Psychological therapies for the management of chronic pain (excluding headache) in adults – a Cochrane review update

Amanda C de Williams1, Emma Fisher2,3, Les Hearn4, Christopher Eccleston2,3,5
1Research Department of Clinical, Educational & Health Psychology, University College London, London, UK
2Cochrane Pain, Palliative and Supportive Care Group, Churchill Hospital, Oxford, UK
3Centre for Pain Research, University of Bath, Bath, UK
4Pain Research and Nuffield Department of Clinical Neurosciences, University of Oxford, UK
5Department of Clinical and Health Psychology, Ghent University, Ghent, Belgium

*Corresponding author: Amanda Williams, Research Department of Clinical, Educational & Health Psychology, University College London, London, UK [amanda.williams@ucl.ac.uk].

Email addresses: emma.fisher@ndcn.ox.ac.uk; les_hearn@yahoo.co.uk; c.eccleston@bath.ac.uk

The authorship of this update is different from the previous versions (Eccleston, Williams, & Morley, 2009; Williams, Eccleston, & Morley, 2012). We recognise the contributions of Stephen Morley who died in 2017 (Eccleston & Williams, 2017).

Rationale
A review was first published in 2009 with the title “Psychological therapies for the management of chronic pain (excluding headache) in adults” (Eccleston et al., 2009). It was updated in 2012 (Williams et al., 2012) and this version was stabilised in 2016. Cochrane reviews are stabilised because a judgement is made that no new evidence that has emerged since the publication is likely to change the estimate of effect. This is normally because there is insufficient novel evidence, or that the effect is judged to be so robust that there would need to be a very large body of high quality new evidence to alter the estimate of effect. Unusually, this review was stabilised because a decision was taken by the authors, supported by the Cochrane editorial team, that there was a large number of small and low quality studies that were adding nothing but noise to the system, and that these were unlikely to change the effect estimates. The authors argued for “…the immediate cessation of new RCTs of CBT against simple alternatives, unless a strong case can be given for the novelty of the population or treatment under investigation” (Williams et al., 2012) p. 15). However, despite repeated calls to address this research waste (Eccleston & Crombez, 2017; Eccleston, Morley, & Williams, 2013) it has continued.

Although the intention of stabilising the review was to positively influence trial and review production, it has paradoxically achieved the opposite. Rather than being considered stable, this 2012 review is now considered to be out-dated and irrelevant. Many reviews on this topic have been published: some are systematic, some are meta-analytic, and others describe the literature narratively. Most reviews do not include modern methods of quality and bias assessment. Typically, they use the date of the stabilised Cochrane review as part of their rationale, and indicate that an update of the literature is needed.

We have rescinded the decision to stabilise this review and will provide an update. The reasons for this decision is to provide an accurate, comprehensive and transparent review of the current state of trials, and to assess the evidence for or against our previous claim that the addition of new evidence leads to greater uncertainty. Updates of Cochrane reviews do not currently require protocols. The assumption is that the original protocol will be followed
and that any deviations from that protocol will be reported. However, the protocol for the
review was first published in 2008 (Eccleston, Morley, & Williams, 2008). In the last 10 years
there have been significant changes in both Cochrane methods and in the way psychological
treatments are discussed and labelled. The protocol is accurate, but incomplete. In the
interest of transparency we outline our a-priori decisions for the conduct of this update.
Please note some parts of the methods are taken from the Williams et al., (2012) review and
Cochrane Pain, Palliative, and Supportive Care (PaPaS) review group guidance (Cochrane
Pain, Palliative and Supportive Care Group, 2012). We have noted where our methods for
this update are new or significantly different from what went before.

Editorial Considerations
Given that this protocol for an updated Cochrane Review is novel within PaPaS, and will
produce changes the Editorial team will need to review we asked the Cochrane Editorial and
Methods group to quality-screen this updated protocol, and a PaPaS action editor and the
PaPaS editorial manager also reviewed it. We also took into account draft guidance from the
Cochrane Infectious Diseases Group, which was based on their consensus paper on how to
develop a protocol for updating reviews (Garner et al., 2016). This guidance is very much
based on making clear the areas of change in any new update that a review group editorial
team will need to focus on.

Objectives
To determine the clinical efficacy and safety of psychological interventions for the treatment
of chronic pain in adults (18 years of age or older) compared with active, wait-list, or
treatment as usual control.

Types of studies
We will include randomised controlled trials (RCTs) comparing a credible psychological
treatment, or a compound treatment with primary psychological content, with placebo,
other active treatment, treatment as usual, or waiting list control. We will exclude studies if
they were concerned with headache or associated with pain from a primary disease such as
cancer. We will exclude studies that were conducted remotely (phone, internet, app, or
equivalent) since these are well-reviewed elsewhere. We will judge a psychological
treatment credible if it is based on an extant psychological model or framework, and is
delivered by a healthcare professional qualified in psychology, or by another healthcare
professional with some psychology training and supervised by a healthcare professional
qualified in psychology.

We will include studies that meet the following criteria:

- are available as a full publication or report of a RCT;
- have a design that placed a psychological treatment as an active treatment of
  primary interest;
- have a face to face psychological treatment with definable psychotherapeutic
  content;
- are published (or electronically pre-published) in a peer-reviewed science journal;
- include participants reporting chronic pain (i.e. at least three months’ duration); and
- have 20 or more participants in each treatment arm at the end of treatment
  assessment.

We will keep the minimum criterion of N \geq 20 per arm at post-treatment assessment, similar
to the decision made in the 2012 update. Studies that include fewer participants at post-
treatment will be excluded.
Types of participants
We will include adults (aged 18 years of age or older) reporting pain of at least three months’ duration in any body site, not associated with a malignant disease. We will exclude patients with only headache or migraine because the psychological treatments for headache and migraine are sufficiently different.

Types of interventions
We will include studies if at least one trial arm consisted of a psychology intervention, with at least one comparator arm of a placebo condition, other active treatment, treatment as usual or waiting list control. Psychological interventions are classed as any intervention with specific content that is designed following psychological theory of behaviour and behaviour change. A typical example of a treatment with psychological content is a coping skills training intervention based on behaviour theory and cognitive theory, developed by an experienced clinical psychologist, and delivered by junior psychologists supervised by a senior and experienced psychologist. At least 50% of the content must be psychology, recognizing that often such treatments are delivered as packages of care alongside education, rest, exercise, relaxation, etc. A typical example of a treatment with insufficient psychological content is a mindfulness meditation treatment that refers only to education and meditation practice and has no theory to support behaviour change, or a treatment that refers to cognitive behavioural principles but is delivered by an unsupervised non-psychologist and has no recognizable psychological content.

Types of outcome measures
We have defined these outcomes in line with the previous two versions of this review, and with reference to the core outcome domains and measurement recommendations in the field (Dworkin et al., 2005). We have specified the following outcomes:

Primary outcomes:
- Pain intensity (e.g., Visual Analogue Scales, McGill Pain Questionnaire)
- Disability (e.g., Brief Pain Inventory interference items)
- Distress (e.g., Beck Depression Inventory).
- Adverse events

NEW: (1) In the 2012 version, we added a new outcome of catastrophic thinking about pain. This has been removed as an outcome for two main reasons. First, we agree with current thinking about catastrophising that it is a process variable rather than an outcome variable (Burns, Day, & Thorn, 2012); second, that its measurement has come under criticism over its conceptual clarity (Crombez, De Paepe, Veirman, Eccleston, & Van Ryckeghem, in submission).

Electronic searches
We will identify RCTs of any psychological therapy through databases including:
- Cochrane Central Register of Controlled Trials
- MEDLINE,
- EMBASE
- PsycINFO

A search strategy is provided below. We will search the databases from the last search (2012) until the end of July 2018. All previous included studies will be eligible for inclusion. We will not apply any language restrictions to the search.
Searching other resources
We will identify additional studies from the reference lists and citations searches of retrieved papers and from discussion with investigators. We will also search online trial registries including clinicaltrials.gov (www.clinicaltrials.gov) and World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (http://apps.who.int/trialsearch/). We will search the reference list of recovered reviews and selected papers for further unconsidered RCTs.

Selection of studies
We will include the trials used in the previous systematic review and meta-analysis provided that they still meet the eligibility criteria for this review (Williams et al., 2012). For post-2012 studies, two review authors [EF, AW] will independently determine eligibility by reading the abstract of each study identified by the search. Independent review authors will eliminate studies that clearly do not satisfy inclusion criteria, and obtain full copies of the remaining studies. Two review authors [EF, AW] will read these studies independently to select relevant studies and, in the event of disagreement, a third author will adjudicate [CE]. We will not anonymise the studies in any way before assessment. We will include a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart in the full review which will show the status of identified studies (Moher, Liberati, Tetzlaff, & Altman, 2009), as recommended in Part 2, Section 11.2.1 of the Cochrane Handbook (Chapter 12) (Higgins & Green, 2011). We will include studies in the review irrespective of whether measured outcome data are reported in a 'usable' way.

Data extraction and management
Two review authors [EF, AW] will independently extract data using a standard form and check for agreement before entry into Review Manager (RevMan, 2014). In the event of disagreement, a third author will adjudicate [CE]. We will extract the following information:

- Design of the study,
- Participants’ characteristics (e.g., age, sex)
- Primary diagnosis,
- Method of treatment
- Outcome measurement tools used.

We will also extract data relating to our chosen outcomes. For disability outcomes, we will preferentially extract disability measures if they are used. Where there is no disability/interference/impact score available, we will extract the physical component of the SF-36, or a physical component of quality of life, or whole scale if the content seems appropriate, although this is unlikely as most include subscales assessing psychological wellbeing. For distress outcomes, we will preferentially extract measures that combine anxiety and depression. If these are not reported, we will extract depression measures, followed by anxiety measures.

Assessment of risk of bias
We assessed risk of bias using the recommended Cochrane guidance (Higgins & Green, 2011). We will assess for failure to include sufficient methods of trial conduct to counter known biases. In Cochrane these methods are described as sources of bias as the assumption is that if they are missing or poorly described, that is the source of the bias. We assess the potential for bias by the extent to which these counter measures have been reported on, and the adequacy of the method taken. Cochrane organises the biases into five domains: (1) Selection bias which has two common sources – random sequence generation and allocation concealment; (2) Performance bias – blinding of participants (both those
delivering and those receiving interventions); (3) Detection bias – blinding of outcome assessment; (4) Attrition bias – incomplete reporting of data or inappropriate imputation; (5) Reporting bias – selective reporting.

Two authors [EF, AW or CE] will independently assess risk of bias for each study using the 'Risk of bias' tool in Review Manager (RevMan, 2014).

For this review we will assess the following sources of bias with the following judgements.

- **Random sequence generation** (checking for possible selection bias). We will assess the method used to generate the allocation sequence as: low risk of bias (any truly random process, e.g. random number table; computer random number generator); unclear risk of bias (method used to generate sequence not clearly stated). Studies using a non-random process (e.g. odd or even date of birth; hospital or clinic record number) will be excluded.

- **Allocation concealment** (checking for possible selection bias). The method used to conceal allocation to interventions prior to assignment determines whether intervention allocation could have been foreseen in advance of, or during recruitment, or changed after assignment. We will assess the methods as: low risk of bias (e.g. telephone or central randomisation; consecutively numbered sealed opaque envelopes); unclear risk of bias (method not clearly stated). Studies that do not conceal allocation (e.g. open list) will be rated as high risk of bias.

- **Incomplete outcome data** (checking for possible attrition bias due to the amount, nature, and handling of incomplete outcome data): We will assess the methods used to deal with incomplete data as being at low risk of bias (less than 10% of participants did not complete the study, or authors used 'baseline observation carried forward' analysis, or both), unclear risk of bias (e.g., used 'last observation carried forward' (LOCF) analysis), or high risk of bias (e.g., used 'completer' analysis).

- **Selective reporting** (checking for reporting bias). We will assess whether primary and secondary outcome measures were pre-specified and whether these were consistent with those reported: [add judgements for low, high and unclear risk].

We will not assess performance bias. Although we recognise that biases from the performance of agents in the trial, in particular actions that allow knowledge of which treatment is being delivered or received, can have a biasing effect on the outcomes of the trial, the standard counter-methods for managing this bias used in the Cochrane risk of bias tool are not relevant to psychotherapy interventions. Specifically, the tool considers performance bias to be best managed by the blinding of both patients and therapists to any knowledge of what treatment is being delivered.

**NEW:**

(2) Previously, we used the Yates scale for quality (Yates, Morley, Eccleston, & Williams, 2005). We will not conduct this assessment in this update and will remove reference to this method. The use of quality measurement is discouraged in Cochrane because such tools are often a mixture of quality and bias judgements. Additionally, our use of Yates pre-dates the adoption of GRADE in Cochrane reviews. Here we judged that the use of the Risk of Bias tool and the use of GRADE covered most of the relevant domains.

(3) The quality rating scale of Yates et al., (2015) did, however, include an item called ‘treatment expectations’ with a binary response of 0 or 1 for the absence or presence of any difference between groups of equivalence in treatment expectations. We decided to keep this item as a potential measure of at least one source of
performance bias. Trials will be graded as low risk of bias if a convincing effort to reduce bias in outcome measurement is reported. Trials will be graded as high risk of bias if no convincing effort to reduce bias in outcome measurement is reported.

Measures of treatment effect
The previous version of this review investigated two classes of psychological treatments: behaviour therapy (BT) and cognitive behavioural therapy (CBT).

NEW:

(4) This updated version will add a category of treatment labelled Acceptance and Commitment Therapy (ACT). The main reason for this addition is to bring the review in line with current developments in psychotherapy. ACT was included in the previous reviews because it was considered a form of cognitive behavioural therapy. Although there are strong arguments for it being considered a variant of CBT, there are also strong views that its differences outweigh its similarities and it should be considered separately (e.g., Hayes, Luoma, Bond, Masuda, & Lillis, 2006). There is precedent for this approach in the Cochrane Library (Churchill et al., 2013; Hunot et al., 2013; Naeem, Asmer, Khoury, Kingdon, & Farooq, 2015).

(5) We will exclude studies primarily using mindfulness protocols because we consider this field to be too heterogeneous (Van Dam et al., 2018). That said, we recognise that some trials of ACT may have components of mindfulness meditation involved. In these cases we will follow a rule previously used with multicomponent trials, that they will be included if the mindfulness component is no more than 20% of its overall content.

(6) Further, we recognise that there are psychological therapies developed which are not recognisable as BT, CBT or ACT, or which we may classify in this way while the originators or practitioners argue strongly for their difference. Therefore, we have created a final category of ‘other’. By definition this category will be either small or heterogeneous. Therefore, no meta-analysis will be attempted on this group, and each area will be reviewed narratively. In future updates we can revisit this category and its membership.

Two classes of comparator treatments are investigated and labelled active control and treatment as usual. The active comparator involves a treatment designed to change pain behaviour such as physical therapy, education or medical regime. Patients randomised to the active control within each trial all receive the same treatment. For patients assigned to a waiting list, trials vary in whether they provide further care, and patients vary in whether they seek further care. For patients assigned to treatment as usual, this treatment can consist of anything from regular consultations to access to care. Thus patients in these conditions receive variable and usually unrecorded treatment.

We will also select two assessment time points: post-treatment and follow-up. Post-treatment is the assessment point immediately following treatment, and follow-up is the assessment point at least six months after the end of treatment, but not more than 12 months, and the longer of the two if there were two follow-up assessments within this timeframe.

Therefore, twelve separate comparisons are designed comprising three classes of psychological treatment under investigation: Behavioural Therapy (BT) Cognitive
Behavioural Therapy (CBT) and Acceptance and Commitment Therapy (ACT). These will be compared with two forms of comparator: active comparators including sham or active other therapies (Active) and Treatment as Usual (TAU). Each treatment will be compared with Active or with TAU at two time-points, immediately post-treatment (T1) and at first follow-up (T2). We will combine data in a meta-analysis using standardised mean differences and 95% confidence intervals if possible. Analyses will be conducted for each of the comparisons below. Where a meta-analysis is not suitable, we will describe findings from studies qualitatively. The twelve comparisons are:

1. BT versus Active at T1
2. BT versus Active at T2
3. BT versus TAU at T1
4. BT versus TAU at T2
5. CBT versus Active at T1
6. CBT versus Active at T2
7. CBT versus TAU at T1
8. CBT versus TAU at T2
9. ACT versus Active at T1
10. ACT versus Active at T2
11. ACT versus TAU at T1
12. ACT versus TAU at T2

NEW:

(7) The ACT planned analyses (9-12) are new due to the separation of ACT trials from the CBT category in the previous versions.

The primary data type will be measurement using continuous scales. We will estimate treatment effects using standardised mean differences by extracting means, standard deviations and sample size at post-treatment and follow-up. Dichotomous outcome data based on clinical improvement are rare but if they exist we will extract these.

Multiple measurement tools are typically used in each trial. For each comparison we will identify four outcomes and labelled them ‘pain’, ‘disability’, ‘distress’ and ‘adverse events’. Although standard trial reporting guidance promotes the definition of primary outcomes (Boutron, Moher, Altman, Schulz, & Ravaud, 2008), most trials do not state a single or preferred a-priori primary outcome, so a judgment must be made. From each trial we will select the measure considered most appropriate for each of the three outcomes. When there was more than one measure for an outcome we gave preference to the measure that has frequent usage in the field as opposed to a novel measure. Also, when there was a choice between single item and multi-item self-report tools, we chose longer tools on the basis of inferred increased reliability. Not all trials will report data on all three outcomes of pain, disability and mood, and not all trials will report follow-up data.

Unit of analyses
The unit of evaluation will be the participant. Where a trial has more than two arms, we will select those which best match our requirements for therapies, and where there is a choice, the most intensive version of either: for example, if a trial had an enriched CBT (that is, CBT with additional non-core components such as vocational guidance), a minimum CBT and a waiting list condition, we compared the enriched CBT with the waiting list. If both options seem similarly ‘intensive’ we will follow the Cochrane handbook guidance (section 16.5.4) and include multiple relevant arms in the same analysis if necessary, for example by splitting
the control group data.

**Dealing with missing data**

We will contact authors where there are missing data. Where data are still missing we will impute with the most conservative method.

**Assessment of heterogeneity**

We will assess heterogeneity according to the standard method using the Chi\(^2\) test and the I\(^2\) statistic, calculated for each comparison on each outcome. I\(^2\) values will be interpreted according to the Cochrane handbook.

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

**Sub-group analyses**

No sub-group analyses are planned, because there is no strong a priori reason in the literature to look at different classes of treatment, characteristic of treatment, or characteristic of participant.

**Assessment of reporting biases**

We will assess reporting biases by assessing funnel plots if there are sufficient studies for such an analysis.

**Data synthesis**

**Quality of the evidence**

In this update, two review authors (EF, AW or CE) will independently rate the quality of the outcomes. We will use the GRADE (Grades of Recommendation, Assessment, Development and Evaluation) system to rank the quality of the evidence using the GRADEprofiler Guideline Development Tool software (GRADEpro, 2015), and the guidelines provided in the Cochrane Handbook for Systematic Reviews of Interventions (Chapter 12) (Higgins & Green, 2011). The GRADE approach uses five considerations (study limitations, consistency of effect, imprecision, indirectness, and publication bias) to assess the quality of the body of evidence for each outcome. The GRADE system uses the following criteria for assigning grade of evidence:

- High: we are very confident that the true effect lies close to that of the estimate of the effect (randomised trials or double-upgraded observational studies).
- Moderate: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different (downgraded randomised trials; or upgraded observational studies).
- Low: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect. (Double-downgraded randomised trials; or observational studies).
- Very low: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect. (Triple-downgraded randomised trials; or downgraded observational studies; or case series/case reports).

Factors that may decrease the quality level of a body of evidence are:

1. limitations in the design and implementation of available studies suggesting high
likelihood of bias;
2. indirectness of evidence (indirect population, intervention, control, outcomes);
3. unexplained heterogeneity or inconsistency of results (including problems with subgroup analyses);
4. imprecision of results (wide confidence intervals);
5. high probability of publication bias.

We will decrease the grade rating by one (-1) or two (-2) (up to a maximum of -3 to 'very low') if we identify:
• Serious (-1) or very serious (-2) limitation to study quality;
• Important inconsistency (-1) or serious inconsistency (-2);
• Some (-1) or major (-2) uncertainty about directness;
• Serious (-1) or very serious impression (-2)
• High probability of reporting bias (-1).

'Summary of findings' table
We plan to include three 'Summary of findings' tables to present the main findings for CBT, BT, and ACT in a transparent and simple tabular format. Where enough data are available, we will present comparisons of intervention vs. active control. In particular, we will include key information concerning the quality of evidence, the magnitude of effect of the interventions examined, and the sum of available data on the outcomes pain (post-treatment and at follow-up), disability (post-treatment and at follow-up), distress (post-treatment and at follow-up), and adverse events.

Sensitivity Analysis
We plan to explore the influence of expected imprecision in measurement that should obtain from the relatively low n of entry of studies of 20 participants with further sensitivity analyses based on a n of 50 participants in the treatment arm at the time point being compared (T1 or T2).

Contributions of authors
All authors contributed to the update protocol. Amanda Williams is the guarantor of the review.

Funding and conflicts of interest
Cochrane Review Group funding acknowledgement: the editing of this updated review will be supported by the National Institute for Health Research, via Cochrane Infrastructure funding to the Cochrane Pain, Palliative and Supportive Care Review Group (PaPas). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS or the Department of Health.

Authors do not have any conflicts of interest.

References


