

# Evaluating the impact of different control states in current psychiatric research design<sup>†</sup>

Harry Williams 

## SUMMARY

A ‘control’ provides a point of clinical comparison for a new intervention, allowing researchers and clinicians to draw more confident conclusions about the effectiveness or potential harm of a given, often novel, therapy. Although this aspect of a trial’s design provides the basis from which interventional impact is measured, it is often less closely examined. This commentary appraises a Cochrane Review that compares various controls in common use in modern psychiatric research and aims to characterise their effects on the outcomes of that research.

## KEYWORDS

Randomised controlled trial; research methodology; control conditions; outcome studies; statistical methodology.

Importantly, the control group provides comparison for both positive and negative effects of the intervention, offering insight into the clinical benefit and the risk involved in offering a particular therapy. The selection of an appropriate clinical control therefore both affects decision-making regarding effective clinical interventions and ensures patient safety in the avoidance of potential harm (Sibbald 1998).

## Particular challenges of psychiatric research

The importance of control design within clinical research methodology has been studied more broadly, including by one of the authors of this month’s Cochrane Corner review in a previous meta-analysis (Hróbjartsson 2010). That study included 234 trials, covering 60 clinical conditions, including psychiatric conditions such as schizophrenia. It showed that the administration of a placebo had small and uncertain positive effects on clinical outcome.

Within psychiatric literature more specifically, the design of control conditions has historically received less attention (Mohr 2014). The specific challenges encountered in psychiatric research are therefore less well-known. Psychiatric research must manage the variability seen between different modalities of commonly used control: pharmacological, psychological and physical controls (Box 2). Each of these introduces its own challenges: psychological controls have demonstrated greater effects than waiting-list controls in anxiety disorders (Patterson 2016), and significant differences were seen between waiting-list controls and usual care (also referred to as treatment as usual or TAU) controls when compared with pill placebo or no-treatment controls in the management of depressive disorders (Mohr 2014). Even within individual and well-established control conditions such as usual care, there is variability introduced in the way this is provided (Bellg 2004).

The design of controls in psychiatric research has therefore been shown to have an impact on outcome effects but broader meta-analysis of these effects has not been undertaken.

**Harry Williams**, MBBS, BSc (Hons), MA, is a foundation trainee within Buckinghamshire Health NHS Trust, currently based at Stoke Mandeville Hospital, Aylesbury, UK. He has an interest in experimental medicine and research methodology.

**Correspondence** Dr Harry Williams. Email: [harry.williams11@nhs.net](mailto:harry.williams11@nhs.net)

First received 13 Feb 2023

Final revision 4 Apr 2023

Accepted 5 Apr 2023

## Copyright and usage

© The Author(s), 2023. Published by Cambridge University Press on behalf of the Royal College of Psychiatrists

<sup>†</sup>Commentary on... Control interventions in randomised trials among people with mental health disorders (Cochrane Corner). See this issue.

## Background

### *The historical and clinical context of control conditions*

New interventions have been compared experimentally with a ‘control’ as far back as Lind’s scurvy trial in 1747, although their commonplace inclusion in research design emerged in the first half of the 20th century (Bothwell 2016). Controls form one aspect of the standardised methodology intended to allow a causative relationship to be drawn between intervention and effect, alongside other cornerstones of clinical research: randomisation, double blinding and intention-to-treat analysis (Schulz 1995; Sibbald 1998).

The control group is intended to reduce, to the greatest degree possible, the effect of a series of confounding factors that might affect the outcome of a study (Box 1) (Mohr 2009, 2014; Patterson 2016). By reducing the variability between intervention and control groups, trial designers aim to causatively link any differences between groups to the intervention under investigation. This in turn allows for more confident conclusions about the specific effects of the intervention being studied.

### BOX 1 Common confounding factors in clinical trials

#### Natural history

Improvement in the condition that would have occurred without any intervention at all.

#### Regression to the mean

The tendency for extreme results generated by limited samples to move closer to the mean result when the sample is expanded or the study repeated.

#### Placebo/nocebo effect

The positive (placebo) or negative (nocebo) effect of an inactive treatment. The exact mechanism of placebo/nocebo is still unclear

but is thought to be due to a combination of patient expectancy, classical conditioning, etc.

#### Comparator bias

A design in which the experimental intervention is compared with a control known to be less effective than another. For example, comparing a novel pharmacological therapy with a placebo as opposed to the current standard of care would compare the novel therapy with a weaker comparator and therefore be likely to misrepresent its effect.

#### *Lack of consensus on a standard approach*

The variance in the effect of the controls outlined above raises a clear concern when trying to discern clinical applicability within psychiatric research: to what degree is the choice of control responsible for the difference seen between the intervention and comparator groups?

The lack of consensus in control design, particularly within psychiatric research, and consequent lack of consistency in the application of control

conditions makes it difficult to discern the specific impact of the intervention separate from the study design (Gold 2017). Added to this, the variability in the control states used makes comparison between studies using different controls more complex. The need for more guidance and consistency in psychiatric research control conditions is an ongoing concern raised by several authors (Schulz 1995; Mohr 2009; Gold 2017).

### The Cochrane Review

#### Summary

This month's Cochrane Corner review (Faltinsen 2022) aimed to assess whether commonly used control states in psychiatric research resulted in differing estimates of intervention effect or different incidences of adverse events when compared with both no-treatment and waiting-list controls. It also compared usual care with both waiting-list and no-treatment both in intervention effect and number of adverse events. Finally, the review authors directly compared waiting-list with no-treatment to see whether any difference in observed effect could be seen.

The population addressed was deliberately broad: namely, any participant with a mental disorder in a randomised trial that included a control arm and a waiting-list or no-treatment arm. The population was not limited by the type of psychiatric diagnosis, although a formal psychiatric diagnosis was an inclusion criterion, and included participants from both in-patient and out-patient settings.

#### Method

The review took as its interventions those methods that were typically control states in the original trials, for example inactive pharmacological, physical or psychological interventions, or usual care, and compared them with no-treatment or waiting-list conditions. Included trials compared a placebo, waiting-list or usual care group with a waiting-list or no-treatment group.

The search strategy for this meta-analysis was wide, using a clear protocol developed from the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019). Trials were drawn from 17 databases and trial registers and were not limited by language, year or publication type.

Initially this returned 64 529 records, which following screening resulted in the final inclusion of 96 randomised trials. Of the included 96 trials, 83 were able to contribute usable data on a total of 3614 participants. Quality of evidence was rated against the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria (Atkins 2004).

### BOX 2 Modalities of control in controlled studies

#### Pharmacological control

Pharmacological controls are a comparator group in which sham medications are given that match the experimental drug as closely as possible (in form, shape, colour, flavour, smell, etc.) but lack the active ingredient thought responsible for the therapeutic effect. The control sample therefore has the same experience as participants taking the treatment under investigation but not the specific intervention under investigation.

#### Physical control

Physical controls include sham procedures (such as sham surgery, electroconvulsive therapy (ECT), deep brain stimulation (DBS), acupuncture/acupressure, etc.) in which the patient has the same experience as if they went through the true procedure but in which the intervention is not in fact performed (e.g. in the case of sham ECT or DBS no current is passed through the electrodes).

#### Psychological control

Psychological controls refer to undirected interaction between clinician and patient which does not feature the directed, structured approach of the psychological intervention being studied. Patients might, for example, engage in a neutral discussion with the clinician but not engage in the more directed conversation involved in, say, cognitive-behavioural therapy.

#### Usual care

The patient receives standard care for their condition as typical for their diagnosis, region and time.

#### Waiting-list

The patient receives no treatment during the trial period but receives the active intervention following the trial period.

#### No-treatment

The patient receives no treatment and, in contrast to a waiting-list control, is not told they will receive an active intervention following the trial period.

The 83 trials included patients with a total of 15 different mental health diagnoses. Any included patient in the trials reviewed must have received a formal diagnosis drawn from either the American Psychiatric Association's DSM (DSM-I to DSM-5, plus DSM-III-R and DSM-IV-TR) or the World Health Organization's ICD (revisions 6 to 11).

## Results

Primary outcomes measured the efficacy of placebo, waiting-list and usual care interventions compared with waiting-list and no-treatment interventions for all diagnoses; and the compared incidence of serious adverse events in these interventions.

Secondary outcomes grouped trials investigating specific mental health diagnoses and again compared efficacy and incidence of adverse events for the various interventions.

Trials that collected continuous data were analysed via standardised mean differences (s.m.d.) and those that collected dichotomous data were reported as risk ratios (RR) (Box 3).

The review authors found a significant difference in favour of all control methodologies when compared with waiting-list or no-treatment (s.m.d. = -0.37, 95% CI -0.49 to -0.25) in trials that generated continuous data. There were also marked differences in favour of the effects of specific modalities of control: psychological placebo versus waiting-list/no-treatment (s.m.d. = -0.49, 95% CI -0.64 to -0.30); physical placebo versus waiting-list/no-treatment (s.m.d. = -0.21, 95% CI -0.35 to -0.08); and pharmacological placebo versus waiting-list/no-treatment (s.m.d. = -0.14, 95% CI -0.39 to 0.11).

No difference was seen in all other meta-analyses of trials that generated continuous data nor in any trials that generated dichotomous data.

Usual care was not found to be significantly beneficial when compared with waiting-list or no-treatment in the seven trials that reported usable data.

The single trial that compared waiting-list and no-treatment directly (Howlin 2007) was not able to yield usable data.

No difference was found in the incidence of adverse events in any of the comparisons undertaken.

## Conclusions

Although the review authors found significant differences in outcome between the control methodologies and waiting-list or no-treatment, these results were tempered by the quality of evidence included, which was rated 'low quality' to 'very low quality' on the GRADE Working Group grades of evidence (Atkins 2004).

## Appraisal of the study

### *The limitations of aggregating different diagnoses*

This review aggregated studies that examined 15 different diagnoses, including anxiety and depression, sleep-wake disorders, autism spectrum disorders and schizophrenia. Attempts by the review authors to analyse the effects of the potential heterogeneity introduced by these varying patient populations were hampered by limited data, allowing them to run analyses of heterogeneity for only 7 of the 15 disorders included. The effect of including such a broad range of pathologies therefore goes partially unexamined. Where analysis was possible for specific disorders, it showed significant and consistent benefit in control groups compared with waiting-list or no-treatment groups but these effects varied depending on the specific disorder studied.

Compare this to an example drawn from Hróbjartsson & Gøtzsche's (2010) meta-analysis of

### BOX 3 Continuous versus dichotomous data

Trials measure either a continuous set of data, for example using a numerical rating to report symptom improvement, or a dichotomy of outcomes, such as recording 'yes' or 'no' when rating whether symptoms have improved. Within the context of meta-analyses, these differing types of result require different outcome measures, outlined below.

#### Standardised mean difference (s.m.d.)

In trials that measure continuous data, results are reported as a numerical change in the mean result. Where the method of rating varies between trials, the units of rating and the range of rating responses may also vary between trials. The mean difference can therefore be divided by the standard deviation to generate the standardised mean difference, allowing

aggregation of these results despite their different methodologies.

#### Risk ratio/relative risk (RR)

In trials that measure dichotomous data, results are reported as a risk ratio (also referred to as relative risk). Relative risk compares the risk of an outcome (e.g. a disease state) between different groups (e.g. a group that has been exposed to an intervention and a group that has not). It is calculated simply by dividing the incidence of outcome in the intervention group with the incidence of outcome in the control group. Relative risk includes this aspect of comparison between groups, whereas 'absolute risk', which is also sometimes reported, simply measures the incidence of an outcome occurring in a given sample.

placebos in all clinical conditions. Hróbjartsson & Gøtzsche highlighted five trials that specifically looked at the effect of acupuncture on various types of pain (Linde 2005; Melchart 2005; Witt 2005; Brinkhaus 2006; Scharf 2006). In these trials, the effect was far more significant than in the broader data-set and the quality of the evidence was consistently higher, with medium to large sample sizes, consistent methodologies focused on a single intervention type (acupuncture), and low drop-out rates. These trials illustrate the firmer conclusions that might be drawn from smaller-scale reviews of trials with more common features and populations. The broad applicability of a review of trials involving all psychiatric diagnoses is counter-balanced by the weaker conclusions that can be drawn from such a heterogeneous sample of trial methodologies, populations and analysis.

### *Risk of bias*

The poor quality of much of the included trials' data is acknowledged. In particular, participant masking ('blinding') was not possible in any of the included trials as participants were certain to know whether they received waiting-list or no-treatment as opposed to an active intervention. The trials also suffered broadly from small sample sizes and variability in randomisation, methodology and populations studied.

### *Inconsistency in the application of control states*

Inconsistency in the design of control conditions was also found within the trials analysed. Ten trials described their control condition as waiting-list but were reclassified by the review authors as no-treatment as they did not meet the pre-selected criteria of a waiting-list control.

Further inconsistency is introduced in the variability of usual care as a control condition, the administration of which can vary significantly from trial to trial (Mohr 2014).

Taken together, these factors limit the applicability of the review authors' conclusions in differentiating intervention effect from effect due to variation in choice and administration of control condition.

### *Waiting-list versus no-treatment*

Importantly, the comparison of waiting-list with no-treatment was undertaken in only one trial, which involved primary school children with autism spectrum disorders (Howlin 2007) and was not able to provide usable data.

This lack of outcome analysis between waiting-list and no-treatment comparators makes it difficult to know whether aggregating these controls, as in this

review, may introduce variable effects on outcome even between these two control conditions.

## **Conclusions**

As the comparator state against which a new intervention will be measured, a control should provide a stable point of reference from which firm comparisons might be made and conclusions drawn. Moreover, a standard methodology for the selection of a control state in a given field allows more confidence in both meta-analysis of the literature and also inter-experimental comparison when making clinical decisions or developing guidelines (Gold 2017). The current lack of uniformity in the use of control states in psychiatric research restricts the confidence with which this research can be read and applied (Hróbjartsson 2010). Perhaps more concerning is the potential for the variability introduced by inconsistent control states to go unanalysed, leading to the over- or underestimation of the effect of interventions, depending on the control used.

In this Cochrane Review, Faltinsen et al (2022) set out to test the equivalence of these different experimental controls. In this, they demonstrate a significant difference in outcome depending on the selection of control in the trials reviewed. Although the results are significant, the review authors acknowledge the overall lack of trials comparing control states and note that the lack of uniformity in their design reveals a paucity of research in an area that is central to psychiatric research methodology. Finally, this review encourages the closer scrutiny of the selection and design of control states and aims to support the formulation of guidelines on the application of controls in future psychiatric research.

## **Funding**

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

## **Declaration of interest**

None.

## **References**

- Atkins D, Best D, Briss PA, et al (2004) Grading quality of evidence and strength of recommendations. *BMJ*, **328**: 1490.
- Bellg A, Borrelli B, Resnick B, et al (2004) Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, **23**: 443–51.
- Bothwell LE, Podolsky SH (2016) The emergence of the randomized, controlled trial. *New England Journal of Medicine*, **375**: 501–4.

- Brinkhaus B, Witt CM, Jena S, et al (2006) Acupuncture in patients with chronic low back pain: a randomized controlled trial. *Archives of Internal Medicine*, **166**: 450–7.
- Faltinsen E, Todorovac A, Staxen Bruun L, et al (2022) Control interventions in randomised trials among people with mental health disorders. *Cochrane Database of Systematic Reviews*, **4**: MR000050 (<https://doi.org/10.1002/14651858.MR000050.pub2>).
- Gold SM, Enck P, Hasselmann H, et al (2017) Control conditions for randomised trials of behavioural interventions in psychiatry: a decision framework. *Lancet Psychiatry*, **4**: 725–32.
- Higgins J, Thomas J, Chandler J, et al (2019) *Cochrane Handbook for Systematic Reviews of Interventions Version 6.0 (Updated July 2019)*. Cochrane. Available from [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook).
- Howlin P, Gordon RK, Pasco G, et al (2007) The effectiveness of picture exchange communication system (PECS) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of Child Psychology and Psychiatry*, **48**: 473–81.
- Hróbjartsson A, Gotzsche P (2010) Placebo interventions for all clinical conditions. *Cochrane Database of Systematic Reviews*, **1**: CD003974.
- Linde K, Streng A, Jürgens S, et al (2005) Acupuncture for patients with migraine: a randomized controlled trial. *JAMA*, **293**: 2118–25.
- Melchart D, Streng A, Hoppe A, et al (2005) Acupuncture in patients with tension-type headache: randomised controlled trial. *BMJ*, **331**: 376–82.
- Mohr DC, Spring B, Freedland KE, et al (2009) The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, **78**: 275–84.
- Mohr DC, Ho J, Hart TL, et al (2014) Control condition design and implementation features in controlled trials: a meta-analysis of trials evaluating psychotherapy for depression. *Translational Behavioral Medicine*, **4**: 407–23.
- Patterson B, Boyle MH, Kivlenieks M, et al (2016) The use of waitlists as control conditions in anxiety disorders research. *Journal of Psychiatric Research*, **83**: 112–20.
- Scharf HP, Mansmann U, Streitberger K, et al (2006) Acupuncture and knee osteoarthritis: a three-armed randomized trial. *Annals of Internal Medicine*, **145**: 12–20.
- Schulz K, Chalmers I, Hayes RJ, et al (1995) Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, **273**: 408–12.
- Sibbald B, Roland M (1998) Understanding controlled trials: why are randomised controlled trials important? *BMJ*, **316**: 201.
- Witt C, Brinkhaus B, Jena S, et al (2005) Acupuncture in patients with osteoarthritis of the knee: a randomised trial. *Lancet*, **366**: 136–43.