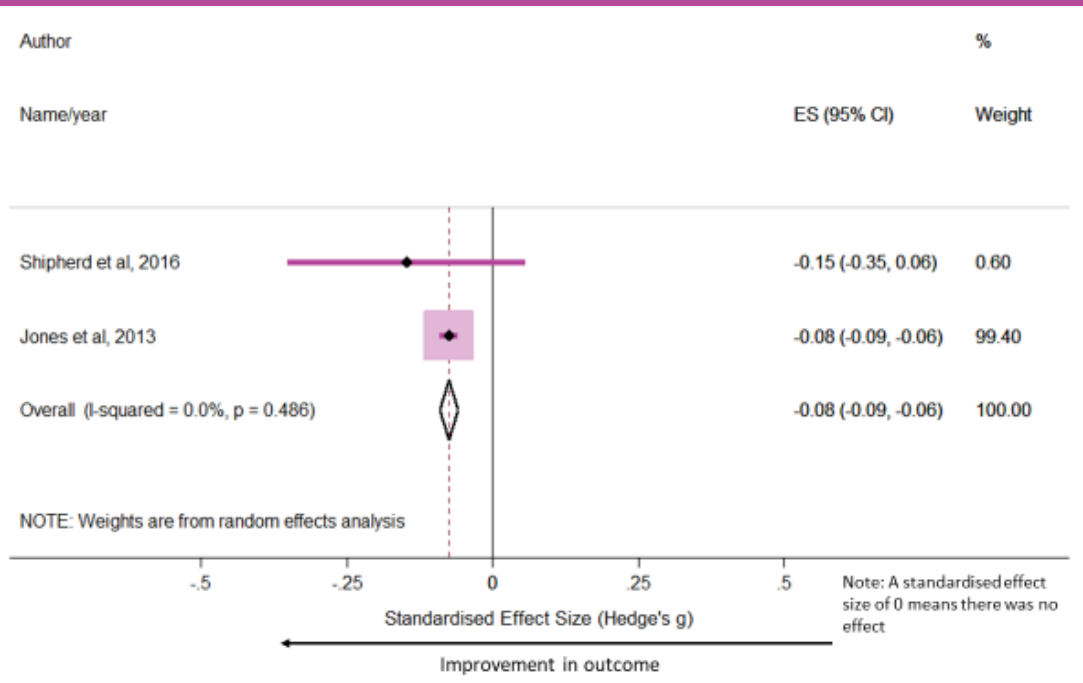


Figure A 5.66 The effect of interventions addressing substance misuse on disorders due to substance misuse or addictive behaviours. Sensitivity analysis : low risk of bias



5.2.10 Resettlement interventions

Figure A 5.67 The effect of resettlement interventions on stress and disorders associated with stress including PTSD



5.2.11 Advice and support interventions

Figure A 5.68 The effect of advice and support interventions on physical health and wellbeing

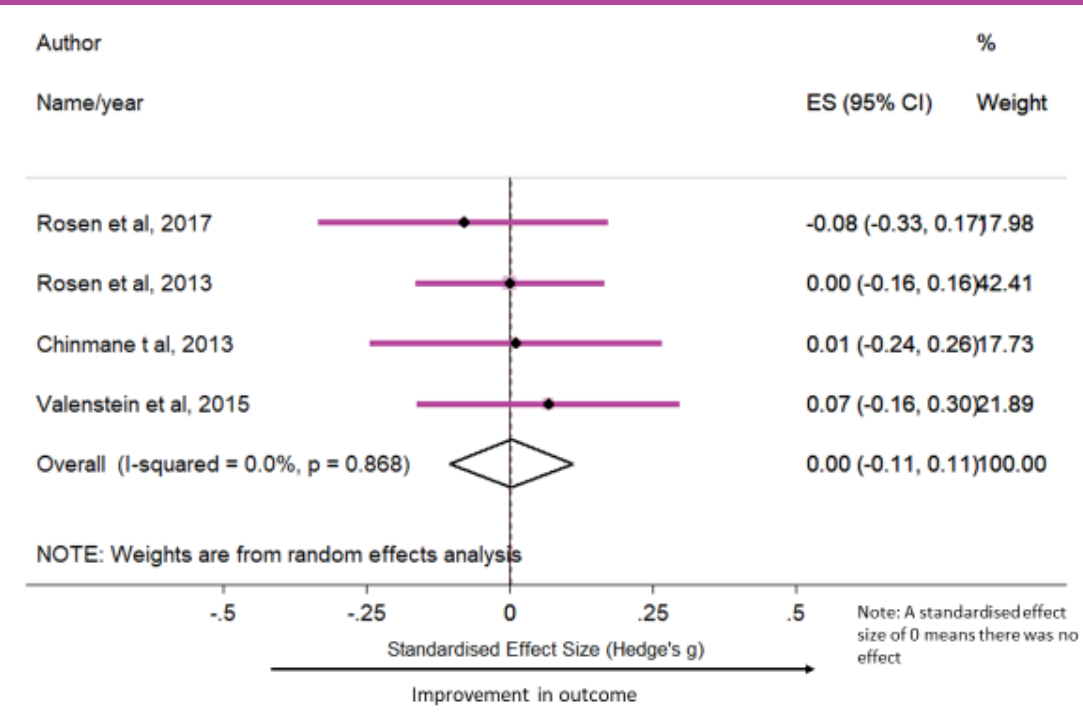


Figure A 5.69 The effect of advice and support interventions on physical health and wellbeing. Sensitivity analysis : low risk of bias

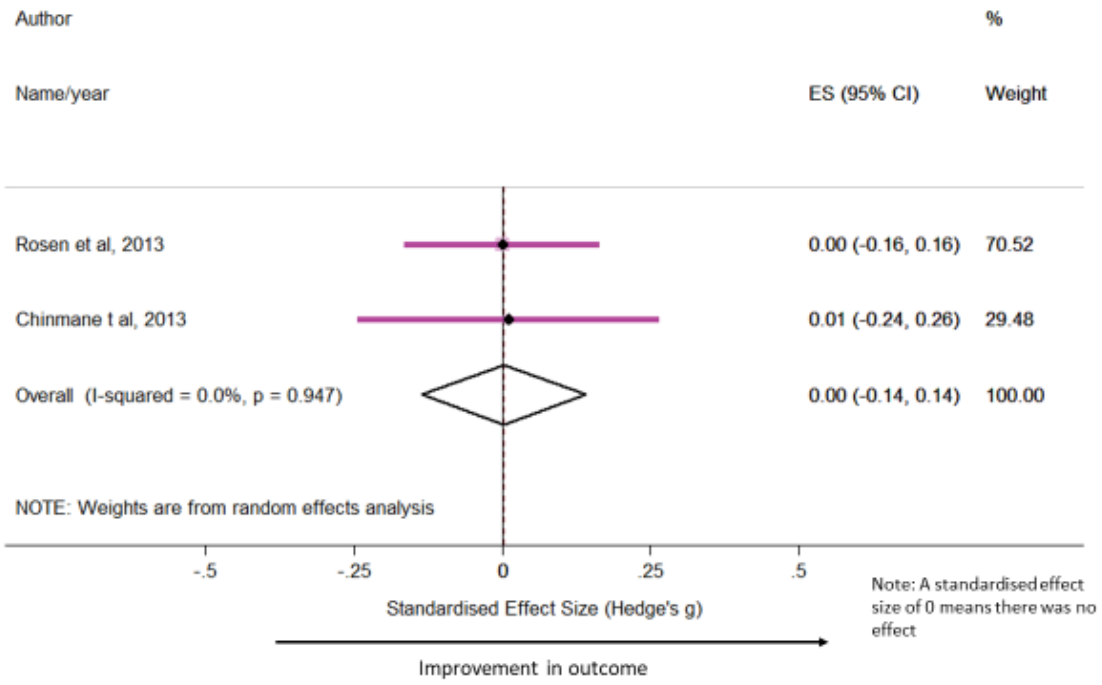


Figure A 5.70 The effect of advice and support interventions on physical health and wellbeing. Sensitivity analysis : 12-month follow-up period

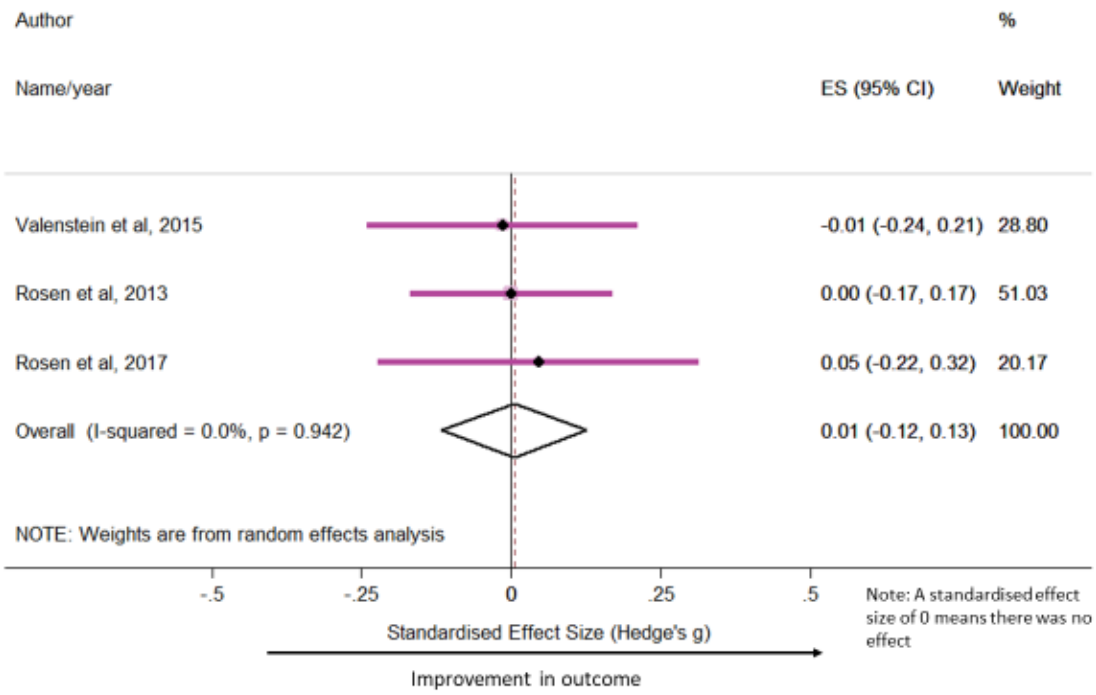


Figure A 5.71 The effect of advice and support interventions on stress and disorders associated with stress including PTSD

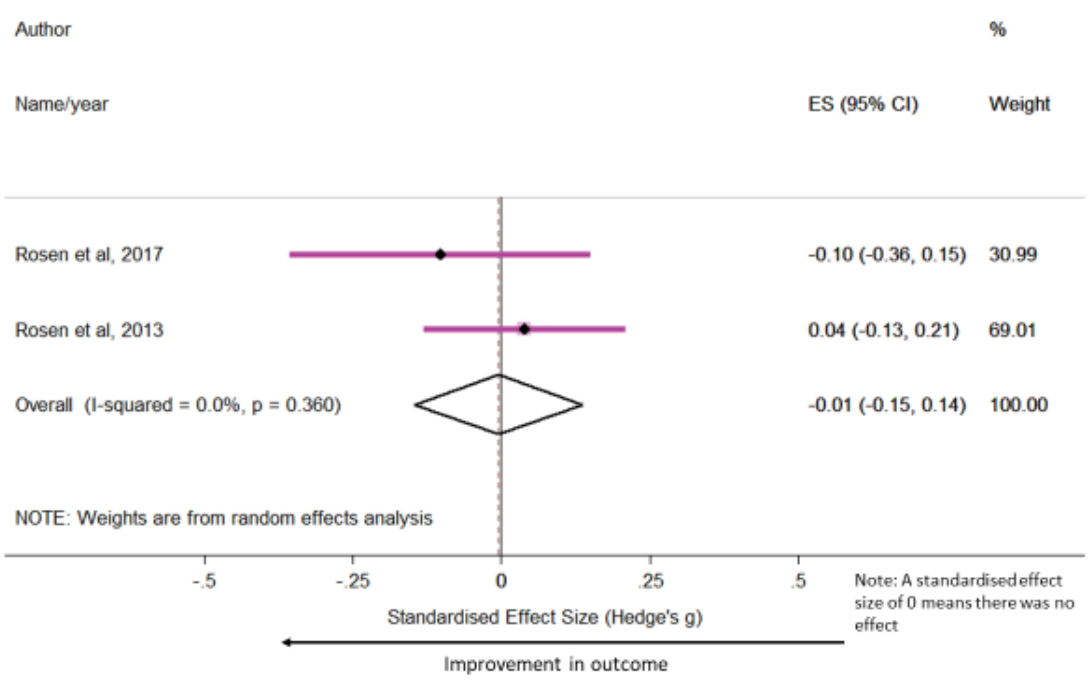


Figure A 5.72 The effect of advice and support interventions on mood disorders

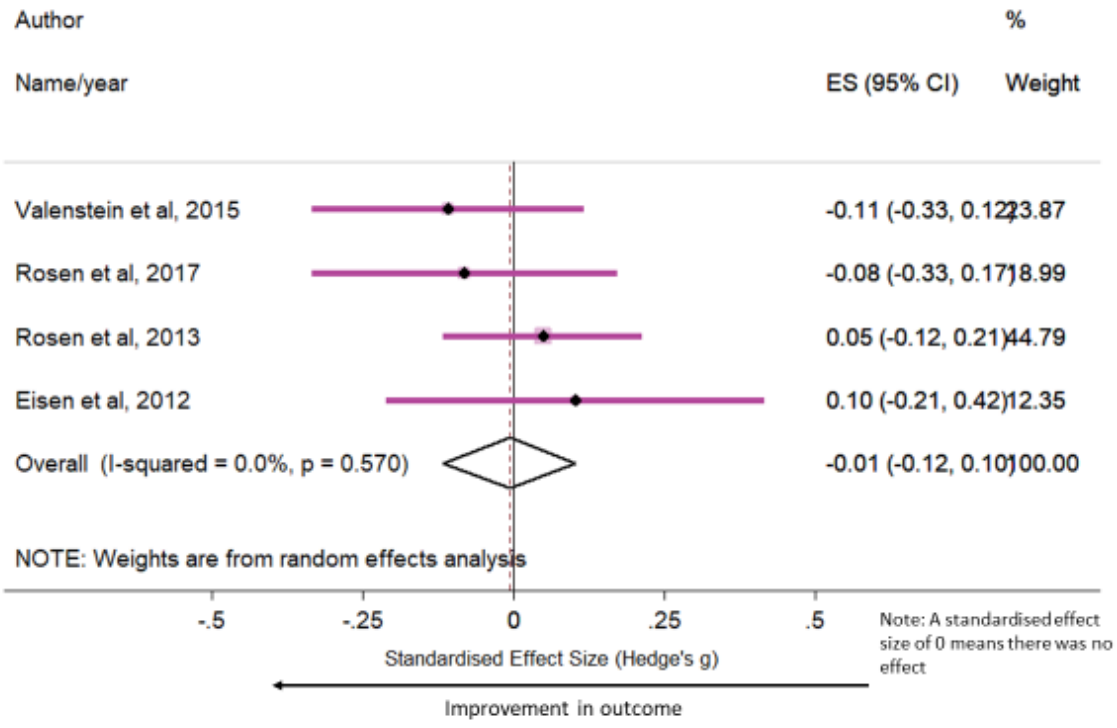


Figure A 5.73 The effect of advice and support interventions on stress and mood disorders. Sensitivity analysis : 12 month follow-up period

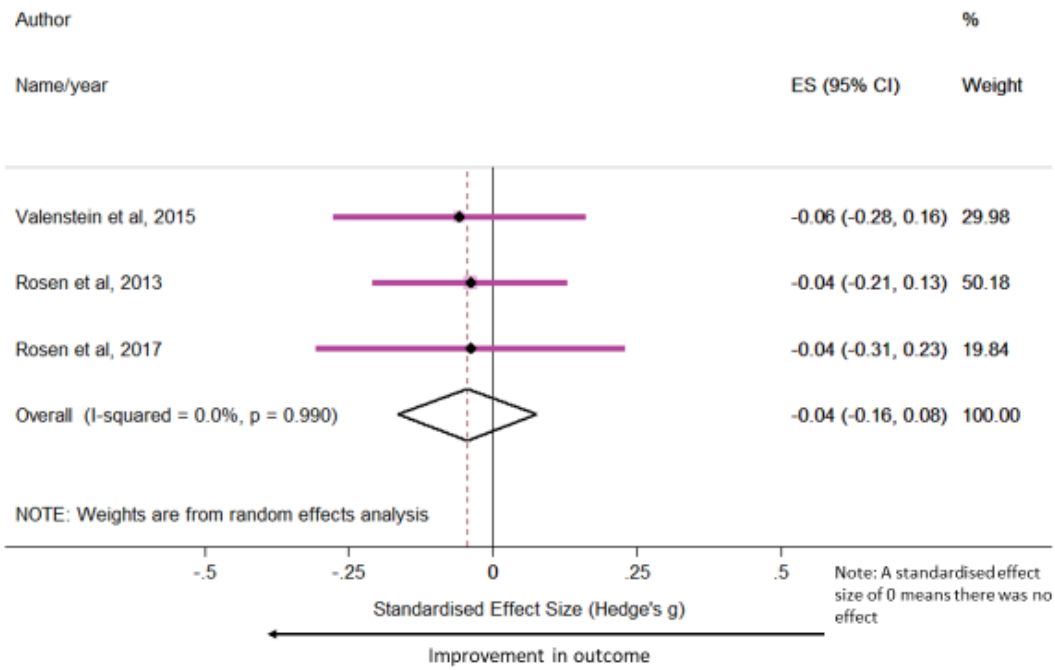


Figure A 5.74 The effect of advice and support interventions on substance misuse

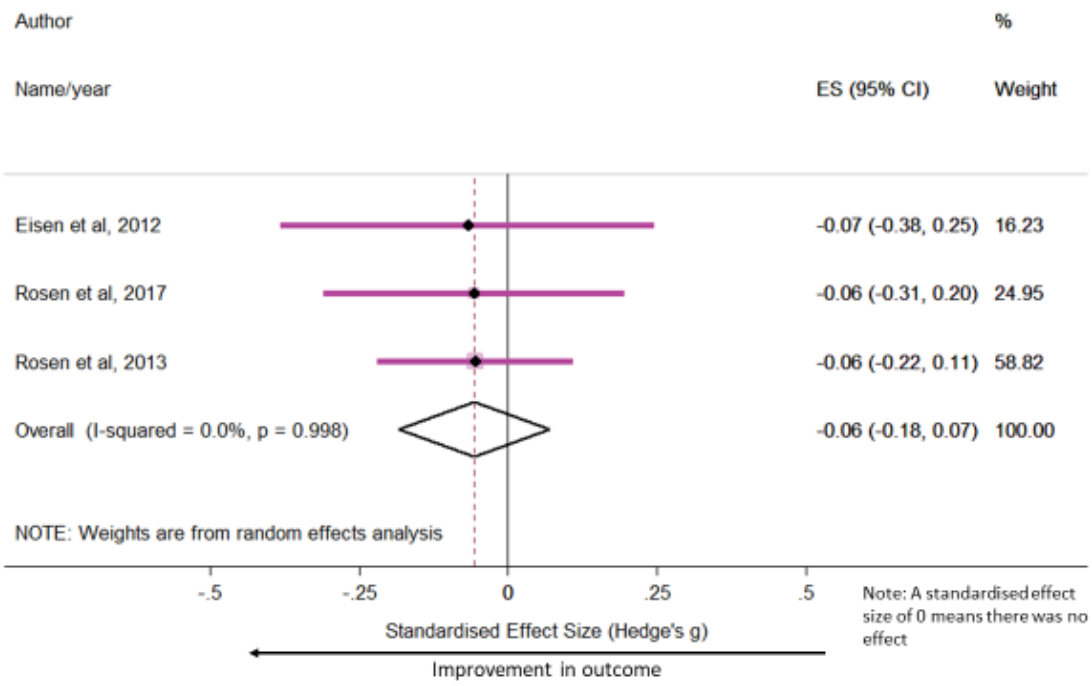


Figure A 5.75 The effect of advice and support interventions on suicide and self-harm

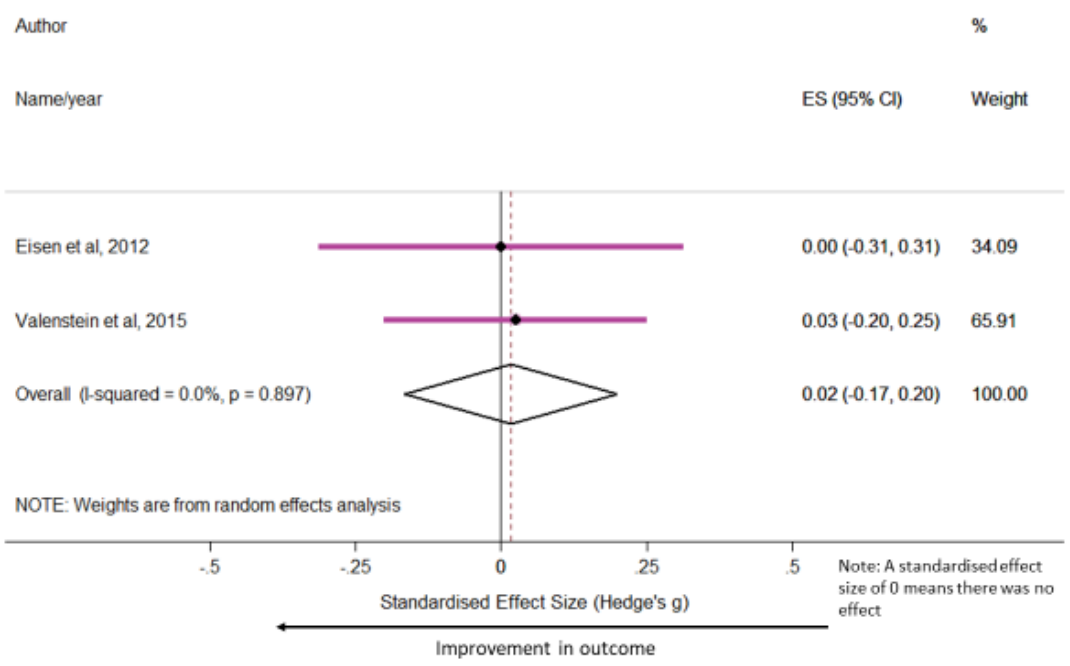


Figure A 5.76 The effect of advice and support interventions on substance misuse. Sensitivity analysis : 12 month follow-up period

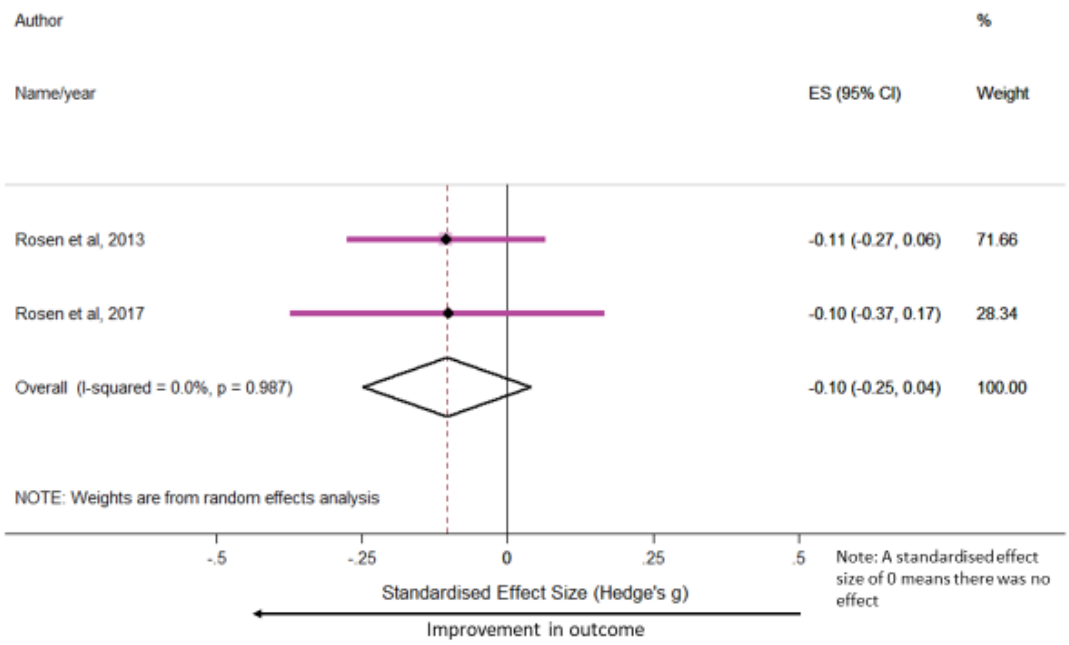
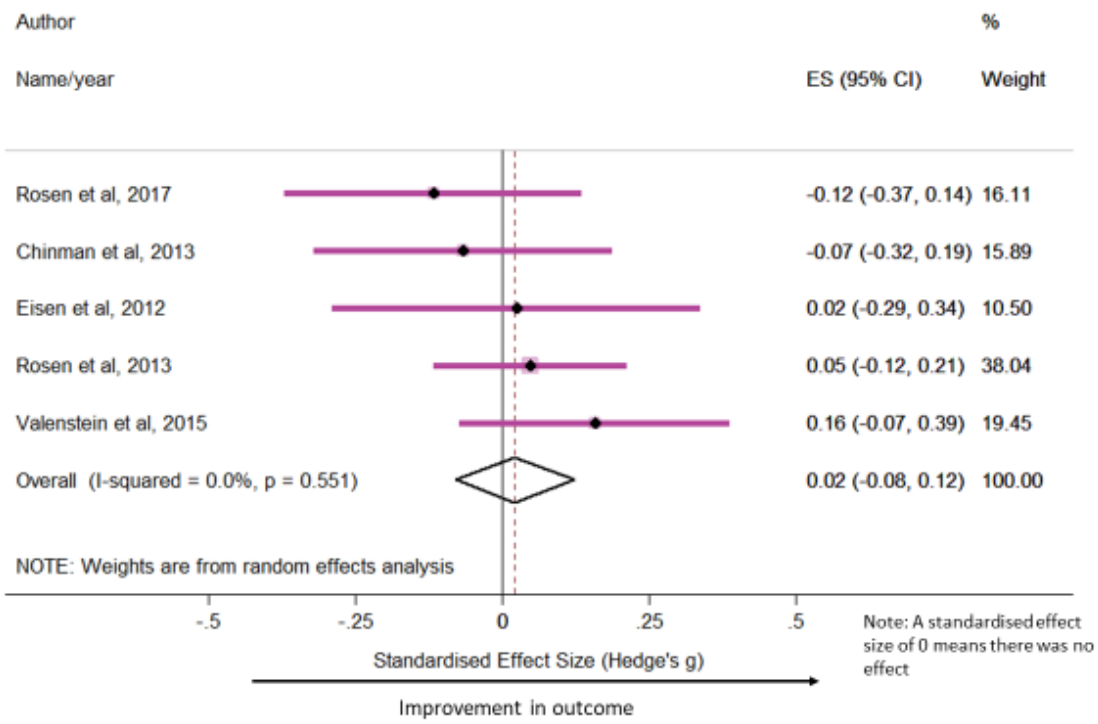


Figure A 5.77 The effect of advice and support interventions on other mental health outcomes



5.3 Sub-group and sensitivity analysis

This section reports the meta-analyses findings for each outcome along with any sub-group or sensitivity analysis that has been conducted. Studies that have not been included within the meta-analyses are also discussed.

5.3.1 Awareness, screening and prevention: other findings

Stress and associated disorders including PTSD

In addition to the two studies included in the meta-analysis, a third study by Mulligan et al. (2011) also looks at the effect of awareness, screening and prevention interventions on stress and associated disorders including PTSD. This study has been excluded from the analysis as insufficient information was reported to calculate the effect sizes. Mulligan et al. (2011) report no statistically significant differences in PTSD scores between military personnel receiving a post-deployment self-help and coping strategy intervention and those receiving a standard post-deployment brief.

5.3.2 CBT: other findings

This section explores the results of the meta-analyses individually for each outcome, alongside a discussion of studies that could not be included in the meta-analysis. Where possible, sub-group and sensitivity analyses have been conducted and these results are also discussed in this section.

Physical health and wellbeing

Two studies assessed the impact of CBT on individuals' sleep quality. The results of the meta-analysis shows that CBT can reduces the likelihood of sleep disturbance by 24.6% (SMD -1.13). Castillo et al. (2016) analysed physical health and quality of life, but these outcome constructs were considered distinct from the two studies assessing sleep quality and therefore wasn't included in this meta-analysis. Castillo et al. (2016) found that CBT significantly improved physical health (SMD 0.94 [0.43, 1.45]).

Stress and associated disorders including PTSD

There is a large evidence base for the effectiveness of CBT on stress and its associated disorders, primarily PTSD. All 22 studies are drawn from the U.S. population. Sub-group analysis that consider the outcomes of serving regulars indicated a much stronger impact from CBT, though this draws on just four studies and is not statistically significant. A sub-group analysis of veterans found smaller impacts than the primary analysis, which was also statistically insignificant.

Table A 5.12 The effect of CBT on stress and associated disorders including PTSD

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.80 [0.15, 1.44]	23	299.27 [0.000]	92.6	1.704	18.5
Sub-group analysis						
Serving regulars	3.32 [-2.89, 9.54]	4	26.94 [0.000]	88.9	28.188	42.8
Veterans	0.44 [-0.00, 0.88]	16	86.88 [0.377]	6.1	0.006	10.7
Sensitivity analysis						
Six-month follow-up period	0.16 [-0.09, 0.41]	6	5.33 [0.377]	6.1	0.006	4.0
Low risk of bias	0.52 [0.12, 0.91]	9	23.53 [0.003]	66.0	0.149	12.5

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Sensitivity analysis observing outcomes at a later follow-up period found no statistically significant effect, though this also used a much smaller sample of studies than the primary analysis. Sensitivity analysis including only studies at low risk of bias confirms the result of the primary analysis, indicating that CBT is associated with a 12.5% reduction in PTSD (SMD 0.52).

One study assessing the effectiveness of CBT on stress and associated disorders including PTSD was not included in the meta-analysis. Ziemba et al. (2014) did not report significance testing and hence the confidence intervals could not be estimated. This study found very similar PTSD (CAPS) scores for both treatment and control and no significance testing could be undertaken with the reported information.

Mood disorders

In total, thirteen studies assessed the effectiveness of CBT on mood disorders (typically depression). The evidence shows that CBT is an effective intervention to

reduce the incidence of mood disorders, by 10.3% (SMD 0.42) though this is not statistically significant.

Table A 5.13 The effect of CBT on mood disorders

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Mood disorders	0.42 [-0.03, 0.86]	13	48.22 [0.000]	75.1	0.364	10.3
Sub-group analysis						
Serving regulars	0.21 [-0.09, 0.51]	3	0.10 [0.952]	0.0	0.000	5.2
Veterans	0.58 [0.13, 1.03]	9	24.94 [0.002]	67.9	0.265	13.9
Sensitivity analysis						
Six-month follow-up period	0.70 [-0.66, 2.06]	4	69.32 [0.000]	95.7	1.514	16.5
Low risk of bias	0.56 [-0.08, 1.20]	8	37.71 [0.000]	81.4	0.455	13.5

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Table A 5.13 shows the results of two sub-group analyses, exploring the effectiveness of CBT for serving regulars and veterans separately. Three studies analysed serving regulars, finding no effect of CBT on mood disorders. Conversely, the studies analysing the veteran population found a significant reduction in mood disorders, by 13.9% (SMD 0.58).

In addition, two sensitivity analyses were conducted. The first used a later follow-up period (six months), which indicates the intervention may have a positive impact on mood disorders (SMD 0.70) though this is not statistically significant. The latter included only studies at low risk of bias and was relatively consistent with the primary analysis.

A single study (Ziemba et al., 2014) could not be included in the meta-analysis as the study did not present significance testing. Ziemba et al. (2014) tested tele-CBT with face-to-face and found both groups experienced a reduction in depression, though a comparison between groups cannot be made as there is insufficient information presented in the study.

Anxiety and fear

The effectiveness of CBT on mediating anxiety and fear was analysed in three studies. The available evidence finds a positive but statistically insignificant impact and sub-group analysis of the two studies focused solely on the veteran population finds consistent results.

Table A 5.14 The effect of CBT on anxiety and fear

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.29 [-0.02, 0.59]	3	0.44 [0.804]	0.0	0.000	7.1
Sub-group analysis						
Veterans	0.31 [-0.13, 0.76]	2	0.41 [0.523]	0.0	0.000	7.8

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

As with mood disorders, the study by Ziemba et al. (2014) could not be included as statistical significance testing was not presented. This study did not find any differences in anxiety after CBT.

Other mental health outcomes
The effectiveness of CBT was also assessed against other mental health outcomes. These were: cognitions (about themselves, the world and self-blame), hopelessness and perceived mental health[8].

Table A 5.15 displays the results of the meta-analysis and a sub-group analysis of studies concerned with the veteran population. The results of both analyses find a slightly stronger impact of CBT on other mental health outcomes of 11.3% (SMD 0.47).

Table A 5.15 The effect of CBT on other mental health outcomes

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.31 [-0.53, 1.15]	3	7.97 [0.019]	74.9	0.019	7.6
Veterans	0.47 [-0.63, 1.56]	2	6.29 [0.012]	84.1	0.522	11.3

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

5.3.3 Wellbeing interventions: other findings

This section details the results of the meta-analyses individually for each outcome as well as any further analyses carried out in addition to the primary analyses. Outcomes not included in the meta-analyses are also discussed.

Physical health and wellbeing
The evidence shows that wellbeing interventions may be effective for treating physical health and wellbeing outcomes, although this estimate is not statistically significant.

[8] Note that to estimate the standard mean difference, the scale of perceived mental health was inverted to be consistent with other outcomes.

A subgroup analysis was conducted using studies that only included ex-Service personnel. This analysis shows that wellbeing interventions are effective at improving physical health and wellbeing amongst veterans (SMD 0.33 or an improvement of 8.1%).

Table A 5.16 The effect of wellbeing interventions on physical health and wellbeing

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.20 [-0.06, 0.50]	3	6.80 [0.033]	70.6	0.031	4.9
Sub-group analysis						
Veterans	0.33 [0.19, 0.47]	2	0.001 [0.941]	0.0	0.000	8.1

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Stress and associated disorders including PTSD

Of the 6 studies included in this meta-analysis, all assessed PTSD outcomes. The primary meta-analysis produces no evidence to indicate that wellbeing interventions are effective at reducing stress and associated disorders including PTSD.

However, the subgroup analysis of veterans indicates that wellbeing interventions may be effective for this group. Stress is reduced by 4.9% (SMD 0.20) for veterans taking part in wellbeing interventions, though this result is not statistically significant.

Table A 5.17 The effect of wellbeing interventions on stress and associated disorders including PTSD

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	-0.01 [-0.24, 0.23]	6	12.25 [0.031]	59.2	0.038	-0.2
Sub-group analysis						
Veterans	-0.20 [-0.61, 0.21]	5	12.13 [0.016]	67.0	0.128	-4.9

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Other mental health outcomes

Outcomes included for this category include anger, emotional regulation and hope. The primary analysis does not indicate that these outcomes are likely to benefit from wellbeing interventions.

A subgroup analysis indicates that wellbeing interventions are detrimental to veteran's other mental health outcomes, though this is not statistically significant. Sensitivity analysis assessing outcomes after three months is consistent with the primary analysis.

Table A 5.18 The effect of wellbeing interventions on other mental health outcomes

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.03 [-0.24, 0.29]	6	15.45 [0.001]	67.6	0.055	0.7
Sub-group analysis						
Veterans	-0.14 [-0.60, 0.32]	5	15.18 [0.004	73.6	0.177	-3.5
Sensitivity analysis						
3 months	-0.06 [-0.61, 0.49]	3	7.25 [0.027]	72.4	0.161	-1.5

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

5.3.4 Meditation and mindfulness: other findings

This section explores the results of the meta-analyses individually for each outcome, alongside a discussion of studies that could not be included in the meta-analysis. Where possible, sub-group and sensitivity analyses have been conducted and these results are also discussed in this section.

Physical health and wellbeing

Five studies assessed the effectiveness of meditation and mindfulness interventions of improving physical health and wellbeing. A single sub-group analysis, looking solely at the veteran population found results consistent with the primary meta-analysis.

Table A 5.19 The effect of meditation and mindfulness on physical health and wellbeing: sub-group and sensitivity analyses

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.58 [0.24, 0.91]	5	6.58 [0.160]	39.2	0.056	13.8
Sub-group analysis						
Veterans	0.58 [0.17, 1.00]	4	6.24 [0.101]	51.9	0.092	14.0

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Stress and associated disorders including PTSD

Fourteen studies investigated how effective meditation and mindfulness interventions were at reducing stress and associated disorders including PTSD. There is evidence that meditation and mindfulness interventions can reduce stress.

Table A 5.20 The effect of meditation and mindfulness on stress and associated disorders including PTSD: sub-group and sensitivity analyses

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.74 [0.25, 1.23]	14	126.58 [0.000]	89.7	0.627	17.4
Sub-group analyses						
Veterans	0.71 [0.18, 1.24]	13	123.91 [0.000]	90.3	0.673	16.7
Sensitivity analyses						
Two-month follow-up period	0.53 [-0.06, 1.11]	4	21.88 [0.000]	86.3	0.303	12.7
Low risk of bias	0.95 [0.18, 1.71]	4	26.03 [0.000]	88.5	0.528	21.4

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Most of the evidence concerns the veteran population and a sub-group analysis including only these studies confirms the result of the primary analysis. The most common follow-up period used in the primary analysis was post-intervention (i.e. as soon as the intervention had finished). Sensitivity analysis finds a reduced and statistically insignificant effect (SMD 0.52) two-months after interventions ended, though this analysis only included four studies. Finally, a further sensitivity analysis, containing all four studies that had a low risk of bias, confirms the result of the primary analysis.

Mood disorders

Ten studies investigated the effectiveness of meditation and mindfulness interventions on mood disorders, of which eight concerned the veteran population. Sub-group analysis solely of veterans found a consistent result with the primary analysis. Most studies analysed outcomes at the end of the intervention, though some studies recorded outcomes at multiple time points. A sensitivity analysis of the impact estimates two-months after interventions ended is consistent with the primary analysis.

Table A 5.21 The effect of meditation and mindfulness on mood disorders

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.42 [0.26, 0.58]	10	6.17 [0.723]	0.0	0.000	10.3
Sub-group analysis						
Veterans	0.37 [0.19, 0.54]	8	6.50 [0.482]	0.0	0.000	9.0

Sensitivity analysis						
Two-month follow-up period	0.53 [0.26, 0.79]	3	1.26 [0.532]	0.0	0.000	12.7
Low risk of bias	0.46 [0.14, 0.79]	3	1.63 [0.443]	0.0	0.000	11.3

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Anxiety and fear

The primary analysis of four studies, shows that meditation and mindfulness can reduce anxiety and fear (SMD 0.73 or a reduction of 17.2%). Two of these studies were at low risk of bias and a sensitivity analysis including only these studies produced a similar standardised mean difference, though this was not statistically significant.

Table A 5.22 The effect of meditation and mindfulness on anxiety and fear

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.73 [0.03, 1.44]	4	18.36 [0.000]	83.7	0.418	17.2
Sensitivity analysis						
Low risk of bias	0.81 [-0.03, 1.66]	2	2.54 [0.111]	60.6	0.232	18.8

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Other mental health outcomes

Six studies of meditation and mindfulness interventions assessed other mental health outcomes. These outcomes included general mental health, hostility, negative affect and fatigue. The analysis indicates that meditation and mindfulness can improve other mental health outcomes, though the result is not statistically significant.

Table A 5.23 The effect of meditation and mindfulness on other mental health outcomes

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.51 [-0.06, 1.08]	6	25.76 [0.000]	80.6	0.407	12.3
Sub-group analysis						
Veterans	0.54 [-0.15, 1.24]	5	25.50 [0.000]	84.3	0.527	13.1
Sensitivity analysis						

Low risk of bias	0.59 [0.13, 1.05]	2	1.25 [0.264]	20.0	0.022	14.2

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Five of the six studies analysed veterans. A sub-group analysis of these studies produced a consistent estimate to the primary analysis. Additionally, a sensitivity analysis including the two studies at low risk of bias found a statistically significant positive impact of meditation and mindfulness on other mental health outcomes.

5.3.5 Interventions for substance misuse and gambling: other findings

This section explores the results of the meta-analyses individually for each outcome, alongside a discussion of studies that could not be included in the meta-analysis. Where possible, sub-group and sensitivity analyses have been conducted and these results are also discussed in this section.

Substance use or addiction

Seven studies assessed the effectiveness of interventions for substance use & gambling on reducing substance use or addiction. A sensitivity analysis, looking solely at studies with low risk of bias, found results consistent with the primary meta-analysis. When only including studies with low risk of bias, the impact of the intervention is larger; a reduction of 6.5% (SMD 0.26).

Table A 5.24 The effect of substance misuse and gambling interventions on physical health and wellbeing: sub-group and sensitivity analyses

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.14 [0.02, 0.27]	7	9.60 [0.143]	37.5	0.009	3.6
Sensitivity analysis						
Low risk of bias studies	0.26 [0.15, 0.38]	4	1.20 [0.753]	0.0	0.000	6.5

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

5.3.6 Advice and support: other findings

This section explores the results of the meta-analyses individually for each outcome, alongside a discussion of studies that could not be included in the meta-analysis. Where possible, sub-group and sensitivity analyses have been conducted and these results are also discussed in this section.

Quality of life

Four studies assess the effectiveness of advice and support interventions on improving quality of life. Two sensitivity analyses were conducted. One looked at a 12-month follow-up period using three studies and find no statistically significant effect, consistent

results with the primary meta-analysis. Similar results are found when looking at studies with low risk of bias.

Table A 5.25 The effect of advice and support on quality of life: sub-group and sensitivity analyses

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.00 [-0.11, 0.11]	4	0.72 [0.868]	0.0	0.000	0.0
Sensitivity analysis						
12-months follow up period	0.01 [-0.12, 0.13]	3	0.12 [0.942]	0.0	0.000	0.1
Low risk of bias	0.00 [-0.14, 0.14]	2	0.00 [0.947]	0.0	0.000	0.1

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Mood disorders

Four studies investigated how effective advice and support interventions are at reducing depression. There is no evidence that the interventions can reduce depression. One sensitivity analysis using a 12-months follow-up period was conducted. The sensitivity analysis finds no significant effect of the intervention at 12-month post-intervention, consistent with findings from the primary analysis.

Table A 5.26 The effect of advice and support on mood disorders: sensitivity analysis

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.00 [-0.10, 0.12]	4	2.01 [0.570]	0.0	0.000	0.2
Sensitivity analysis						
12-month follow-up period	0.05 [-0.08, 0.17]	3	0.02 [0.990]	0.0	0.000	1.1

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Substance use or addiction

The primary analysis is based on three studies and shows no evidence that advice and support can reduce substance use or addiction. Two of these also had results for 12-months follow-up and a sensitivity analysis including only these studies produced a similar non-significant standardised mean difference.

Table A 5.27 The effect of advice and support on anxiety and fear

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.06 [-0.07, 0.18]	3	0.00 [0.998]	0.0	0.000	1.4
Sensitivity analysis						
12-month follow-up period	0.10 [-0.04, 0.25]	2	0.00 [0.987]	0.0	0.000	2.6

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

Other mental health outcomes

Five studies of advice and support interventions assessed other mental health outcomes. These outcomes included aggressivity, general mental health, interpersonal relationship.

Table A 5.28 The effect of advice and support on other mental health outcomes

Analysis	Standard mean difference [95% CI]	Number of studies	Q [P-value]	I^2 (%)	Tau^2	Standardised percentage (%)
Primary analysis	0.02 [-0.08, 0.12]	5	3.04 [0.551]	0.0	0.000	0.5
Sensitivity analysis						
12-months follow-up period	0.02 [-0.10, 0.14]	3	0.47 [0.792]	0.0	0.000	0.4
Low risk of bias	0.01 [-0.12,0.15]	2	0.54 [0.462]	0.0	0.000	0.4

Note that the sign of standardised mean differences may be switched in tables for comparability across outcomes.

The primary analysis found no evidence of impact and this was reflected in the sub-group and sensitivity analyses. Three out of five studies had collected results at 12-months post-intervention. Therefore, a sensitivity analysis using studies providing this time point only was conducted. It found similar non-significant results to the primary analysis. Two studies out of five had low risk of bias and a sensitivity analysis using only those studies found similar non-significant results as well.

Outcomes not included in meta-analyses
A single study one the effectiveness of advice and support interventions was not included in the meta-analysis as it was the only study with delivery as an outcome. The study by Rosen et al. (2017) assessed the impact of an added telephone care management on top of usual outpatient mental health care for U.S. veterans. The study found that the intervention can increase the attendance to sessions (0.26 [0.05, 0.47]) at three-month follow-up.

5.4 Publication bias

If the publication or non-publication of a study depends on the nature or direction of its results, this can introduce bias to a meta-analysis. To assess if publication bias affects the meta-analyses in this review the "tandem procedure" outlined in Ferguson and Brannick (2012) was implemented.

The tandem procedure consists of three statistical tests that are individual tests of publication bias. These tests are:

- Orwin's fail-safe N: This indicates the number of "fail-safe" or unpublished studies finding no effect that would need to be included in the meta-analysis for the result to change the standardised mean difference to a "trivial" value. If this number is greater than the number of studies in the meta-analysis it is an indication of bias.
- Egger's test: A "funnel plot" is the concordance between the precision of each study and the size of the standardised mean difference. This will produce a symmetrical distribution in the absence of publication bias. Egger's test is a regression that tests for asymmetry. If the results of the regression are statistically significant, this indicates publication bias.
- The trim and fill procedure: This procedure first trims studies from the meta-analysis that fall outside of the symmetry in the funnel plot and imputes estimates of the standardised mean difference and study precision for studies "missing" from the meta-analysis. If the pooled standardised mean difference produced by this procedure is statistically significantly different to the pooled standardised mean difference from the primary meta-analysis, this indicates publication bias.

Under the tandem procedure, a meta-analysis is considered to suffer from publication bias if multiple tests indicate publication bias. The results indicate that publication bias does not impact upon any of the primary analyses in this review, as no meta-analysis failed more than one of these tests.

Table A 5.29 and Table A 5.30 present the meta-analyses that failed each of the tests. No meta-analysis failed the "trim and fill" procedure and hence these results are not reported.

Table A 5.29 Meta-analyses failing Orwin's fail-safe N test

Intervention	Outcome	Number of studies	Orwin's FSN
Behavioural Activation Therapy	Mood disorders	2	2
Cognitive Behavioural Therapy	Satisfaction with care	2	44
	Physical health and wellbeing	2	21
	Stress and associated disorders including PTSD	23	160
	Mood disorders	13	41
	Anxiety and fear	3	6

	Suicide and self-harm	2	5
	Other mental health outcomes	3	6
Family therapy	Other mental health outcomes	2	3
Wellbeing interventions	Mood disorders	2	3
Meditation and mindfulness	Physical health and wellbeing	5	24
	Sleep	2	17
	Stress and associated disorders including PTSD	14	90
	Mood disorders	10	27
	Anxiety and fear	4	25
	Other mental health outcomes	6	24
Other therapeutic wellbeing interventions	Stress and associated disorders including PTSD	2	7
Medication	Stress and associated disorders including PTSD	3	6
	Mood disorders	2	10
	Substance use or addictions	2	5
	Other mental health outcomes	2	7
Advice and support	Mood disorders	4	4
	Substance use or addictions	3	5

Table A 5.30 Meta-analyses failing Egger's test

Intervention	Outcome	Number of studies	P-Value
Wellbeing interventions	Stress and associated disorders including PTSD	6	0.000
	Other mental health outcomes	6	0.014

6 Tables of characteristics

The following section presents Tables of Characteristics with a list of the studies included in the Prevalence, Experience and Effectiveness domains of the review

6.1Prevalence: Table of Characteristics

Table A 6.1 presents a Table of Characteristics for studies included in the Prevalence domain. It provides details on author(s), date of publication, populations of interest, study design, sample and estimates of prevalence of mental health conditions and behaviours.

Table A 6.1 Table of Characteristics for studies reporting on the prevalence of mental health issues

Author(s) and date	Population	Study design	Mental health conditions and behaviours - estimates of prevalence
Aguirre et al., 2014a and 2014b	Serving personnel	**Method:** Pilot study of an Enhanced Mental Health Assessment questionnaire. The questions were taken from validated screening tools including PC-PTSD screen, AUDIT-C, GAD-7 and PHQ-9. **Sample:** personnel undergoing routine and discharge medicals at four defence medical centres (n=325).	8% (mental health condition) 65% (higher risk drinking)
Ashwick and Murphy, 2017	Ex-Service personnel	**Method:** Postal questionnaire collecting data on demographics, physical and mental health in help-seeking veterans. A comparative analysis of participants by country (England, Scotland, Wales and Northern Ireland) was conducted to examine relationship between UK nations and mental health outcomes. Mental health measures used included the PCL-5, GHQ-12, DAR-5 and AUDIT. **Sample:** random sample of veterans seeking treatment from the charity Combat Stress (n=403)	Study found no differences in PTSD, CMD or anger across UK veterans. 82.3%-86.2% (PTSD) 70.7%-80.5% (Depression/Anxiety) 70.76%-82.8% (Anger) 15.7%-37.9% (Alcohol misuse). Higher rates of alcohol misuse in Scottish and Welsh veterans.
Bennett, 2017	Families	**Method:** Cross-sectional study involving an online survey administered to a sample of UK military partners at different	47.6% (trauma symptoms) 9.4% (mental health condition)

		stages of deployment. Survey based on validated questionnaires (DASS-42, PSS-10 and PCL-5). **Sample:** partners of military personnel (n=380).	
Bergman et al., 2017	Ex-Service personnel	**Method:** Data from a retrospective cohort study (Scottish Veterans Health Study) **Sample:** Scottish veterans (n= 56,205) and matched non-veterans (n=172,741) to compare risk of suicide and fatal self-harm.	0.48% n=267 (veteran suicides) 0.53% n=918 (non-veteran suicide).
Cawkill et al., 2015	Serving personnel	**Method**: Data collected as part of the KCHMR cohort study from respondents in a medical role who had completed the Phase 2 questionnaire. **Sample:** analysis of data of medical personnel (doctors, nurses and medical support personnel) who had been deployed to Iraq or Afghanistan (n=321). Sample consisted of 129 forward located medics (FMs) and 192 rear located medics (RLMs).	25% (Alcohol abuse - AUDIT>16, FMs) 16% (Alcohol abuse - AUDIT>16, RLMs) 9% (PTSD, FM) 3% (PTSD, RLM)
Dighton et al., 2018	Ex-Service personnel	**Method**: Comparison of rates of gambling, self-harm and suicide in veterans and non-veterans. **Sample:** random sample of veterans (n = 257) and sex and age-matched controls (n= 514) drawn from the 2007 Adult Psychiatric Morbidity Survey.	1.4% (veterans problem gambling), 0.2% (non-veterans problem gambling) 15.63% (male veterans self-harm), 33.4% (female veterans self-harm) 4.85% (male veterans attempted suicide), 15.71% (female veterans attempted suicide)
Goodwin et al., 2015	Serving personnel	**Method**: Used general population data from Health Survey for England (2003 and 2008) and serving military data from KCMHR (Phase 1 and 2) to calculate and compare rates of probable CMD, assessed by GHQ-12. **Sample:** military sample restricted to males, in-service, aged 18-44.	CMD (18.7%)
Goodwin et al., 2017	Serving personnel	**Method:** Longitudinal study on alcohol use in a sample of serving personnel. Baseline survey and follow-up questionnaire collecting data on weekly alcohol consumption across an eight-year period. **Sample:** Study based on sample at second follow-up (n =667)	Baseline weekly alcohol consumption (units): 9 (median), 4-20 (interquartile range) 55% (Mid-average drinkers); 4% (Abstainers); 19% (low level drinkers); 3% (decreasing drinkers); 19% (heavy drinkers)

Harden and Murphy, 2018	Ex-Service personnel	**Method**: Individuals completed questionnaires and data were then linked to risk assessments extracted from clinical records to explore the risk factors associated with suicidal ideation **Sample:** Participants randomly selected from veterans seeking help from Combat Stress 2015-16 (n=144)	No overall prevalence estimate
Head et al., 2016	Serving personnel, ex-Service personnel	**Method**: Questionnaire data from Phase 2 of KCMHR cohort study (n= 9984) used to assess participants for probable PTSD and alcohol misuse **Sample:** n=4725, regular military personnel (n= 4,246) and reservists (n=479)	4% (PTSD) 13% (Alcohol misuse) 1.8% (PTSD + Alcohol misuse)
Hines et al. 2014a	Serving personnel, Reservists	**Method:** Postal survey data from Phase 2 of KCMHR cohort study (2007-2009) Participants asked about general medical problems, stress or emotional problems, and alcohol problems resulting from their last deployment. **Sample:** n=4,725, regular military personnel (n= 4,246) and reservists (n=479)	6% (alcohol problem) 19% (emotional problem)
Hines et al., 2014b	Serving personnel	Systematic review of studies to determine prevalence of PTSD in military personnel deployed to Iraq or Afghanistan (n=49). Meta-analysis was used to assess PTSD prevalence across subgroups	12.9% (PTSD - Iraq deployed personnel) 7.1% (PTSD - Afghanistan deployed personnel)
Kwan et al.,2017	Serving personnel	**Method:** Questionnaire data from Phase 2 of KCMHR cohort study (2007-2009). Analysis of questions on premilitary antisocial behaviour (ASB), service history, experiences prior to and during deployment and measures of post-deployment physical and mental health. **Sample**: n=6,711 (5,741 regulars and 970 Reservists) deployed UK military personnel	3.6% (Family directed violence - post deployment) 7.6% (Stranger directed violence - post deployment) 2.3% (Both family and stranger directed violence - post deployment)
Kwan, 2016	Reservists	**Method:** questionnaire data from Phase 1 and 2 of the KCMHR cohort study. Participants asked about socio-demographic and military characteristics, pre-enlistment antisocial behaviour, deployment experiences, post-deployment mental health, and self-reported interpersonal violent behaviour. **Sample:** n=1,710 (including Reservists	3.5% n=60 (Violent behaviours)

MacManus et al., 2014	Serving personnel, ex-Service personnel, Reservists	Literature review examining the mental health outcomes and apparent resilience of the UK military (in comparison with the US).	1.3-4.8% (PTSD - returning from deployment) 3.8% (PTSD - during deployment) 16-20% (Alcohol misuse) 10.5% (Self harm - ex-Service personnel) 4.2% (Self-harm - serving personnel)
MacManus et al., 2015	Serving personnel, ex-Service personnel	A systematic review of articles describing the prevalence of, or statistical relationships between, risk factors and post-deployment violence among military populations (n=17). Meta-analysis (n=3) explored the association between combat exposure and aggressive or violent behaviour	8% (Violence convictions - post deployment) 29% (Aggressive/violent behaviour - post deployment)
MacManus, 2013	Serving personnel, Reservists	**Method**: Data linkage study to assess risk factors of violent offending using data from Phase 1 and 2 of KCMHR study and Ministry of Justice national criminal records. Questionnaire gathered information on gather information on socio-demographics, experiences/behaviour since joining the military and health/behaviour after deployment. **Sample:** 13,856 randomly selected, serving and ex-Service UK military personnel	15.7% (committed one or more violent offences in their lifetime) 64% of all offenders in sample were violent. Alcohol/drug related offences: 6.4%
Ministry of Defence (MOD), 2016	Serving personnel	**Method:** Secondary analysis of data on suicides and open verdict deaths among serving UK regular Armed Forces 1996-2015 aged 16-59 to calculate suicide rate. Data on deaths **Sample:** Analysis of male population aged 16-59 years only	0.009% (Suicide - Armed forces, male) 4% (PTSD - all Armed forces) 7% (PTSD - combat troops)
Ministry of Defence (MOD), 2018a	Serving personnel, Reservists	**Method:** Secondary analysis of data provided by DCMH and in-patient providers. Summary of all attendances for a new episode of care of Armed Forces personnel at MOD Specialist Mental Health services only.	0.2% (PTSD) 0.1% (Psychoactive substance abuse) 1% (Mood disorders) 1.9% (Neurotic disorders) 3.1% in 2017/18 (UK armed forces personnel receiving treatment at an MoD specialist mental health service)
Ministry of Defence (MOD), 2018b	Serving personnel	**Method:** secondary analysis of data from the Defence Medical Information Capability Programme to calculate rate of UK Armed Forces personnel who were assessed in primary health care for a disorder	3.5% (% of armed forces personnel assessed with a mental disorder) 4% (PTSD - all armed forces personnel) 7% (PTSD - combat troops)

Murphy and Turgoose, 2018a	Ex-Service personnel	Rapid evidence review presenting research relating to prevalence rates, challenges in defining TBI, and co-morbidities with PTSD and other mental health difficulties	No overall prevalence estimate
Murphy and Turgoose, 2018b	Serving personnel, Ex-Service personnel	**Method:** Survey on alcohol consumption in male veterans using the AUDIT. Comparisons with data on the UK Armed Forces and the general public **Sample**: n=403 veterans who had sought support from Combat Stress. Analysis of alcohol use was restricted to male veterans only	42% (hazardous alcohol use); 22% (harmful alcohol use); 37% (alcohol related harm) Active military personnel 24% (alcohol related harm)
Murphy et al., 2015	Ex-Service personnel	**Method:** Data on a range of mental health measures collected from veterans at initial assessment, and at follow up to explore the prevalence rate of brain injury and symptoms of post-concussion syndrome **Sample:** n=123. Drawn from new admissions to Combat Stress from January - July 2014.	80% (Depression) 73% (PTSD) 69% (Generalised anxiety) 53% (Anger problems) 49% (Alcohol problems) 63% (Mild Traumatic brain injury)
Murphy et al., 2016	Families	**Method**: Postal survey of partners of UK veterans diagnosed with PTSD. **Sample:** female partners of male veterans (n=100)	45% (alcohol difficulties) 39% (Symptoms of depression) 37% (Generalised anxiety) 17% (Probable PTSD)
Rona et al., 2016	Serving personnel, ex-Service personnel	Meta-analysis of UK studies selected from the database of King's Centre for Military Health Research (n=9) which included personnel assessed for mental health outcomes after their most recent deployment (PCL-C, GHQ-12, AUDIT)	2-2.9% PTSD 6 months post-deployment 2.5-4.3% PTSD18 months after deployment
Royal British Legion, 2014	Ex-Service personnel, Families	**Method:** questions placed on a nationally representative Omnibus Survey of UK adults. The report summarises the size, profile and needs of the ex-Service community in 2014. **Sample: n=** 2,121. The questionnaire included 19 screening questions asked to identify members of the Armed Forces community.	10% (Depression, ex-Service personnel) 1% (Alcohol-related illness) 12% (Psychological difficulties)

Short et al., 2018	Ex-Service personnel, Serving personnel	**Method**: Analysis of data from 29 Liaison and Diversion (L&D) services across England from the period 2015-2016. Veterans and non-veterans were compared in regard to socio-demographic factors, offending behaviour and mental health characteristics **Sample:** 2.4% of the sample (n=1,215) reported previous or current service in the UK Armed Forces.	Ex-Service personnel: 37% (Anxiety); 32% (Depression); 5% (Schizophrenia); 38% (Substance misuse); 2% (Dementia); 7% (Adjustment disorder); 7% (Personality disorder); 0.5% (ADHD); 4% (Bipolar disorder)
Stevelink et al., 2015	Ex-Service personnel	**Method:** Telephone interviews with male ex-Service personnel with a visual impairment from the charity Blind Veterans UK to assess their mental wellbeing **Sample:** (n=74)	28.4% (Probable PTSD, anxiety or depression) 45% (hazardous drinking, combat related visual impairment) 20.4% (hazardous drinking, non-combat related visual impairment)
Stevelink et al., 2018	Serving personnel, ex-Service personnel, Reservists	**Method:** Data from three phases of KCMHR study. Self-completion questionnaire **Sample:** Third phase (2014–2016; n = 8093). The sample was based on participants from previous phases (2004–2006 and 2007–2009) and a new randomly selected sample of those who had joined the UK armed forces since 2009.	Service personnel: 10.2% (Alcohol misuse); 4.8% (Probable PTSD); 21.9% (CMD symptoms) Ex-Service personnel: 10.3% (Alcohol misuse); 7.4% (Probable PTSD); 21.5% (CMD symptoms) Mobilised: 21.3% (CMD, not deployed); 22.% (CMD, deployed); 5.2% (PTSD, not deployed); 6.9% (PTSD, deployed); 8.3% (Alcohol misuse, not deployed); 11.4% (alcohol misuse, deployed) Reservists: 18.8 % (CMD, not deployed); 27.5% (CMD, deployed); 3.2% (PTSD, not deployed); 6.9% (PTSD, deployed); 5.5% (Alcohol misuse, not deployed); 9.9% (Alcohol misuse, deployed)
Thandi et al., 2015a	Reservists	**Method:** Phase 2 of KCHMR cohort study. Questionnaire on hazardous drinking, risky driving, physical violence, smoking and attendance at accident and emergency (A&E) departments as a result of risk-taking behaviours **Sample:** (n=1710)	46% (Hazardous drinking) 3% (Physical violence)

Thandi et al., 2015b	Serving personnel	**Method**: Analysis of data from Phase 1 and 2 of the KCMHR cohort study to assess changes in AUDIT scores over time **Sample**: Random representative sample of regular UK military personnel who were surveyed in 2004–2006 (phase1) and again in 2007–2009 (phase 2) (n=5239)	5.3% (Alcohol dependence, Phase 1) - 4.9% (Phase 2)
Turgoose and Murphy, 2018a	Serving personnel, ex-Service personnel	Brief review highlighting some of the main issues regarding mild Traumatic Brain Injury (m TBI), with a focus on military personnel and veterans	9.5% (mTBI, combat roles)
Turgoose and Murphy, 2018b	Ex-Service personnel	**Method:** Questionnaire of random sample of UK veterans seeking help for mental health difficulties from the charity Combat Stress, collecting data on anger, aggression and mental health and sociodemographic variables **Sample: (**n =403)	74% (anger difficulties) 28% (aggressive behaviours)
Whybrow et al., 2015	Serving personnel	**Method:** Survey to assess prevalence of mental health disorders and associations with military and operational characteristics **Sample:** Deployed Royal Navy personnel (n=1393)	41.2% (CMD) 7.8% (Probable PTSD) 17.4% (Alcohol misuse)
Williamson et al., 2018	Ex-Service personnel, Reservists, Serving personnel	Meta-analysis of the findings of several multiple observational studies to estimate mental health disorder prevalence among Armed Forces personnel (serving, regulars and reserves, veterans). Used data obtained via Freedom of Information requests and raw data	Regular serving: 10% (mental health problem, deployed); 17% (mental health problem, non-deployed); 4% (PTSD, deployed); 3% (PTSD, non-deployed); 14% (alcohol misuse, deployed); 11% (alcohol misuse; non-deployed) Veterans: 10% (mental health problem, regulars), 14% (mental health problem, non-deployed); 9% (PTSD, deployed); 5% (PTSD, non-deployed), 13% (alcohol misuse, deployed); 10% (alcohol misuse, non-deployed) Reservists: 11% (mental health problem, deployed mobilised); 17% (mental health problem, non-deployed mobilised); 6% (mental health problem, deployed veteran); 19% (mental health problem, non-deployed veteran) 5% (PTSD, non-deployed veterans); 2% (PTSD, non-deployed veterans);

			4% (PTSD, deployed mobilised); 2% (PTSD, non-deployed mobilised) 9% (alcohol misuse, deployed mobilised reservists); 8% (alcohol misuse, non-deployed mobilised); 12% (deployed and non-deployed veterans)
Woodhead, 2013	Serving personnel, ex-Service personnel	**Method:** Quantitative element of a mixed method study on mental health of female service personnel. **Sample:** Female participants (n=1185) were surveyed as part of a cohort study of UK military personnel	<5% (PTSD, female service personnel) Nearly 25% (CMD, female service personnel) Nearly 40% (Hazardous alcohol use, female service personnel)

6.2 Experience: Table of characteristics

Table A6.2 presents a list of the studies included in the Experience domain, including details on author(s) and date of publication, populations of interest, study design and sample size.

Table A 6.2 Table of Characteristics for studies reporting on the experience of mental health issues

Author(s) and date	Population	Study design
Brian Parry Associates, 2015	Families	An initial Stakeholder Roundtable meeting with a selected group of 11 organisations; a series of seven deliberative style Stakeholder Events; and an Online Response Form designed to enable written submissions on the topic areas of the stakeholder events. Input received from 159 individuals from 67 different organisations.
Bull et al., 2015	Serving personnel	Interviews with medical and welfare staff currently serving in the UK Armed Forces (n=21).
Caddick et al., 2015	Ex-Service personnel	Life history interviews and participant observation with male combat veterans (n=15).
De Rond and Lok. 2016	Ex-Service personnel	Ethnographic observation of a medical military team deployed in Afghanistan. A personal journal of “headnotes” with ethnographer’s own experiences, anxieties, and reflections was also kept.
Dighton et al., 2018	Ex-Service personnel	Data analysis with samples drawn from the 2007 Adult Psychiatric Morbidity Survey. Sample includes veterans (n = 257) and sex and age-matched controls (n = 514).
Doncaster et al., 2015	Families	Fifteen semi-structured interviews with partners of veterans with initial phase including a systematic review.

Farrand et al., 2018	Veterans and family members	Qualitative interviews with veterans and family members (n=14).
Fertout et al., 2015	Mobilised and non-mobilised Reservists, Serving personnel	Principal component analysis from three surveys data with 3405 personnel (15% of deployed UK force surveyed on each of the three occasions).
Hallett, 2012.	Ex-Service personnel	Semi-structured interviews with 6 participants. Interpretive Phenomenological Analysis (IPA) used to explore how a phenomenon is experienced from the perspective of those involved.
Hatch et al., 2013	Serving personnel	Data collected by the self-completion of questionnaires in a UK-based cohort study designed to monitor the physical and mental health of a UK Armed Forces personnel. Sample including 9395 participants from a randomly selected group of deployed and non-deployed personnel from the 2003 Iraq war; 1789 randomly selected personnel who had been deployed to Afghanistan between April 2006 and April 2007; and a replenishment sample of 6628 randomly selected individuals who had joined the UK Armed Forces since the phase 1 cohort was recruited in 2003.
Hatton, 2016	Ex-Service personnel	Semi-structured interviews with 9 male veterans. Interpretive Phenomenological Analysis used to analyse interview transcripts to seek to privilege the accounts of the participants themselves and how they make sense of their experiences.
Hines et al., 2014a	Serving personnel, Mobilised and non-mobilised Reservists	Questionnaire completed by personnel returning from deployment in Iraq and Afghanistan (n=4725). Longitudinal sample drawn from KCHMR Phase 2 (cross-sectional analysis).
Hunt et al., 2016	Serving personnel	Secondary analysis of two combined datasets of an RCT (Battlemind), participants completed a baseline survey (n=2443) and a survey 4-6 months later (n=1636).
Jones and Coetzee, 2018	Serving personnel	Data collected from questionnaire which asked DCMH attendees to rank order 10 potential reasons for seeking help from mental health services. Study sample comprised 2 sub-groups: 1) DCMH attendees: UK armed forces personnel referred to DCMH (n=549) 2) Cohort sample: Participants in the (n=3682).
Jones et al.,2013	Serving personnel	Participants completed a questionnaire (n=484). Cross-sectional study detecting a change of +/-5% in reporting one or more stigma scale item among 212 personnel with 95% confidence.
Jones et al.,2015	Serving personnel	Secondary analysis of an RCT of a post-deployment intervention known as Battlemind. Participants completed a survey (n=1636) immediately after returning from deployment and 4-6 months later.

Jones et al.,2018	Serving personnel	Secondary analysis of data derived from a cluster randomised controlled trial of an intervention known as UK Battlemind. Participants were surveyed twice - at baseline, following deployment (n= 2510) and 4–6 months later (n= 1636).
Kantor et al., 2017	Serving personnel	A synthesis study about trauma survivors' perceived barriers and facilitators regarding MHS utilization.
Keeling et al., 2017	Serving personnel	21 semi-structured telephone interviews.
Kiernan et al., 2017	Ex-Service personnel	The study contains three phases of data collection: 1) Interviews with planners/providers/commissioners (n=6) 2) Interviews with veteran service users, almost all had severe alcohol misuse problems (n=22) 3) Focus groups with wider ex-armed forces community (no substance misuse problems) (n=9).
Kiernan et al., 2018	Ex-Service personnel	19 in-depth semi-structured interviews with UK veterans who have a history of alcohol misuse with the use of purposive sampling.
Lovatt, R. 2017.	Ex-Service personnel	7 semi-structured interviews with former Army personnel who had been diagnosed with PTSD following assessment by Combat Stress with the use of purposive sampling.
Mellotte et al., 2017	Ex-Service personnel	Semi-structured interviews used to explore veterans' experiences of help-seeking (n=17). Quantitative data regarding participant's demographic info and current mental health used to select the sample.
Murphy et al., 2014	Serving personnel	Interviews carried out with UK service personnel (n=8) accessing mental health care through two DCMHs. Interpretative phenomenological analysis used to explore the lived experiences of participants during their pathways to accessing mental health services.
Murphy et al., 2016	Families	Data were collected via sending questionnaires to potential participants in two waves (n=141). Assessed associations between endorsing a particular barrier to care and severity of mental health score using Univariate logistic regression models (odds ratios and 95% confidence intervals calculated).
Murphy et al., 2017a	Ex-Service personnel	Quantitative data linkage from UK multiple deprivation indices and neighbourhoods lived in by population of veterans who have contacted Combat Stress (n=3,120).
Murphy et al., 2017b	Families	Semi-structured interviews (n=8). Interpretative Phenomenological Analysis (IPA) used to understand the lived experiences of participants from their perspective.
NHS, 2016	Ex-Service personnel	Feedback from a formal engagement on NHS veterans' mental health services to help inform future service provision and improve care. The engagement was a questionnaire (other responses via letter, phone, emails).
		Most responses were from veterans (n=715).

Palmer et al., 2016	Ex-Service personnel	Qualitative interviews explored the lived experience of PTG for veterans who have recently received treatment for PTSD and reported growth according to the PTGI. Interpretative Phenomenological Analysis (IPA) was chosen as the preferred methodology to explore the meaning of phenomena through a researcher's interpretation of individuals' accounts.
Rafferty et al., 2017	Ex-Service personnel	62 in-depth telephone interviews.
Ramchand et al., 2015	Serving personnel	Synthesis study.
Rowe et al., 2013	Serving personnel	Data drawn from the first stage of a cohort study comparing the physical and mental health outcomes between those who took part in the 2003 invasion of Iraq (January 18 to June 28, 2003) and those who were in the military at that time but not deployed. 17,812 military personnel were sampled, with 10,272 questionnaires received (response rate = 58.4%).
Rowe et al., 2014	Families	Service personnel (regular and reserve) with one or more children (<18 years) were included. Data were taken from a large UK military cohort study completed between 2007 and 2009 (n= 3198). Unadjusted multinomial regression analyses were conducted to calculate unadjusted multinomial odds ratios (MORs) for the associations between perceived impact of military career on children and all explanatory variables.
Sharp, 2015	Ex-Service personnel, current service personnel	Mixed methods design based on 2 qualitative studies and 1 quantitative study: Study 1: sample (help seekers and non-help seekers) taken from KCMHR Phase 2 (N=16). Study 2: sample (N=10) was recruited from help seeking beneficiaries (all ex-Service and all diagnosed mental health problems - mainly PTSD) of Combat Stress. Study 3: used data collected by a KCMHR clinical telephone interview study, which recruited participants from phase 3 of the KCMHR cohort military study (N=453).
Stevelink and Fear, 2016	Ex-Service personnel	Study contains two phases: Phase 1: telephone questionnaire examining participants' mental health; Phase 2: face-to-face interviews using an 11-item semi-structured interview schedule.
Turgoose et al., 2018	Ex-Service personnel	Observational design to explore the impact of Tele-therapy on the mental health of UK help-seeking veterans diagnosed with PTSD (n=21). Participants completed the outcome measures at three time points: before therapy, immediately after therapy, and at a three-month follow-up point (p.20). Sub-sample of participants completing semi-structured interviews post-therapy. Quantitative analyses were used to assess the effectiveness of tele-therapy using self-report measures of PTSD, anxiety, depression, anger and alcohol use.
Wainwright et al., 2016	Ex-Service personnel	In-depth interviews (n=20) with male ex-armed forces personnel in prison with the use of purposive sampling ensure individuals with a range of ages, type of service, rank and offence types were interviewed.

Woodhead, 2013.	Serving personnel	A quantitative analysis using data from the KCMHR cohort study to identify statistically significant relationships between specific risk factors and health outcomes. A qualitative approach (semi-structured telephone interviews) was used to explore work, family and interpersonal stressors.

6.3 Effectiveness: Table of characteristics

Table A 6.3 presents a list of the studies included in the Effectiveness domain, including details on author(s) and date of publication, populations of interest, focus country (nationality of the population being studied), intervention evaluated and study design.

Table A 6.3 Table of Characteristics for studies reporting on the effectiveness of interventions to address of mental health issues

Short Title	Population	Setting	Study Design	Intervention evaluated	Outcomes
Acierno et al., 2017	Veterans	USA	Randomised Controlled Trial (RCT)	Prolonged exposure (PE) via home-based telehealth compared to standard in-person PE	PTSD
Badura-Brack et al., 2015	Veterans	USA, Israel	RCT	Attention bias modification involves computerized cognitive training strategies designed to alter biases in attention	PTSD
Batki et al., 2014	Veterans	USA	RCT	A randomized, double-blind, placebo-controlled, flexible-dose (25 to 300 mg/d) pilot trial of topiramate augmentation treatment.	PTSD, Disorders due to substance use or addictive behaviours
Bormann et al., 2013	Veterans	USA	RCT	Mantram repetition program (MRP) is a portable meditation-based intervention that teaches three tools for training attention and regulating emotion.	PTSD, Anxiety
Bourque et al., 2015	Veterans	Canada	RCT	Assertive Community Treatment (ACT) is a community-based, multidisciplinary mental health intervention team that is available 24/7, with a typical service provider-to-user ratio of 1:10.	Socio-economic, Quality of life
Bremner et al., 2017	Veterans	USA	RCT	MBSR intervention provides systematic and intensive training in mindfulness through formal meditation and mindful hatha yoga exercises.	PTSD

Brief et al., 2017	Veterans	USA	RCT	VetChange is designed to motivate veterans to make changes in drinking and to develop skills necessary to reduce drinking to a safer level (either moderation or abstinence).	Disorders due to substance use or addictive behaviours
Brown, 2013	Veterans	USA	RCT	An initial five-week psychoeducation group known as “PTSD 101.”to help bridge the gap between initial diagnosis and treatment	Depression, PTSD, Anxiety
Buffington et al., 2016	National Serviceman, Other	USA	RCT	The COPE 10-session intervention is based on CBT, which contends that how an individual thinks is directly related to how he or she feels and behaves.	Anxiety, PTSD
Carter et al., 2013	Veterans	USA	RCT	A modified SKY programme on group yoga instructions for veterans	PTSD, Quality of life, Anxiety, Disorders due to substance use or addictive behaviours, Depression
Castillo et al., 2016	Veterans	USA	RCT	A 16-week, 90-min, three-module treatment protocol consisted of five imaginal exposure, five cognitive, and four behavioural skills sessions.	Physical health and wellbeing (perceived physical health, QoL) Other mental health (perceived mental health), PTSD
Chinman et al., 2013	Veterans	USA	RCT	Use of Peer specialists in traditional VHA case management teams	Quality of life, Other mental health
Church et al., 2013	Veterans	USA	RCT	Emotional freedom techniques coaching therapy	PTSD, Other mental health, Anxiety
Davis et al., 2018	Veterans	USA	RCT	Individual placement and supported employment	Socio-economic
Dretsch et al., 2014	Serving: regulars	USA	RCT	Active treatment was 2.5 g/d of EPA+DHA ethyl esters provided in 3 capsules of Lovaza (Lovaza consists of 47% EPA, 38% DHA, and 4% docosapentaenoic acid with no other FA greater than 2%).	Sleep, Anxiety
Egede et al., 2015	Veterans	USA	RCT	Telemedicine-delivered psychotherapy (group therapy received via videophone so participants received treatment in their home). 8 one hour sessions delivered once a week.	Depression

Eisen et al., 2012	Veterans	USA	RCT	One weekly peer led recovery group (Vet to-Vet) and one weekly clinician led recovery group	Self harm, Disorders due to substance use or addictive behaviours, Other mental health, Depression
Engel et al., 2014	Serving: regulars	USA	RCT	PTSD care (UPC) plus eight 60-minute sessions of acupuncture conducted twice weekly or to UPC alone.	PTSD, Depression
Engel et al, 2016	Serving: regulars	USA	RCT	Central Assisted Collaborative Telecare (stepped psychosocial and pharmacologic treatment with nurse telecare management of caseloads, symptoms, and treatment)	PTSD, Depression
Foa et al., 2018	Serving: regulars	USA	RCT	-Massed Prolonged Exposure Therapy (daily over 2week period) -Spaced Prolonged Exposure Therapy (delivered over 8 weeks). -Present centred therapy	PTSD
Fortney et al., 2015	Veterans	USA	RCT	The Telemedicine Outreach for PTSD (TOP): provision of care via telephone/video	PTSD, Quality of life, Depression
Gelkopf et al., 2013	Veterans	Israel	RCT	Nature Adventure Rehabilitation (NAR)	PTSD, Depression, Quality of Life, Other mental health (Hope, interpersonal relationships, control of PTSD)
Geronilla et al., 2014	Veterans	USA	RCT	Emotional freedom techniques	PTSD, Physical health and wellbeing, Other mental health, Sleep, Anxiety
Gmel et al., 2013	National Servicemen	Switzerland	RCT	Brief integrative multi-substance intervention. It involves exploring the use of tobacco, cannabis, alcohol and other substances by introducing and discussing behaviour change perspectives in a non-judgmental, empathic and collaborative manner	Disorders due to substance use or addictive behaviours, Smoking, Disorders due to substance use or addictive behaviours
Gray et al., 2017	Veterans	USA	RCT	Cognitive intervention - reconsolidation of traumatic memories (RTM). Three 120-minute sessions were delivered	PTSD

Greenberg et al., 2010	Serving: regulars	UK	RCT	Traumatic Risk Management - peer delivered psychological support process	PTSD, Other mental health
Harris et al., 2015	Veterans	USA	Matching, including Propensity Score Matching (PSM)	Extended release naltrexone (XRN)	Disorders due to substance use or addictive behaviours,
Hobfoll et al., 2016	Veterans	USA	RCT	Vets Prevail - seven online CBT lessons, a community message board, and peer chat support	Depression, PTSD
Johnson et al., 2018	Veterans	USA	RCT	Therapeutic horseback riding (THR)	Other mental health (loneliness, self-efficacy, emotional regulation), PTSD
Jones et al., 2013	Serving: regulars	UK	Matching, including Propensity Score Matching (PSM)	Third Location Decompression (TLD)	PTSD, Disorders due to substance use or addictive behaviours, Physical health, Other mental health outcome
Kahn et al., 2016	Veterans, Families of the above groups of people	USA	RCT	Web-based and mobile app video and audio instruction in a set of mindfulness-related stress reduction and contemplative practices, as well as partner massage for reciprocal use	Depression, PTSD, Stress
Kearney et al., 2013	Veterans	USA	RCT	Mindfulness-Based Stress Reduction (MBSR): group-based, includes meditation, yoga	Depression, PTSD, Physical health and wellbeing, Other mental health, Quality of life
Kearney et al., 2016	Veterans	USA	RCT	Mindfulness group meditation	Depression, Other mental health, PTSD, Physical health and wellbeing
Kilbourne et al., 2014	Serving: regulars	USA	RCT	Re-Engage program to contact veterans with mental illness, assess patient clinical status, and facilitate appointments to VA health services.	Attendance at appointments

Kip et al., 2013	Serving: regulars, Veterans	USA	RCT	Accelerated resolution therapy	PTSD, Disorders due to substance use or addictive behaviours, Anxiety, Other mental health
Krupnick et al., 2017	Veterans	USA	RCT	Online writing intervention as an adjunct to treatment as usual	PTSD, Physical health, Disorders due to substance use or addictive behaviours, Depression
Kuckertz et al., 2014	Serving: regulars	USA	RCT	Attention bias modification+normal treatment	PTSD, Depression
LaCroix et al., 2018	Other	USA	RCT	Post-Admission Cognitive Therapy (PACT)	Depression, Other mental health, PTSD, Suicide
Luxton et al., 2016	Veterans	USA	RCT	behavioural activation treatment for depression (BATD), either home-based or in office.	PTSD, Depression, Other mental health, Anxiety
Mack, 2013	Veterans	USA	RCT	Cognitive behavioural therapy for insomnia (CBT-I) and imagery rehearsal therapy (IRT)	Sleep, PTSD
Mackintosh et al., 2017	Veterans	USA	RCT	Anger Management Treatment (AMT). Compared anger management treatment (AMT) with AMT augmented by a mobile application (app) system, Remote Exercises for Learning Anger and Excitation Management (RELAX)	Other mental health, Depression, PTSD, Delivery
Maguen et al., 2017	Veterans	USA	RCT	CBT called Impact of Killing in War, A 6- to 8-session, weekly, individual, CBT, lasting 60–90 minutes and focused on key themes, including physiology of killing responses, moral injury, self-forgiveness, spirituality, making amends, and improved functioning.	Anxiety, Depression, PTSD
Maieritsch et al., 2016	Veterans	USA	RCT	Video-conference technologies (cognitive processing therapy)	Depression, PTSD
Margolies et al, 2013	Veterans	USA	RCT	CBT	Depression, PTSD, sleep
Martens et al., 2015	Veterans	USA	RCT	Personalized Drinking Feedback Intervention	Disorders due to substance use or addictive behaviours
McDevitt-Murphy et al., 2014	Veterans	USA	RCT	Brief alcohol interventions including information on hazardous drinking, PTSD symptoms, depression, and coping	Disorders due to substance use or addictive behaviours

McLay et al., 2017	Serving: Regulars	USA	RCT	Virtual reality exposure therapy	PTSD
Mithoefer et al., 2018	Veterans, Other	USA	RCT	MDMA- assisted psychotherapy	PTSD, Depression, Sleep, Other mental health
Moriarty et al., 2016	Veterans, Families of the above groups	USA	RCT	Veterans In-Home program (6 home visits and 2 telephone calls delivered by occupational therapists over a 3- to 4-month period. Family members were invited to participate in the 6 home sessions)	Depression, Other mental health outcomes
Morland et al., 2014	Veterans	USA	RCT	Cognitive processing therapy-cognitive only version (CPT-C) delivered via video teleconferencing to in-person in a rural, ethnically diverse sample of veterans with PTSD	Other mental health, Depression, PTSD, Delivery
Mulligan et al., 2011	Serving: regulars	UK	RCT	Post-deployment screening. Battlemind, Trauma Risk Management	PTSD, General health, Disorders due to substance use or addictive behaviours, Physical health and wellbeing
Nacasch et al., 2015	Veterans	USA	RCT	20 min Imaginal exposure & 60 min prolonged exposure	Depression, Other mental health, PTSD
Oman and Bormann, 2015	Veterans	USA	RCT	Case management + mantram repetition program (MRP), a portable meditation-based intervention	Depression, PTSD
Oslin et al., 2014	Veterans	USA	RCT	Primary care-based Alcohol Care Management (ACM) program for Disorders due to substance use or addictive behaviours disorder and treatment engagement in veterans	Disorders due to substance use or addictive behaviours, Attendance
Pedersen et al., 2017	Veterans	USA	RCT	Web-based personalized normative feedback alcohol intervention for young adult veterans, looking to change perceived norms and intended drinking behaviour	Disorders due to substance use or addictive behaviours, Depression
Polusny et al., 2015	Veterans	USA	RCT	Mindfulness based stress reduction therapy - group sessions	Depression, PTSD, Quality of life
Possemato et al., 2016	Veterans	USA	RCT	Mindfulness therapy	Depression, PTSD

Raskind et al., 2018	Veterans	USA	RCT	Receive prazosin for 26 weeks	PTSD, Depression, Sleep, Disorders due to substance use or addictive behaviours, Suicide, Other mental health
Rauch et al., 2015	Veterans	USA	RCT	Prolonged exposure	PTSD
Reger et al., 2016	Serving: regulars	USA	RCT	Prolonged exposure (PE) and Virtual reality exposure (VRE)	Depression, PTSD, Satisfaction with care
Resick et al., 2015	Serving: regulars	USA	RCT	Cognitive Processing Therapy, compared With Group Present-Centred Therapy	Depression, PTSD
Resick et al., 2017	Serving: Regulars	USA	RCT	Group vs Individual Cognitive Processing Therapy	Depression, PTSD, Suicide
Rona et al., 2017	Serving: regulars	UK	RCT	Tailored help seeking advice	Depression, PTSD, Disorders due to substance use or addictive behaviours, Anxiety
Rosen et al., 2013	Veterans	USA	RCT	A telephone care management protocol to usual aftercare	PTSD, Depression, Other mental health, Disorders due to substance use or addictive behaviours, Quality of life, Other mental health outcomes
Rosen et al., 2017	Veterans	USA	RCT	A telephone care management protocol to usual aftercare	PTSD, Depression, Other mental health, Disorders due to substance use or addictive behaviours, Quality of life, Delivery, Other mental health outcomes

Sautter et al., 2015	Veterans, Families of the above groups	USA	RCT	Structured approach therapy (SAT), a couples-based treatment to reveal and discuss trauma-related memories and emotions with their partners	PTSD, Anxiety, Depression, Other mental health
Sayer et al., 2015	Veterans	USA	RCT	Online expressive writing	PTSD, Other mental health, Physical health and wellbeing, Quality of life, Socioeconomic
Seppälä et al., 2014	Veterans	USA	RCT	Mediation	Depression, PTSD, Anxiety
Shea et al., 2013	Veterans	USA	RCT	Adapted cognitive behavioural intervention (CBI)	Anxiety, Other mental health, Delivery
Shipherd et al., 2016	Serving: Mobilised and non-mobilised Reservists	USA	RCT	Training as usual (TAU), psychoeducation only (psychoeducation on intrusive cognitions [PIT]), psychoeducation plus change-based skills (CONTROL), and psychoeducation plus acceptance-based skills (RESET).	PTSD, Depression, Anxiety
Surís et al., 2013	Veterans	USA	RCT	CBT	Depression, PTSD
Tuerk et al., 2018	Veterans	USA	RCT	Prolonged exposure therapy with dose of yohimbine prior exposure	Depression, PTSD
Tylee et al., 2017	Serving: regulars	UK	RCT	Cognitive intervention - reconsolidation of traumatic memories (RTM). Three 120-minute sessions were delivered.	PTSD
Wahbeh et al, 2016	Veterans	USA	RCT	Mindfulness group meditation	Depression, PTSD, Stress, Sleep, Other mental health
Wesemann, 2016	Serving: regulars, Serving: mobilised and non-mobilised reserves	Germany	RCT	Chaos Driven Situations Management Retrieval System (CHARLY), a computer-aided training platform with a biofeedback interface.	Other mental health

Valenstein et al., 2015	Veterans	USA	RCT	Telephone-based mutual peer-support intervention	Suicide, Other mental health, Physical health and wellbeing, Quality of life, Depression
Verduin et al., 2013	Veterans	USA	RCT	Computer simulation to practice relapse prevention skills	Disorders due to substance use or addictive behaviours, Other mental health,
Wolf, 2016	Veterans, Serving: regulars	USA	RCT	Prolonged exposure vs. present centred	Dissociative disorders, PTSD
Yehuda et al., 2014	Veterans	USA	RCT	Prolonged exposure therapy	Depression, PTSD
Yuen et al., 2015	Veterans	USA	RCT	Prolonged exposure (PE) therapy	Anxiety, Depression, PTSD
Ziemba, 2014	Serving: regulars, Serving: mobilised and non-mobilised reserves	USA	RCT	10-weekly sessions of CBT telemedicine	Anxiety, Depression, PTSD

7 References

Baird, S., Ferreira, F., Özler, B. and Woolcock, M. (2014). Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. Journal of Development Effectiveness, 6(1), pp.1-43.

Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). Introduction to Meta-Analysis. Chichester: John Wiley & Sons Ltd.

Ferguson, C. and Brannick, M. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. Psychological Methods, 17(1), pp.120-128.

Gleser, L. J., & Olkin, I. (2007). Stochastically Dependent Effect Sizes. Technical report No. 2007-2. Stanford, California: Stanford University.

Harris, R., Deeks, J., Altman, D., Bradburn, M., Harbord, R. and Sterne, J. (2008). Metan: Fixed- and Random-Effects Meta-Analysis. The Stata Journal: Promoting communications on statistics and Stata, 8(1), pp.3-28.

Hombrados, G. J. and Waddington, H. (2012). Internal validity in social experiments and quasi experiments: An assessment tool for reviewers. Unpublished working document. International Initiative for Impact Evaluation (3ie).

Keef, S. P. & Roberts, L. A. (2004). The meta-analysis of partial effect sizes. British Journal of Mathematical and Statistical Psychology, Volume 57, pp. 97-129.

Lipsey, M. & Wilson, D. (2001). Practical meta-analysis. Thousand Oaks, CA: Sage.

Ouzzani, M., Hammady, H., Fedorowicz, Z. and Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. Systematic Reviews, 5(1).

Oya, C., Johnston, D., Muchiri, E., Schaefer, F., Skalidou, D., Dickson, K. and Stansfield, C. (2015). Protocol for a Systematic Review: Effects of Certification Systems for Agricultural Commodity Production on Socio-economic Outcomes in Low and Middle-Income Countries. Campbell Systematic Reviews, 11(1), pp.1-71.

Phillips, D., Marcinkiewicz, A., Wishart, R., Forsyth, E., Nguyen, A., Lynch-Huggins, S., Gogescu, F., Gilbert, A., Dinos, S., Vojtkova, M., Eyers, J., Denney, D. (2020a) The mental health needs of serving and ex-Service personnel: A systematic review.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36 (3), 1–48.

Wilson, D. B. N.D. Practical Meta-Analysis Effect Size Calculator