REVIEW
Blinding in trials of interventional procedures is possible and worthwhile [version 1; peer review: 3 approved]

Karolina Wartolowska, David Beard, Andrew Carr
Botnar Research Centre, Nuffield Department of Orthopaedics Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford, UK

Abstract
In this paper, we have used evidence from our earlier review of surgical randomised controlled trials with a placebo arm to show that blinding in trials of interventional procedures is feasible, and that many creative methods can be used to make the active and the placebo procedure as similar as possible. We give examples of ingenious strategies used to simulate the active procedure and make the placebo control indistinguishable from the active treatment. We discuss why it is important to blind of patients, assessors, and caregivers and the types of bias that may occur in interventional trials. Finally, we describe the benefits of blinding, from the obvious ones such as avoiding bias, as well as less evident benefits such as avoiding patient drop out in the control arm.

Keywords
blinding, RCT, randomised controlled trials, placebo, surgery, procedures, non-pharmacological
Introduction
The aim of a trial is to produce unbiased evidence. As randomised controlled trials (RCTs) with a placebo arm control for many types of bias, they are regarded as the most reliable method of demonstrating treatment efficacy and provide the highest level of evidence. RCTs of interventional procedures are rare, partly because they are challenging; however, they are not impossible to perform, even if they involve a placebo arm. In this paper we will discuss why trials should be blinded and summarise the methods which have been used in the published placebo-controlled trials of interventional procedures to achieve blinding.

Blinding in interventional trials is often necessary because nowadays many procedures are performed to reduce pain and improve function and quality of life. Pain, function and quality of life are sometimes regarded as preferable outcome measures because they reflect patients' needs and point of view. However, as these outcomes depend on patients' subjective perception, they are prone to bias and may lead to an exaggerated treatment effect in open-label trials. Using subjective outcomes in an open-label study undermines its internal validity because it is not possible to determine how much of the reported effect is related to the investigated treatment and how much is related to bias.

Blinding of patients, surgeons, outcome assessors and caregivers

Blinding of patients
Blinding means concealing the treatment allocation from patients and any other people involved in the trial who may bias the results of the trial by knowing which groups the patients were randomised to.

Blinding of patients prevents reporting bias in patient-reported measures. For example, it has been demonstrated that non-blinded patients exaggerate the effects size by 0.56 standard deviations and that the effect is even larger in trials on interventional procedures, such as acupuncture. This bias may be caused by patients' expectations of treatment effect and information given to them before the treatment. Patients may also report symptoms depending on their “hunches” about treatment being effective or they may give answers they believe are “correct” or expected from them, for example, because it would have been impolite not to report improvement. Therefore, it has been suggested that patients should be blinded whenever possible.

Blinding of patients also reduces adherence bias, i.e. patients in the control group not following the protocol/treatment. It may also prevent so called “contamination of the control group”, i.e. seeking additional treatment/help outside the trial and receiving concomitant treatment. Blinding improves patient retention in the trial. Risk of attrition in blinded trials is about 4% whereas in non-blinded trials it is 7%. Specifically in placebo-controlled surgical trials, subject retention is often reported as “excellent”, and in our analysis the withdrawal rate was low (4%) and comparable between the treatment and the placebo arm.

Blinding of surgeons
Unlike drug trials, in which the physician gives a tablet prepared somewhere else, the surgeon has to perform a specific procedure considered to be therapeutic; therefore, blinding of surgeons may not always be possible.

There have been attempts to blind surgeons, for example a surgeon inserted a catheter under fluoroscopic guidance and handed over the procedure to a technician who delivered the radiofrequency energy (or not) according to the allocation. In other trials, a palatal implant delivery system was prepared by the manufacturer to either contain the implant or not, which allowed for blinding of surgeons.

Blinding of outcome assessors
In 81% of placebo-controlled surgical trials both patients and assessors were blinded. It has been demonstrated that, non-blinded assessors of subjective outcomes cause less bias in trials than non-blinded patients reporting their symptoms.

Blinding of assessors prevents observer-related bias, detection bias, and the Pygmalion effect. The Pygmalion effect refers to a situation when investigators looking for a particular response are predisposed to interpret the result in a way that shows the response they expect even if it is objectively absent. For example, a study by Hrobjartsson and colleagues demonstrated that non-blinded assessors were over-optimistic and “over-rated” patients in the treatment group rather than “under-rated” patients in the control group. In some trials the assessment was done by people not involved in the surgery, for example blinded researchers, staff at another hospital that they were operated on at, or by pathologist blinded to the treatment allocation.

Blinding of caregivers
Apart from blinding patients and assessors, it is important that caregivers and clinical/research staff do not know patient treatment allocation, because their behaviour and attitudes may influence patient responses. Patient-clinician interaction plays an important role in treatment response, and patients in trials do better as they get more attention and time from clinical staff than patients receiving standard care. Therefore, the interactions between patients and the trial team should be standardised so that the “treatment context” (similar attention from doctors, expectations, and settings) are comparable between the groups.

Strategies used to maintain blinding in interventional placebo-controlled trials
A placebo control is necessary if we want to know whether improvement is really caused by the investigated procedure. It compares the intervention of interest with a procedure that seems identical but does not involve the crucial element of believed to be “the cure”. The aim of a placebo arm is to control for effects of receiving treatment not specifically related to the investigated intervention.

It is often difficult to determine what is a specific and what is a non-specific effect in a trial, and to disentangle placebo
response from response bias or the effect of patient-doctor interactions. It is beyond the scope of this review to discuss definitions of placebo. Whether something is or is not a placebo depends on the intervention and chosen outcome variables, but in order for blinding to be successful, the control procedure has to be as similar as possible to the investigated procedure. Interventional trials differ from drug trials as they require access to the anatomical structure of interest; therefore, they involve a skin incision or an insertion of a scope.

Many published surgical trials used general anaesthesia or heavy sedation, which made blinding easier because patients were unaware of the details of surgical procedures. In such trials, only the surgical wound had to be similar in both groups. Some studies did not add any placebo procedure but simply omitted part of interventional procedure, for example, in the trial by Stone and colleagues, all patients underwent a percutaneous coronary intervention and maximal medical therapy but only patients in the active arm also had percutaneous transmyocardial revascularisation.

When light sedation or local anaesthesia are used, surgical staff have to simulate the actual intervention to preserve the blinding. The complexity of a surgical procedure can make blinding challenging, and ingenious ideas are required to make the real and placebo interventions indistinguishable.

**Imitation of incision/ surgical access point**

If a procedure requires open surgery, then it will leave an obvious mark where the incision has been made, which will have to be imitated in the placebo group. There have been very few trials involving full skin incision, in both the surgical and placebo arms. In the seminal trials on internal mammary artery ligation, a skin incision was made to expose the arteries in all patients but no ligation was made in the placebo group. Similarly, Guyuron and colleagues used a skin incision to expose superficial nerves and muscles, which were cut during the active surgery but, in the placebo group, the integrity of these structures was maintained. Trials investigating transplantation of dopaminergic neurones as a treatment for Parkinson’s disease not only required skin incision but also burr holes in the skull.

However, most of the published placebo-controlled surgical trials used minimally-invasive methods to access the structure of interest. For example, the placebo procedure involved laparoscopy but without ablation, endoscopy without radiofrequency energy delivery, bronchoscopy without radiofrequency energy delivery, or bronchoscopy without valve placement. Therefore, most of the studies required a small incision to mimic the portals created during the laparoscopy or arthroscopy, or to mimic the incision through which an intravascular catheter was inserted. Interestingly, Sutton and colleagues used three incisions in both groups so that patients could not tell apart a diagnostic laparoscopy from a laparoscopic surgery; even though the third instrument port was not necessary in the placebo group. Trials using endoscopy and bronchoscopy were even easier to blind as natural orifices were used to insert the scope, and the incision or actual procedure site was not visible to patients, caregivers, and assessors.

**Simulation of an interventional procedure**

Typically, the preparation for the placebo procedure and the active procedure was as similar as possible and imitated the visual, auditory, and physical cues. In order to mimic the sounds, surgeons were required to talk through the procedure steps, ask for instruments, use suction, or ask for a laser or other device to be activated, even though it was not used in the placebo group.

Clinical staff performing the intervention were screened from the patients’ view either by a surgical drape or by arranging the operating room in a way that the patient could not see the procedure. In the trial by Stone and colleagues, patients were heavily sedated and wore opaque goggles. In a trial by Maurer and colleagues, the manufacturer delivered tools that looked identical but did not contain an implant, which allowed for blinding of patients and clinical staff.

Surgeons also attempted to imitate sensory cues, for example, by manipulating the knee as if the actual arthroscopy were performed, injecting saline to imitate tidal irrigation, or by splashing saline on the knee to simulate lavage. In a trial on meniscectomy, surgeons used a mechanised shaver (without the blade) pushing it firmly against the patella to simulate the sensations the patient would experience during the surgery. In a trial on intragastric balloon for obesity, operators manipulated the endoscope as during the balloon insertion to create the sensation of resistance in the stomach.

Even smell during the surgery was imitated to make the placebo procedure indistinguishable from surgery. For example, in the trial by Deviere and colleagues there were concerns that patients could have known the allocation because the copolymer used in the active arm had a distinct smell. In trials on vertebroplasty, a container with cement was opened during placebo procedure to help with blinding.

**Inactive nature of the placebo**

It is important that the procedure used for blinding does not have any therapeutic effect. For example, the results of the vertebroplasty trials were criticised because the elements of placebo procedure could have had an effect on the reported pain, namely, a potential pharmacological anaesthesia due to injection of an anaesthetic into the facet capsule and periosteum.

On the other hand, the procedure used for blinding may have diagnostic use, as with diagnostic laparoscopy or diagnostic laparoscopy with biopsy. In the trial by Sihvonen and colleagues, all participants underwent diagnostic arthroscopy, but only after they had been confirmed to be eligible for inclusion in the trial was the envelope with the assignment opened and the assignment revealed to the surgeon.
Duration of the procedure
Many trials specifically stated that the duration of procedure in the surgical and control arms were matched, either by imitating the elements of the surgical procedure or by keeping all patients in the operation room for the same duration of time. However, in some trials, the placebo procedure was shortened in comparison to the actual surgery because it was believed it would have been ethically unacceptable to prolong the placebo intervention.

Additional procedures that may reveal allocation
Interventional treatment often requires additional procedures, such as diagnostic scans or medication to prevent infection, blood clots, transplant rejection, or epileptic fits. For example, a trial by Freed and colleagues, both groups received identical preoperative evaluation, intraoperative sedation and pain control, underwent two PET scans and a MRI scan, and received phenytoin. In some trials, the same medication was given in both groups, whereas in others unnecessary treatment was omitted or imitated, for example by injecting saline instead of antibiotics.

Standardisation of interventions and care
The active and placebo procedure have to be indistinguishable but they also have to be stable and standardised. Standardisation of the procedure itself may be difficult but is important because surgeons vary in their experience, and gain experience throughout the trial.

Some of the changes observed in a trial may not be related to the treatment or the placebo intervention, but may be caused by the natural course of the disease, spontaneous remissions or fluctuations in the severity of symptoms or regression to the mean. Some changes may be a result of just being in a trial either because of lifestyle changes that are part of the protocol such as self-monitoring using diaries or avoiding alcohol or due to so called “Hawthorne effect”, which refers to change in the behaviour when people, both patients and doctors, know that they are being observed. Finally, it has been demonstrated that adhering to a protocol improves the performance of doctors, and that patients who adhere to treatment regimes have better outcomes. Therefore, it is important to standardise pre- and post-operative care, and the explanations given to the patients. For example, in a trial by Sihvonen and colleagues all procedures were standardised and recorded on video; the post-operative care was also standardised, and all patients received the same exercise programme and walking aids.

Blinding after the surgery
Most trials blinded the assessors while the surgeon and other staff in the operating room were aware of the group assignment, and did not participate in further treatment, post-operative care or follow-up of the patient. In a trial by Thomsen and colleagues, the post-operative care and assessment was done at a different hospital than the surgery. In a trial by Cotton and colleagues a post-operative care was provided by the referring physician, who was blinded when deciding on treatment, and when this was not sufficient, by the evaluating physician at the study site (who was also blinded).

Bias specific to surgical trials
There are other types of bias that are specific to surgical trials. For example, in trials on upper gastrointestinal tract bleeding, the endoscopic procedure was not performed if the rate of blood loss was too fast, or the endoscopy was judged to be life-threatening and posed an unacceptable risk. In other trials, some patients were excluded because they could not tolerate endoscopy, or due to anatomical conditions that made the surgery impossible to carry out (laser could not be aimed at the bleeding arteries). Alternatively, some patients were not included in a trial because they were no longer eligible, for example because the bleeding had stopped or they no longer reported the symptoms on the day of the study.

Conclusions
Blinding in trials of interventional procedures is possible and many creative methods have been used to maintain the blinding. Interventional procedures are challenging to blind, but the effort is worthwhile because of the obvious benefits, such as avoiding bias, as well as the less evident benefits, such as avoiding patient drop-out in the control arm.

Competing interests
No competing interests were disclosed.

Grant information
This work has received funding from the NIHR Oxford Musculoskeletal Biomedical Research Unit.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References
5. Cook JA: The challenges faced in the design, conduct and analysis of surgical
Open Peer Review

Current Peer Review Status: ✔ ✔ ✔

Version 1

Reviewer Report 04 December 2017

https://doi.org/10.5256/f1000research.13566.r25901

© 2017 Hróbjartsson A et al. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Helene Moustgaard
Centre for Evidence-Based Medicine, Odense University Hospital (OUH), University of Southern Denmark, Odense, Denmark

Asbjørn Hróbjartsson
Centre for Evidence-Based Medicine, Odense University Hospital (OUH), University of Southern Denmark, Odense, Denmark

The overall topic of this paper is important: Blinding facilitates unbiased trial results, but blinding is less used in surgical trials (and less frequently discussed) than in trials of pharmacological interventions. A prevalent preconception is that blinding is not possible, or not relevant, in surgical trials. The main strength of the paper is that it illustrates well this to be ill founded.

However, we wondered whether the manuscript could be further improved. First, we were in doubt what the specific aim of the paper was. We assumed that the authors aimed to discuss the methodological role of binding in surgical trials (and not primarily to present and reflect on their previous review of surgical placebo trials)? If so, it would have strengthened the paper if it had provided some reflections on the barriers to the use of binding in surgical trials, the ethical issues involved (and how we can best balance the need for reliable evidence for effects of surgery against the risks or discomforts patients may be exposed to in blinded surgical trials), as wells as discuss what measures could be taken to remedy the current situation where binding is used infrequently in surgical trials. In general, the reader would be better prepared if the paper clearly stated whether the topic it addresses is binding in surgical trials or use of placebo in surgical trials and whether the main focus is providing practical examples or providing a general overview of the topic.

It would be helpful if the paper defined central terms used. For example, what is meant by “interventional procedures”? Does that include acupuncture, radiotherapy, and physiotherapy? Also, how do the authors conceptualize “subjective outcomes vs. objective outcomes” [1].

The central section Bias specific to surgical trials appears compressed. It would help the reader if the section could more clearly show how the example provided introduces bias, and in what way this bias is specific too surgery (rather than just more frequently encountered).
Of minor importance was that we were puzzled by some of the references used. For instance, in the Introduction, there is a reference to a study by Wood et al [2], but several more recent studies have looked at this issue [3]. Also, in the section Blinding of outcome assessors the authors refer to a study by one of the undersigned, to the effect that the study documents a “Pygmalion effect” in un-blinded studies [4]. The term, however, was not used in the original publication. It would possibly help the reader if the authors referred to publications that provide a general overview of blinding in trials (and not only surgical trials).

References

Is the topic of the review discussed comprehensively in the context of the current literature?
Partly

Are all factual statements correct and adequately supported by citations?
Yes

Is the review written in accessible language?
Yes

Are the conclusions drawn appropriate in the context of the current research literature?
Yes

**Competing Interests**: No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

**Author Response 14 Jan 2018**

**Karolina Wartolowska**, Nuffield Department of Orthopaedics Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford, UK

Our aim was to discuss the methodological role of blinding in surgical trials and to demonstrate that blinding can and should be done in surgical trials, especially, if the outcomes are subjective. The need to control for bias in surgical trials is often not recognised. We agree with the Reviewer that there is a preconception/assumption that blinding in surgical trials is not possible or not necessary. Surgery may mean an open surgery undertaken to preserve life. In such cases the outcome is
objective, binary and with low risk of bias. However, nowadays, many invasive procedures are performed not to save lives but to improve the quality of life, function or to reduce pain. As a consequence, their outcomes are highly subjective, and as such, prone to bias. This problem seems to be unappreciated in clinical and research community. In this paper, we wanted to demonstrate that blinding can be achieved in surgical trials and that it is necessary if the outcomes are subjective.

Placebo-controlled surgical trials represent a single type of trial that requires blinding, but similar arguments apply to trials comparing two different surgical methods.

In this paper, we did not discuss the placebo because our main focus was blinding as a way to control bias. We used placebo-controlled surgical trials as an example but similar arguments apply to trials comparing two different surgical methods. We have already written about barriers in completing placebo-controlled surgical trials (Wartolowska et al., 2016), about the balance between harms and benefits of such trials (Wartolowska et al., 2014) and about the ethical implications of placebo in surgery (Savulescu, Wartolowska and Carr, 2016)(George et al., 2016).

We have used acupuncture as an example of an experimental study on placebo effect because acupuncture is the only type of invasive procedure used in such studies.

We have included radiofrequency procedures because they change anatomy. In this paper, we have defined surgery as “any interventional procedure that changes the anatomy and requires a skin incision or the use of endoscopic techniques; dental studies were excluded. We used the term placebo to refer to a surgical placebo, a sham surgery, or an imitation procedure intended to mimic the active intervention; including the scenario when a scope was inserted and nothing was done but patients were sedated or under general anaesthesia and could not distinguish whether or not they underwent the actual surgery”.(Wartolowska et al., 2014) We have added these definitions to the manuscript.

“Outcomes were classified as “subjective”, i.e., patient-reported and depending on the patients’ perception and cooperation, “assessed”, i.e., subjective ratings by external assessors and depending upon their judgment, and “objective”, i.e., measured using devices or laboratory tests and independent of patients’ or observers’ perception, for example, weight.”(Wartolowska et al., 2016)

As suggested by the Reviewers, we have explained in the section titled “Bias specific to surgical trials” that by including only patients with non-life-threatening symptoms or with sufficiently severe symptoms on the day of surgery a bias is created, which is difficult to control for.

As suggested by the Reviewers, we have added the paper by Page et al. as a reference in this paragraph. We have also added a reference that provides a general overview of blinding in trials (Hróbjartsson and Boutron, 2011) and provided references to Rosenthal’s papers to explain the “Pygmalion effect”.(Rosenthal and Jacobson, 1968)

REFERENCES

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 04 December 2017

https://doi.org/10.5256/f1000research.13566.r28006

© 2017 Walach H. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Harald Walach**
Institute for Transcultural Health Studies, European University Viadrina, Frankfurt, Germany

This review is, as its self-set goal goes, very well substantiated, balanced and timely. There is no factual correction in itself that is necessary in my view.

However, I wish to give a few critical feedbacks to the authors that they may wish to integrate. I would prefer if they did, but in my view it is unfair to formally demand such an extension. However, the debate itself would profit.

My point is regarding the presuppositions of the current model the authors implicitly subscribe to, for instance when they say randomised trials “provide the highest level of evidence” (1st para). This is true, but only from a - debated – point of departure that accepts the hierarchisation of evidence as a given and true proposal. As it happens, we, and others, have debated this [1,2]. The hierarchisation of evidence accepts – without further reflection – that internal and external validity are compatible, linear concepts, where first internal evidence can be gleaned and is prior and superior to external validity. This I find an unsubstantiated point of view that needs at least acknowledgement. All the arguments in this paper address the improvement of internal validity, and as such they are very good.

The argument that “specific” effects have to be tested against “non-specific” or placebo-effects and therefore blinding is necessary also makes presuppositions that are not clarified and critically debated. The presupposition is that there is such a thing as separability of these effects. This separability argument is highly speculative, very little researched and where and when researched found to be wrong [3, 4]. One can, of course, try and disentangle various components of interventions. And especially where interventions are costly or fraught with side-effects this is even an ethical requirement, because else
patients would be submitted to problematic procedures with little benefit or where other interventions might be just as good, all taken together. But one should be clear about the fact that this is not something that automatically gets rid of all problems in evaluating therapeutic procedures.

Therefore, I think the arguments the authors present are valid within the context of improving internal validity of intervention trials that offer costly and potentially dangerous or problematic interventions. But I think the authors should be clear themselves that this is not a universally valid strategy and it would be wise to integrate some caveats and restrictions in their argument.

References
Publisher Full Text

Is the topic of the review discussed comprehensively in the context of the current literature?
Yes

Are all factual statements correct and adequately supported by citations?
Yes

Is the review written in accessible language?
Yes

Are the conclusions drawn appropriate in the context of the current research literature?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research methodology, clinical and other trials

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 14 Jan 2018

Karolina Wartolowska, Nuffield Department of Orthopaedics Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford, UK

We are grateful for the comment regarding the hierarchy of evidence. We have rephrased the sentence so that it now reads “high level of evidence because they attempt to minimise bias due to confounding factors”. We have also added the suggested reference (Walach and Loef, 2015).

We agree that RCTs have high internal validity because they are performed in a standardised way.
and because they attempt to control for confounding factors. However, owing to the controlled conditions under which they are performed, RCTs often have low external validity. In this paper, we wanted to draw attention to the fact that surgical trials often use subjective outcomes, and are therefore prone to bias, which has to be minimised or controlled for.

We agree with the Reviewer, that the “specific” and “non-specific” effects are difficult to separate or to unequivocally define. As surgical procedures are complex and consist of many elements, it is sometimes difficult to identify the crucial surgical element that is believed to be therapeutic. We have described the CSAW trial as an example. The aim of this trial was to investigate whether the improvement was due to change of anatomy (arthroscopic removal of a bony spur on the acromion) or it was not related to the change in anatomy and the spur removal was not necessary because a placebo procedure (arthroscopy only) results in similar change in pain and function.

We wanted to draw attention to the fact that in the case of surgical trials, with their high costs and risks, one would expect to see a much larger effect after a surgical procedure than after a placebo or no treatment.

We are familiar with the debate in the literature regarding specific and non-specific effects. In the context of surgical trials, by “specific effects” we meant changes directly related to change in anatomy, for example, a bony spur removal in the trial of arthroscopic shoulder decompression surgery for subacromial pain. All other possible factors resulting in improvement, such as the effect of arthroscopy, the placebo effect, influence of patient-doctor interactions, symptoms fluctuation as well as spontaneous improvement were regarded as “non-specific”. We also agree with Enck and colleagues that the effect in the placebo and active arms may not be additive; therefore, the improvement related to the crucial surgical element may not equal the difference between the effect in the active and placebo arms. (Enck et al., 2011) We would like to clarify that blinding and placebo control help to minimise bias and aid in the interpretation of trial results; especially if the outcomes are subjective and the symptoms fluctuate or may spontaneously improve.

REFERENCES

**Competing Interests:** No competing interests were disclosed.
This manuscript is a narrative overview of methods of blinding in trials of procedures, using the previous review by these authors as the source of data. It explains the rationale for blinding and gives examples (subgrouped) of how blinding was achieved in previous trials. The authors wisely avoid a detailed discussion of the placebo effect and concentrate on the need for blinding based on empirical evidence of bias resulting from not blinding.

The paper reinforces the need for blinding in procedural trials which is a reasonable demand that is increasingly being heeded. In other words, the message is important.

The paper is very well written. I only have minor comments, below:

1. Page 2, second last paragraph, "crucial element of believed to be the cure" should read "crucial element believed to be the cure"

2. The evidence of bias resulting from not blinding may be one sided. Opposing evidence should be discussed, such as Berkman ND, The Empirical Evidence of Bias in Trials Measuring Treatment Differences which shows that the effect of blinding may even be in the opposite direction and the classic Schultz article from 1995 which gives a lower estimate of effect exaggeration from unblinding. There are several other summaries and meta-epidemiological studies in this area.

References

Is the topic of the review discussed comprehensively in the context of the current literature?
Yes

Are all factual statements correct and adequately supported by citations?
Yes

Is the review written in accessible language?
Yes

Are the conclusions drawn appropriate in the context of the current research literature?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Orthopaedics, Epidemiology

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
Karolina Wartolowska, Nuffield Department of Orthopaedics Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford, UK

We are grateful for the comment and we have added an explanation that trials have limited external validity, as by controlling for confounding factors and standardising the trial conditions, they become very different from usual clinical practice. They are also liable to have study populations that are so unique that the results of the trial cannot be easily generalised.

In the case of blinding, the magnitude of effect may be lower in a blinded trial than in everyday practice, because blinding causes uncertainty about treatment allocation. In clinical practice, there is usually no doubt as the patient-doctor relationship is based on trust and on the assumption that doctors offer patients the best available treatment. In an RCT, there is an inherent uncertainty as to which treatment group a patient has been allocated. (Enck et al., 2011)

REFERENCES

Competing Interests: No competing interests were disclosed.