Evidence-based psychological interventions for adults with chronic pain: precision, control, quality, and equipoise

Amanda C de C Williams[1*], Emma Fisher[2,3], Leslie Hearn[2], Christopher Eccleston[3].

[1]Research Department of Clinical, Educational and Health Psychology, University College London, London, UK;
[2]Cochrane Pain, Palliative and Supportive Care Review Group, Pain Research Unit, Churchill Hospital, Oxford, UK;
[3]Centre for Pain Research, University of Bath, UK.

No. of text pages: 11
No. of figures: 0
No. of tables: 2

*Corresponding author. Address: Research Department of Clinical, Educational and Health Psychology, University College London, 1-19 Torrington Place, London, WC1E 7HB, UK. E-mail address: amanda.williams@ucl.ac.uk. Tel.: +44 207 6791608.

Evidence-based psychological interventions for adults with chronic pain: precision, control, quality, and equipoise

1. Introduction

Psychological interventions for adults with chronic pain are common and desirable treatment options in multidisciplinary pain treatment centers operating in high resource countries. Evidence for the efficacy and safety of psychological treatments, typically from randomized controlled studies, has been repeatedly summarized in systematic reviews and meta-analyses over the last 40 years. Our group has maintained the Cochrane Library review on 'Psychological interventions for adults with chronic pain (excluding headache)' since its inception in 1999, updating it with the latest evidence, most recently in August 2020 [7,22,29,30].

In the 1999 review, we included 25 randomized controlled trials (RCTs); in 2009, 40; in 2012 (with stricter inclusion criteria), 35; and, in 2020, 75. The main findings of the 2020 review are summarized in Table 1. Our aim in this topical review is to draw attention to key features of RCTs from which we draw our evidence base of psychological treatments, including design, conduct and reporting of trials. We do not address here other valuable methodologies, such as single case and process studies. If we are to produce trusted, high-quality evidence that is useful, then we must attend to issues that can potentially undermine that trust, issues that add to uncertainty around the evidence. These include, but are not limited to, the *precision* of our effect estimates as a function of trial size; the *quality* of treatment content; *control* over therapy content in the use of comparator treatments; and the degree of *equipoise* in how questions are asked and answered. We conclude with recommendations for the future direction of the field.

2. Critical methodological concerns

2.1. Precision

Precision refers to the variability of an effect estimate, usually indicated by the width of the confidence interval [2]. RCT sample size is a critical factor when determining the estimate of effect in a meta-analysis. Of course, small trials are sometimes useful, for example when establishing feasibility. However, when determining efficacy and harms of an intervention, individual trials with small sample sizes risk imprecision. This is true even when many (small) studies are pooled with the rigor of Cochrane methodology [18,27]. Imprecision does not, by itself, bias in any single direction, but it does inflate the risk of type II error, i.e., the failure to reject a false null hypothesis. However, it is well documented that publication bias for positive trials leads to over-representation in the literature of *small* positive trials [9,10,15,20]. When included in meta-analyses, smaller trials, whatever their absolute size, generate, on average, larger effect sizes than do larger trials [4,5]. The inflation can be substantial: the quartile of smallest-sized studies in the Dechartres et al. [4] meta-epidemiological study produced estimates a mean 23% higher than the other three quartiles.

Concern about small trials, and the imprecision of their pooled results in meta-analyses, encouraged the Cochrane Pain, Palliative, and Supportive Care review group (https://papas.cochrane.org) to investigate size as a risk-of-bias (RoB) domain for assessment. Reanalysis of existing systematic reviews using only larger trials, including psychological studies [e.g. 8], often shows smaller effect sizes that can change conclusions. In our recent Cochrane review and its forerunner [29,30], we excluded studies with fewer than 20 participants per arm post-treatment, but this still allowed inclusion of trials whose size occasioned a high RoB. An inclusion criterion of 20 per arm is still overly liberal. The required size of a trial varies with the true event rate in both comparison and intervention groups. For a given rate in the comparison group, required size increases with the expectation of a clear statistically significant effect, and with lower event rates in the intervention group. For example, in single-dose pharmaceutical trials [20] over 500 per arm

are required. With 40 participants per arm, a common sample size in chronic pain trials, we can only have confidence in the direction, not size, of effect. Including 50 participants per arm as the minimum is the prerequisite for merely unclear, rather than high, RoB.

We revisited the evidence base used in the 2020 Cochrane review and conducted sensitivity analyses, repeating the meta-analyses using only those trials with at least 50 in either arm and reassessed the evidence using criteria from GRADE (Grading of Recommendations Assessment, Development and Evaluation), a transparent and systematic method of grading certainty of evidence [26]. These sensitivity analyses are presented in Table 1.

Analyses in the 2020 review were organized according to psychological therapy type (cognitive behavioral therapy (CBT); behavioral therapy (BT); acceptance and commitment therapy (ACT); and other therapies) and by control comparison (active control (AC) or treatment as usual/waiting list control (TAU)). There were insufficient trials for re-analyses of BT vs either control, of ACT vs TAU, or of disability outcomes for ACT vs AC.

Table 1 here

Including only larger trials changed the effect estimates. For all ACT analyses, in line with previous studies in pain, including only larger sample sizes substantially reduced the effect estimates. The evidence remained of very low certainty although the confidence intervals narrowed. For CBT, however, the findings were surprising. For all analyses of CBT vs AC, effect sizes and confidence intervals remained largely unchanged, but GRADE ratings improved from low or moderate certainty to high certainty. By contrast, for CBT vs TAU, the size of the effect estimates increased, and GRADE ratings improved from low or moderate certainty to all moderate certainty.

In summary, these sensitivity analyses revealed that the evidence base for ACT is of very low certainty and the true effect could lie outside the confidence intervals [14]. However, for CBT, increasing the precision improves the certainty of the evidence, leaves the estimates unchanged for CBT vs AC, and improves them for the comparison with TAU, showing that effect estimates for CBT are stable.

2.2. Treatment quality

In standard RoB methods, the inability to blind an operator (e.g. surgeon or therapist) to the treatment being delivered is automatically considered a risk of bias [11], introducing unwanted variability. However, non-pharmacological interventions such as surgery, physical therapy and psychological therapy cannot separate the production of the treatment from its delivery [6]. While treatments and delivery can be standardized by protocols and training, treatment content and delivery remain somewhat flexible in response to dynamic features of the context, such as the patient presentation.

Because double blinding is not possible in trials of psychological interventions, more attention is typically paid to characteristics of delivery. All of the 75 trials in our recent review featured qualified psychological therapists delivering psychotherapeutic content face-to-face (a condition of inclusion). We know less about the specific protocols for therapy content in each treatment, largely due to scant reporting. The recent promotion of the TiDier checklist [12], which encourages and guides content reporting in full, will improve this situation. Also rarely discussed, though important, are the treatment models adopted in non-pharmacological trials – the theories of mechanism that underlie delivery. The rich pre-clinical science in pain is often poorly translated to clinical trials, or not obvious from their reporting. Finally, there is also discussion on the potential role of variables common to all forms of therapy, such as therapist alliance [13], and their potential role trans-diagnostically [1,28,31]. For secondary data analysis of published studies to fulfill its promise, we need to share not only data but manuals, content, and theoretical rationales.

2.3. Control

In the RCT of a psychological intervention, the choice of comparator, ideally a placebo, follows directly from concern about treatment content and quality. A sham treatment should be built to have 'carrier' characteristics similar to those of the active treatment (same therapist, same dose, same location, etc.) but no therapeutically active content. Although sham treatments are possible, they are uncommon as they are considered difficult to create and sustain. More common is the use of random allocation to another treatment considered equally engaging. In this review, 39 of the 75 trials used various ACs. The assumption is that an active treatment will mirror some 'carrier' characteristics, such as dose and location, but may also introduce placebo characteristics, such as patient expectation of benefit and therapist belief in the treatment. Common active comparators are education, exercise, and relaxation training, the commonest in our review being education.

Forty-four of the 75 trials employed a waiting-list or treatment-as-usual (both subsumed under TAU) comparison (eight studies used both AC and TAU). TAU has the value of ecological validity, representing normal practice, although it may be that the 'Hawthorne effect' of being in a trial alters key outcomes. Simply put, in chronic pain, TAU in the real (non-trial) world covers a range of possibilities, from no treatment or engagement with health care to attentively delivered treatment: these are distributed at random across control groups.

More work is needed to identify the optimal control condition for psychological interventions. However, given the effects for CBT reported in Williams et al. [30], supported by the sensitivity analyses reported here, it is likely that trials in which one arm is no/inert treatment will be increasingly rejected as unethical. In future, consideration will need to be given to alternative and emerging techniques, such as ACT or emotion-focused therapy [17], being entered into non-inferiority comparisons with a standardized CBT protocol to replace both active and TAU comparisons.

2.4. Equipoise

Therapists have an ethical duty to deliver the treatment they believe to be right for their patients, so permitting randomization to an inert control could violate patient rights and call into doubt this duty. The principle of equipoise allows for ethical use of selection to an inert treatment if the evidence for a treatment is in equipoise [16] or, more simply, if there is sufficient professional uncertainty about the efficacy and safety of an intervention [6,16,25]. From this perspective, equipoise is not a clinician quality, but an ethical position adopted in response to uncertainty about treatment. Holding the likelihood of efficacy and safety in the balance throughout a study is an important principle. The extent to which an individual actor within a trial, whether as initiator, operator, regulator, statistician, author, or broadcaster, has individual balance is less important than a) whether the actors transparently disclose their prior beliefs about evidence, and b) whether attempts have been made to counter the potentially biasing effect of those beliefs on trial conduct and outcome.

On the first point – disclosure – it is relatively rare in trials of psychological interventions to assess or disclose actors' prior beliefs on the potency or otherwise of an experimental treatment. Not only can it indicate a potential threat to equipoise (e.g. if all actors believe that a treatment is ineffective) but it can also go some way to addressing concerns for what have been called academic conflicts of interest (e.g. [3]), in which a person might be judged to indirectly benefit from the result of a study in areas such as professional advancement or reputation.

On the second point – blinding to patient allocation – it is rare to blind actors other than patients, although it is possible to attempt to manage the influence of prior therapist belief by blinding staff involved in assessment (only reported in 17 of our 75 trials) and by the

operation of an independent clinical trial committee. In psychological trials, blinding of therapists is impossible, but patient and therapist expectations of outcome and the strengths of those beliefs can be recorded. Analyses of the influence of therapist expectations and beliefs on outcomes should be included in the protocol and in the clinical trial report.

3. Discussion

In this topical review, we have outlined the importance of precision, treatment quality, control, and equipoise being core considerations in the next generation of trials, in order to bring value to patients. Unheeded, the next generation of evidence will not advance the field but merely add more 'me-too' trials to the literature. Replication is of course important: however, we have reached a stage where replication of CBT vs TAU or AC incurs a high cost in lost opportunities for alternative developments. In 2013 [23] and in 2018 [6], we called for an end of trials comparing CBT to control, the case for the efficacy of CBT compared with doing nothing having been made, and yet we are still seeing a large number of these trials added to the literature each year. It is time for a change in direction.

The next generation of evidence needs to be innovative, yet still include the core features of modern trial design, summarized in Table 2. From this discussion, we have highlighted that the next generation of trials should: i) have at least 50 participants completing treatment in each arm; ii) fully disclose the content and rationale of treatments, including at least a TiDieR checklist and supporting documentation; iii) include a standard CBT control for testing both non-inferiority or superiority; iv) use a standardized CBT comparator, and consider a sham comparator, when investigating new therapies; and v) require principal actors in trial design, conduct, and delivery to disclose a priori beliefs about the outcome and the value of treatments under investigation to allow for an independent exploration of equipoise. While these do not guarantee capture of all unaccounted variance in trials, they do represent improved standards from which we can further progress.

Table 2 here

We still face challenges. First, we know that not all psychological treatments are alike. Whilst CBT shows efficacy (and we are discouraging future trials comparing CBT to TAU or AC), there is still important work to be done in determining treatment mechanisms and, indeed, tailoring of treatments. Pre-clinical science can usefully guide the improvement of specific CBT content. Dropout is still fairly common in these trials and understanding more about patients who do not continue treatment is an important future direction for the field. Second, more investment should be put into novel therapeutic options, to be compared with a standard CBT treatment. ACT in particular is in need of a large-sized, multi-site trial delivered to strict protocol, with a majority of independent contributors with no strong prior beliefs about the value of the therapy. The same applies to other novel therapies, so that clinicians are equipped with a range of evidence-based therapies to offer patients with chronic pain. Third, there is a need for grant agencies, journal editors, and systematic reviewers to only fund or accept those studies that follow modern evidence-based practice and to reject or exclude those that do not.

In conclusion, this is very much a turning point at which we can choose to forge a new direction in psychological treatments for chronic pain, by considering precision, treatment quality, control, and equipoise in how evidence is generated and how it is presented to clinical and scientific communities.

Acknowledgments

Thanks to Andrew Moore for useful discussions.

Conflicts of interest

There are no conflicts of interest for any author.

References

1. Baier AL, Kline AC, Feeny NC. Therapeutic alliance as a mediator of change: a systematic review and evaluation of research. Clin Psychol Rev 2020;82:101921. https://doi.org/10.1016/j.cpr.2020.101921
2. Brand A, Bradley MT. The precision of effect size estimation from published psychological research: surveying confidence intervals. Psychol Rep 2016;118:154-70.
3. Cochrane Collaboration. Policy: conflicts of interest and Cochrane Reviews. 2020. https://documentation.cochrane.org/display/EPPR/Policy%3A+conflicts+of+interest+and+Cochrane+Reviews (accessed 24 Nov 2020).
4. Dechartres A, Altman DG, Trinquart L, Boutron I, Ravaud P. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. JAMA 2014;312:623-30. doi: 10.1001/jama.2014.8166.
5. Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 2013;346:f2304. doi: 10.1136/bmj.f2304.
6. Eccleston C, Crombez G. Advancing psychological therapies for chronic pain**.** F1000Research 2017;6:461. doi: 10.12688/f1000research.10612.1.
7. Eccleston C, Williams ACdeC, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2009;(2):CD007407. doi: 10.1002/14651858.CD007407.pub2.
8. Fisher E, Law E, Dudeney J, Palermo TM, Steward G, Eccleston C. Psychological therapies for the management of chronic and recurrent pain in children and adolescents. Cochrane Database Syst Rev 2018;9(9):CD003968. doi: 10.1002/14651858.CD003968.pub5.
9. Glasziou P, Chalmers I. Research waste is still a scandal – an essay by Paul Glasziou and Iain Chalmers. BMJ 2018;363:k4645 doi: 10.1136/bmj.k4645.
10. Guyatt GH, Mills EJ, Elbourne D. In the era of systematic reviews, does the size of an individual trial still matter? PLoS Med 2008;5(1):e4. doi: 10.1371/journal.pmed.0050004.
11. Higgins JPT, Savović J, Page MJ, Sterne JAC (eds) on behalf of the RoB2 Development Group. Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). 2019. https://riskof bias.info/welcome/rob-2-0-tool.
12. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, Altman DG, Barbour V, Macdonald H, Johnston M, Lamb SE, Dixon-Woods M, McCulloch P, Wyatt JC, Chan A-W, Michie S. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687 doi: 10.1136/bmj.g1687.
13. Horvath AO. The alliance. Psychotherapy 2001;38:365-72.
14. Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, Alper BS, Meerpohl JJ, Murad MH, Ansari MT, Katikireddi SV, Ostlund P, Tranaeus S, Christensen R, Gartlehner G, Brozek J, Izcovich A, Schünemann H, Guyatt G. The GRADE Working Group clarifies the construct of certainty of evidence. J Clin Epidemiol 2017;87:4-13. http://dx.doi.org/10.1016/j.jclinepi.2017.05.006
15. Ioannidis JPA. Why most published research findings are false. PLoS Med 2005;2(8):e124.
16. John A. Equipoise in research: integrating science and ethics in human research*.* JAMA 2017;317(5):525-6. doi: 10.1001/jama.2017.0016.

17. Lumley MA, Schubiner H, Lockhart NSA, Kidwell KM, Harte SE, Clauw DJ, Williams DA. Emotional awareness and expression therapy, cognitive behavioral therapy, and education for fibromyalgia: a cluster-randomized controlled trial. PAIN 2017;158(12):2354-63.
18. Moore RA, Derry S, McQuay HJ. Fraud or flawed: adverse impact of fabricated or poor quality research. Anaesthesia 2010;65(4):327-30. doi: 10.1111/j.1365-2044.2010.06295.x.
19. Moore R, Eccleston C. Safe methods of imputation for clinical trials of interventions for chronic pain: promoting transparency and comparison**.** PAIN 2020;161(10):2225-6.
20. Moore RA, Gavaghan D, Tramèr MR, Collins SL, McQuay HJ. Size is everything – large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. PAIN 1998;78:209-16. doi: 10.1016/S0304-3959(98)00140-7.
21. Moore AR, Straube S, Eccleston C, Derry S, Aldington D, Wiffen P, Bell RF, Hamunen K, Phillips C, McQuay H. Estimate at your peril: Imputation methods for patient withdrawal can bias efficacy outcomes in chronic pain trials using responder analyses. PAIN 2012;153(2):265-8.
22. Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. PAIN 1999;80(1-2):1-13.
23. Morley S, Williams A, Eccleston C. Examining the evidence of psychological treatments for chronic pain: time for a paradigm shift? PAIN 2013;154(10):1929-31.
24. Palermo TM, Slack K, Loren D, Eccleston C, Jamison RN. Measuring and reporting adverse events in clinical trials of psychological treatments for chronic pain. PAIN 2020;161(4):713-7.
25. Sackett DL. Equipoise, a term whose time (if it ever came) has gone. CMAJ 2000;163(7):835-6.
26. Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, Guyatt GH. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions, version 6.1, Cochrane, 2020. Available from www.training.cochrane.org/handbook.
27. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PLoS One 2013;8(3):e59202. doi: 10.1371/journal.pone.0059202.
28. Wampold BE, Imel ZE. The great psychotherapy debate: The evidence for what makes psychotherapy work. Hove UK: Routledge; 2015.
29. Williams ACDC, Eccleston C, Morley S. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2012;11(11):CD007407. doi: 10.1002/14651858.CD007407.pub3.
30. Williams ACdeC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2020;8(8):CD007407. doi: 10.1002/14651858.CD007407.pub4.
31. Williams ACdeC, Morley SJ. Conducting and evaluating treatment outcome studies. In: Turk DC, Gatchel RJ (eds). Psychological approaches to pain management (3rd ed). New York, Guilford Press, 2018.
32. Yates SL, Morley S, Eccleston C, Williams ACdeC. A scale for rating the quality of psychological trials for pain. PAIN 2005;117:314-25.

Table 1. Meta-analyses of cognitive behavioral therapy vs active control and treatment as usual, and acceptance and commitment therapy vs active control, showing results of trials including ≥ 50 participants per arm and results of any trial including ≥ 20 participants per arm.

Table 2. Core features of the best trials

Table 1

Outcome	Studies	Participants	SMD (IV, Random, 95% CI)	GRADE certainty of evidence
Cognitive behavioral therapy vs active control				
Pain post-treatment				
n ≥ 50/arm	12	2395	-0.08 [-0.16, -0.00]	High
n ≥ 20/arm	23	3235	-0.09 [-0.17, -0.01]	Moderate
Disability post-treatment	10	1893	-0.11 [-0.20, -0.01]	High
n ≥ 50/arm	19	2543	-0.12 [-0.20, -0.04]	Moderate
n ≥ 20/arm				
Distress post-treatment				
n ≥ 50/arm	12	2394	-0.11 [-0.19, -0.03]	High
n ≥ 20/arm	24	3297	-0.09 [-0.18, 0.00]	Moderate
Pain follow-up				
n ≥ 50/arm	10	1943	-0.01 [-0.11, 0.08]	High
n ≥ 20/arm	16	2362	-0.08 [-0.19, 0.04]	Moderate
Disability follow-up				
n ≥ 50/arm	8	1454	-0.05 [-0.17, 0.07]	High
n ≥ 20/arm	15	1919	-0.12 [-0.26, 0.02]	Low
Distress follow-up				
n ≥ 50/arm	10	1943	-0.07 [-0.16, 0.02]	High
n ≥ 20/arm	16	2362	-0.13 [-0.25, -0.01]	Moderate
Cognitive behavioral therapy vs treatment as usual				
Pain post-treatment				
n ≥ 50/arm	5	728	-0.40 [-0.55, -0.25]	Moderate
n ≥ 20/arm	29	2572	-0.22 [-0.33, -0.10]	Moderate
Disability post-treatment	6	843	-0.42 [-0.59, -0.24]	High
n ≥ 50/arm	28	2524	-0.32 [-0.45, -0.19]	Low
n ≥ 20/arm				
Distress post-treatment				
n ≥ 50/arm	7	1037	-0.41 [-0.56, -0.26]	Moderate
n ≥ 20/arm	27	2559	-0.34 [-0.44, -0.24]	Moderate
Pain follow-up				
n ≥ 50/arm	6	867	-0.27 [-0.41, -0.14]	High
n ≥ 20/arm	15	1674	-0.16 [-0.27, -0.04]	Moderate
Disability follow-up				
n ≥ 50/arm	5	694	-0.45 [-0.68, -0.22]	Moderate
n ≥ 20/arm	15	1581	-0.21 [-0.37, -0.05]	Low

Distress follow-up				
n ≥ 50/arm	6	869	-0.31 [-0.48, -0.13]	High
n ≥ 20/arm	16	1757	-0.25 [-0.37, -0.13]	Moderate
Acceptance & commitment therapy vs active control				
Pain post-treatment				
n ≥ 50/arm	2	228	-0.39 [-1.25, 0.48]	Very low
n ≥ 20/arm	5	385	-0.54 [-1.20, 0.11]	Very low
Distress post-treatment				
n ≥ 50/arm	2	228	-0.41 [-1.39, 0.56]	Very low
n ≥ 20/arm	5	385	-0.61 [-1.30, 0.07]	Very low
Pain follow-up				
n ≥ 50/arm	2	212	-0.18 [-0.99, 0.63]	Very low
n ≥ 20/arm	3	265	-0.38 [-1.03, 0.27]	Very low
Distress follow-up				
n ≥ 50/arm	2	212	-0.35 [-1.10, 0.40]	Very low
n ≥ 20/arm	3	265	-0.58 [-1.24, 0.07]	Very low

GRADE certainty of evidence: “Very low” means that the true effect is probably markedly different from the estimated effect; “low” means that the true effect might be markedly different from the estimated effect; “moderate” means that the true effect is probably close to the estimated effect; “high” means that the authors have a lot of confidence that the true effect is similar to the estimated effect [26].

Table 2

Feature	Rationale
Features discussed in topical review	
Precision: trials including ≥ 50 participants/arm, preferably ≥ 100/arm	Gives greater power and precision; reduces type II error
Content of intervention and comparison based on theory; clear statement of presumed mechanisms of change	Enables more confident interpretation of results
Placebo controls designed for specific aspects of treatment, based on theory or presumed mechanisms of change	Allows more confident interpretation of differences between intervention/s and control/s
CBT used as a standard comparison treatment	CBT has robust efficacy; new treatments should be evaluated against it before being adopted in practice
Treatment content described; TiDieR checklist used [12]; quality and delivery of treatment reported [32]	Allows replication and detailed analysis
Disclosure by treatment staff of prior beliefs about efficacy and value of treatment(s) under investigation.	Avoids bias in delivery, assessment, analysis and reporting; allows for an independent assessment of equipoise
Other essential features	
Protocol publicly pre-registered	Commits to conducting and reporting results according to protocol
Randomization method adequate	Reduces risk of flaws in results due to biased randomization and allocation procedures
Participants blinded; expectations of outcome assessed at baseline	Minimizes or controls for differences in expectations between intervention/s and control/s
Adverse events and full explanations of attrition included in outcome assessment	Identifies possible harms [24], including those associated with dropout or treatment failure

Primary analyses only undertaken when participants have more than mild symptoms, e.g. at least moderate pain at baseline in an analgesic trial	Only includes those with needs relevant to the intervention
Baseline observation carried forward for missing data	Reporting data only from participants who have completed the intervention inflates effects [19,21]
Data published	Reduces selective reporting and publication bias; allows pooling of datasets and individual patient analyses

ICMJE Conflict of Interest Form--REQUIRED from ALL authors of submission at revision stage

ICMJE Conflict of Interest Form--REQUIRED from ALL authors of submission at revision stage

ICMJE Conflict of Interest Form--REQUIRED from ALL authors of submission at revision stage

ICMJE Conflict of Interest Form--REQUIRED from ALL authors of submission at revision stage