Placebo effect studies are susceptible to response bias and to other types of biases

Asbjørn Hróbjartsson[1], **Ted J Kaptchuk**[2], and **Franklin G Miller**[3]

Asbjørn Hróbjartsson: ah@cochrane.dk

[1] The Nordic Cochrane Centre, Rigshospitalet. Copenhagen, Denmark

[2] Beth Israel Deaconess Medical Center, Harvard Medical School. Boston, MA. USA

[3] Department of Bioethics, National Institutes of Health. Bethesda, MD. USA

Abstract

OBJECTIVE—Investigations of the effect of placebo are often challenging to conduct and interpret. The history of placebo shows that assessment of its clinical significance has a real potential to be biased. We analyse and discuss typical types of bias in studies on placebo.

STUDY DESIGN AND SETTING—a methodological analysis and discussion.

RESULTS—The inherent nonblinded comparison between placebo and no-treatment is the best research design we have in estimating effects of placebo, both in a clinical and in an experimental setting, but the difference between placebo and no-treatment remains an approximate and fairly crude reflection of the true effect of placebo interventions. A main problem is response bias in trials with outcomes that are based on patients reports. Other biases involve differential co-intervention and patient drop-outs, publication bias, and outcome reporting bias. Furthermore, extrapolation of results to a clinical settings are challenging because of lack of clear identification of the causal factors in many clinical trials, and the non-clinical setting and short duration of most laboratory experiments.

CONCLUSIONS—Creative experimental efforts are needed to assess rigorously the clinical significance of placebo interventions and investigate the component elements that may contribute to therapeutic benefit.

Keywords

Placebo; placebo effect; response bias; bias; randomised trial; experiment

Introduction

Scientific interest in the placebo effect has grown dramatically in the past two decades. Nevertheless, assessing the clinical relevance of this fascinating phenomenon is hampered by the difficulty in developing unbiased assessment of the effects of placebo interventions. In this article, we review the pervasive and complex connection between the placebo effect and bias.

After World War II, placebos were widely adopted as concurrent controls in randomized clinical trials to facilitate recruitment and retention and to control for bias (1). Interest in the "placebo effect" emerged out of the observation that patients in the placebo arm of these clinical trials often demonstrated significant improvement (2).

Henry Beecher popularized the concept of the placebo effect and brought it to the attention of the medical community in his classic 1955 *JAMA* article, "The Powerful Placebo" (3). Presenting a review of assorted placebo-controlled trials, he argued that the substantial improvement in the condition of patients receiving placebo was caused by the placebo intervention. His position implied two strong, but methodologically weak, claims: first, that placebo interventions caused large clinical effects on many patients across many clinical conditions, improving both patient-reported and observer-reported outcomes; second, that this assessment was reliable because it was based on randomized trials. Beecher saw placebo effects as a major source of bias in the assessment of treatment efficacy and his chief purpose in publicizing such a potential threat was to advocate for placebo-controlled trials (2). The concept of 'powerful placebo' became an established dogma in biomedicine.

Nevertheless, Beecher's analysis committed the very fallacy that underlies the need for controlled trials. The observed response to placebo in randomized trials does not itself provide any reliable, unbiased, evidence of a placebo *effect*—an outcome caused by receiving a sham treatment disguised to be indistinguishable from an active (verum) medical intervention. Unbiased assessment of the placebo effect requires comparison of placebo interventions with a suitable control group in order to distinguish an effect of the placebo intervention from confounding factors, for example the natural history of the condition under investigation or regression to the mean (4). The flaws in Beecher's approach were clearly recognized in the late 1990's (5), but by that time the notion of 'powerful placebo' became deeply rooted.

Enhancing the scientific credibility of the placebo effect was the finding in 1978 that placebo analgesia could be blocked by the opioid antagonist naloxone (6), indicating that placebo analgesia can involve endogenous opiates (7,8). A further development in thinking about placebo was the application of brain imaging techniques from 2001 onwards sparking growing experimental investigation into neurobiological mechanisms of placebo effects, especially in pain (9).

Also in 2001, in sharp contrast, the power of the placebo was challenged by a systematic review published in *The New England Journal of Medicine*. The review identified 114 randomized clinical trials including placebo and no treatment groups, and reported no evidence of overall effects of placebo for objective and binary outcomes and a small, and doubtfully clinically relevant, effect for continuous subjective outcomes, such as pain (10). The findings were clearly incompatible with Beecher's classic position. Some media commentators interpreted the result as demonstrating the placebo effect to be a myth (11), while other academic commentators either pointed out that worthwhile effects could still exist in some settings (12), or saw the review as a necessary scientific correction to set the bar differently for claims concerning placebo (13).

The review was updated in 2004 with similar findings (14), but the latest update from 2010 reported more multifaceted results (15). Large analgesic effects of placebo interventions were found in several well conducted trials. Furthermore, a considerable variation in effect could in part be explained by differences in trial design, for example, effect of placebo was larger when the intervention was a device (as compared with pill placebo).

The history of placebo shows that the assessment of the clinical significance of placebo has a very real potential to be biased. On the one hand, Beecher's approach of analysing placebo

groups without comparing with an untreated control group generates inflated estimations of placebo effects. Additionally, popular fascination with the placebo effect fuels unrealistic assessments of its therapeutic potential. The above-mentioned meta-analyses (13–15), involving progressively larger numbers of studies and subjects, challenges the belief that in general the placebo is powerful. Yet it is unwarranted to conclude that placebo interventions are incapable of producing clinically meaningful benefit. It is generally not possible to prove a negative; moreover the meta-analyses (13–15) identified several well-designed clinical trials with relatively large analgesic effects of placebo, and a general tendency for effects on patient-reported continuous outcomes. This, in addition to collateral evidence from laboratory experiments, points to the conclusion that placebo analgesia is a real phenomenon with the potential for clinical significance in some settings. However, estimating the size of the effect of placebo is subject to considerable uncertainty. The challenge in rigorously assessing the clinical benefit of placebo interventions is to reliably distinguish the magnitude of any real effect of placebo from the noise embedded in the human interaction of an experiment or a clinical trial. In this paper we explicate the types of bias that complicate the detection and measurement of clinically meaningful placebo effects.

Bias and placebo

In clinical epidemiology, bias constitutes deviation from the truth. As opposed to error caused by chance, bias represents a systematic distortion (16,17). The number of possible biases is large, and we make no claim for completeness.

The term 'placebo effect' is often used in a narrow meaning, implying the effect caused by a placebo intervention (or the treatment ritual). However, it is also used in a broader meaning implying the effect caused by a manipulation of the patient-doctor relationship, not necessarily involving a placebo intervention. In this paper we focus on trials and experiments comparing placebo with no-treatment, though the biases are equally relevant for trials and experiments comparing manipulations of the patient-provider interaction.

Biases in placebo research, such as response bias, can affect the internal validity of studies, raising doubts about whether, or to what extent, genuine placebo responses have been observed. Biases, such as publication bias, can also affect external validity, making the findings from experimental studies not accurately generalizable. In general, different types of biases may lead to overestimating or underestimating of the effects of placebo interventions. There is no clear evidence to support a ranking of which type of bias is more important, however, we consider response bias to play a major role (Table). Discussions of bias in controlled trials often focus on randomization and the need for concealment of allocation as a means to minimize selection bias. Selection bias is a general problem in clinical and experimental research (Table), and there is a large literature on the subject (18). In this paper, we have prioritized types of biases that we belief are more specific to the field of placebo.

The conundrum of response bias

The assessment of the placebo effect faces a basic conundrum. Patients may desire to please the researcher, or just give a "correct" or expected answer that fits with the experimental situation (19–22). When patients report that they feel better after receiving a placebo intervention how do we know to which degree this reflects genuine symptomatic improvement, such as pain relief, that can be attributed to the placebo effect or a response bias? Patient-subjects who receive placebo interventions in clinical trials or laboratory experiments, believing that it was or may have been a real treatment, might be disposed to report positive outcomes to please the investigators with whom they had a clinical

relationship. Conversely, those who did not receive any study intervention might be disappointed and disposed to report negative or "correct" outcomes.

This conundrum is posed by two different considerations. First, there is no blinded control for the placebo effect. Second, the placebo effect is most likely to play a role in the treatment of conditions in which the outcome targets are subjective (15), and necessarily based on introspective subject self-reports, for example pain. In controlled trials to assess the placebo effect (whether in clinical or laboratory settings), the placebo intervention is usually compared with a no-treatment control. Research subjects in the no-treatment arm necessarily know that they are not receiving treatment.

In assessing treatment efficacy of a pharmacological intervention with respect to subjective outcomes, such as relief of pain, blinded placebo-controlled trials are able to discriminate real effects from response bias. Patient-subjects may be disposed to report favorable outcomes by virtue of trial participation. But since they are randomized to masked drug or placebo, significant improvement in the treatment group as compared with the placebo group can be attributed to the efficacy of the treatment, as long as the masking conditions were successful. Response bias may operate to inflate the apparent drug effect (the difference between pre-trial baseline and the time of study outcome measurement); likewise, it may account for all or part of the response in the placebo arm. However, in view of randomization and blinding conditions, there is no reason to infer that the effect of response bias is greater in one arm than the other. In contrast, controlled trials to assess the placebo effect are not able to factor out response bias in this way, because they cannot be blinded.

Another important aspect of response bias is that it is likely to be closely associated with the same causal factors hypothesized to cause placebo effects: a warm patient-provider interaction and the doctor's verbal and non-verbal suggestion of an important beneficial treatment effect. Thus, the more a physician signals friendliness and confident expectation of improvement, the less likely is the patient to disappoint the doctor who is making such an effort. Recent qualitative studies of patients in randomized clinical trials have demonstrated that patients can become dramatically attached to the research team and very committed the 'success' of a trial (23).

The conundrum of response bias is not limited to the typical clinical trial design. To elucidate the placebo effect, Benedetti and colleagues have deployed an experimental paradigm that compares the responses of patients to analgesic drugs in conditions of open and hidden administration (24). For post-surgical patients receiving open injections of drugs in the manner of a typical clinical encounter, a given dose of an opioid drug appears to produce a substantially greater reduction in pain as compared with patients who receive the same dose of drug via a computerized infusion pump but are not informed about when the drug will be administered (25,26).

This paradigm has been interpreted as demonstrating a clinically meaningful placebo effect, or the placebo component of active treatment, without the use of a placebo control. The results are impressive, but can we reliably distinguish between a real, greater reduction in pain in the open treatment group from a response bias, given that the patients knew that they were being given an analgesic drug and that they were participating in an experiment to assess analgesia? Likewise, those receiving the hidden infusion may have been negatively biased in their assessment of pain relief, knowing that they were suffering from pain but not knowing when pain medication would be administered. The open/hidden design is not itself able to rule out the alternative possibility of response bias.

A possible solution to the conundrum of response bias is to design trials assessing placebo effects with objective outcomes, not susceptible to patient behavior, and to blind the

outcome assessor. This may be possible in some situations, such as studies of wound healing. However, even wound healing may be susceptible to variations in patient behavior; there is scant reliable evidence that placebo interventions modify objective outcomes in clinical trials (15); and what is important to patients is usually reduction in symptoms. Thus, trials with (only) objective outcomes would reflect a fairly limited number of clinically relevant problems.

Some have argued that neuroimaging technologies such as functional magnetic resonance (fMRI) and positron emission tomography (PET) can help determine whether placebo effects are independent of response bias (27). For example, one team of researchers has reported that placebo responses occur in pain-related areas of the brain *during the time of stimulation* and not only during assessment (28), while other researchers have shown that spinal cord mechanisms are involved with placebo analgesia (29). These experiments seem to indicate that at least some of the observed effect of placebo in an experimental setting is independent of response bias; however, they cannot rule out the hypothesis that some of the observed clinical effect is due to response bias.

In fact, other neuroimaging experiments point to potential involvement of response bias. For example, one study compared no-treatment to placebo treatment and placebo treatment plus naloxone. The placebo group reported significant pain reduction and the pain ratings of placebo treatment plus naloxone partially blocked the placebo behavioral response (30). But when one examines the simultaneous brain activation patterns of placebo with and without naloxone, there are inconsistencies. In the placebo treatment group, the average blood-oxygen-level-dependent (BOLD) response across all pain responsive brain regions decreased compared to controls. In the placebo treatment plus naloxone group, the BOLD actually shows that the naloxone group had increased activation compared to controls. For this group, instead of the partial blocking of pain sensitivity found in the behavioral data, there was brain activation that usually represents a worsening of pain (not a partial blocking of pain reduction). This finding suggests that what is reported is not necessarily congruent with what is felt, and, at least some of the time, pain self-reports due to placebo treatment are unrelated to the organic process of nociception.

While fMRI can measure hemodynamic blood oxygenation level dependent (BOLD) effect and PET can monitor regional cerebral blood flow and volume and map specific neuroreceptors using radiopharmaceuticals, neither method has advanced far enough to clearly and unequivocally distinguish to which extent the activations observed are "really" being felt by patients. In sum, just as there is no way to construct a blinded controlled trial to assess the placebo effect, so there is no way to eliminate subjectivity from patient-reported outcomes. This simply reflects the familiar but philosophically deep fact that there is no objective access to subjective experiences.

Other biases related to lack of blinding

Patients receiving placebo, and believing they are receiving (or have a fair chance of receiving) genuine treatment are less likely to seek alternative treatment, or to modify their basic care treatment, so-called co-intervention bias. For example, in several large acupuncture trials for various pain conditions, the patients in the placebo group in general tended to take less analgesic medication as compared with the no-treatment group at post-treatment, despite similar levels of treatments at baseline (31). If the basic care of the patients was sub-optimal, and an increase in, for example, analgesic dose would influence pain, a difference in analgesic drug use would tend to underestimate the true effect of placebo.

Another type of bias relevant for trials assessing the effect of placebo is attrition bias -that is, the bias caused by patients dropping out of the trial. Patients in the no-treatment group may tend to drop out of the trial, or to skip examinations, or not follow the trial protocol, more frequently than the patients receiving placebo. In one well performed three-armed acupuncture pain trial running for 12 weeks, the drop-out rate in the acupuncture and placebo acupuncture group were comparable, 9% vs 6% (32). In the no-treatment group it was 16%. This illustrates the difference in patient behavior when included in placebo groups and no-treatment groups. The degree of bias involved, and its direction, is difficult to predict, and depend on whether those leaving the no-treatment group had better or worse outcomes than those who stayed.

Both co-intervention bias and attrition bias seem more likely to have an effect in trials of long duration. Thus, their potential impact would be smaller in studies of short duration. However, such studies face the problem that what is usually clinically most relevant is the longer-term result of a treatment.

Publication bias and outcome reporting bias

Bias also can occur during analysis and reporting. It has become painfully clear that studies with positive results are more likely to be published than studies with a negative or neutral result, so-called publication bias (33–35). Similarly, it has become clear that positive results within studies are more likely to be reported, and more likely to be reported in sufficient detail to be usable for meta-analyses, than negative results, so-called outcome reporting bias (36). Though most of the studies analyzing selective publishing of studies, and selective reporting of results, have focused on randomized trials, it seems likely that the same phenomenon applies to experimental research.

In experimental studies of placebo effects with no third active arm, the 'interesting' result is a positive outcome for those receiving placebo. It seems likely that the same motivations driving authors, reviewers, and editors to enhance publication of studies with large effects of experimental interventions in general, also apply to the field of placebo research. This would mean that studies reporting large effects of placebo are published in prestigious journals, whereas studies with no effects would be dismissed as 'failed experiments', and are published less often, later, in less prestigious journals and more often as abstracts.

Matters are somewhat more complicated in three-armed trials aimed at assessing specific treatment efficacy, as the primary drive for publication bias and outcome reporting bias relate to the effect of the active treatment, and it is difficult to predict how this affects the publishing and reporting of results of placebo. However, the pooling of results from placebo and no-treatment groups is a problem. Several examples of such pooling based on a lack of statistical difference between placebo and no treatment has been published (22,36), though it is difficult to assess how often it takes place. In one trial of immunotherapy for ragweed allergy the placebo and no-treatment group were combined because of 'no statistically difference' in symptom score (33). Hopefully, the ongoing efforts to establish trial registers, and to have public access to trial protocols, will diminish this problem in the long run, also for placebo research (37).

The problem of the extrapolation of study results

It is always challenging to extrapolate study results from the controlled framework of a clinical trial or a laboratory-based experiment to the normal clinical encounter. However, within placebo research this mental and methodological leap is more difficult than usual. One reason for this in clinical trials assessing placebo effects is what could be called 'causal indeterminateness". A normal placebo-controlled trial is designed to isolate the hypothesized

causal factor that is the only difference between the treatment and the placebo control. For example, in a trial of the effect of a drug on irritable bowel syndrome the hypothesized causal factor (the drug) is an inherent component of the intervention (the tablet). However, if the trial also involved a comparison between the placebo intervention and a no-treatment/ usual care group, the causal factors that may be responsible for improved symptoms in the placebo group are not isolated, as they are for the drug group. Instead, they form a set of putative causal factors, including the physical appearance of the pill, the ritual of pill taking, and positive expectations associated with taking an intervention, which may be believed to be an active treatment (because it is indistinguishable from the drug intervention). In other words, the placebo intervention serves as a kind of causal surrogate for the true underlying causal factors.

One example of 'causal indeterminateness' concerns the possible interpretations of the effects of placebo acupuncture. Large three-arm trials including traditional acupuncture, sham intervention, and no-treatment/usual care groups report considerable effect of real and sham acupuncture on pain, but no difference between them (31). Acupuncture as well as sham acupuncture are complex interventions, which include a host of factors that may be causally responsible for placebo effects. One interpretation is that the true causal factors are the repetitive and prolonged treatment sessions involving a high degree of interaction between convinced acupuncturists and patients with a positive view of alternative medicine. In addition, the perception of needling—an invasive intervention, which may seem novel and exotic for acupuncture naïve patients—may contribute to positive expectations. Another interpretation is that the effects reflect, in part, lack of disclosure to patients that they could receive placebo, which is common (and ethically problematic) in these studies (38). This interpretation is in accord with the notion that the process of randomisation and informed consent may, at least for some outcomes, influence the effects of placebo (39). A third interpretation is that sham needling has a physiological analgesic effect, unrelated to the patient-provider interaction. Thus, we have no clear grasp of what are the true causal factors at play, and the various interpretations would have quite different clinical implications.

In laboratory studies it is possible to control the experimental situation more than in the standard clinical trial, but the main problem with laboratory studies is that they are far removed from the normal clinical situation. Non-clinical experimental studies on placebo almost always evaluate outcomes of very short duration, lasting for hours or a few days, and often involve healthy volunteers. Furthermore, the highly controlled environment with interactions of relatively long duration with laboratory personnel differs substantially from the hectic and uncontrolled environment of the normal clinical encounter, which usually is of comparatively short duration. The mechanistic laboratory studies are aimed at exploring the biological basis for placebo effects, and not at evaluating the effect of placebo in the normal clinical situation.

Thus, it is difficult to reliably extrapolate the results of placebo studies without 1) a better understanding of the causal factors and 2) a study framework that is designed to isolate the causal factors of interest and is as close to the clinical situation as possible. One strategy to address this problem of causal indeterminateness is to design clinical trials in which the presumed causal factors are characterized *a priori*, and implemented in accordance with a discrete protocol. For example, a trial may be designed to test the effect of a 'placebo intervention package' consisting of a well-defined treatment provider approach to a patient, for example a specific placebo intervention in combination with a positive consultation. Further studies might vary the components of the placebo intervention package to determine the relative contributions of discrete elements to the size of placebo effects. Although hampered by logistical and financial challenges, this may be a promising avenue for further research (40,41). Progress in translating placebo research into clinical practice depends on

developing hypotheses relating to the causal factors responsible for placebo effects in the clinical setting and designing rigorous experiments to test them.

Conclusion

Despite methodological limitations, randomization to placebo and no-treatment is the best research design we have in estimating effects of placebo, both in a clinical and in an experimental setting. However the design remains an approximate and fairly crude method for assessing the true effect of placebo interventions. The conundrum of response bias when outcomes are patient-reported, and other biases make study results challenging to interpret. The placebo effect and its various mechanisms of action is of great interest for neuroscience; however, its interest for medicine depends primarily on the extent to which placebo interventions can be found to produce reliable and clinically relevant therapeutic benefits. Creative experimental efforts are needed to assess rigorously the clinical significance of placebo interventions and investigate the component elements that may contribute to therapeutic benefit.

Acknowledgments

References

1. Kaptchuk TJ. Intentional ignorance: A history of blind assessment and placebo controls in medicine. Bull Hist Med. 1998; 72:389–433. [PubMed: 9780448]

2. Kaptchuk TJ. Powerful placebo: the dark side of the randomized controlled trial. Lancet. 1998; 351:1722–25. [PubMed: 9734904]

3. Beecher HK. The powerful placebo. JAMA. 1955; 159:1602–6.

4. Hróbjartsson A. What are the main methodological problems in the estimation of placebo effects. J Clin Epidemiol. 2002; 55:430–5. [PubMed: 12007544]

5. Kienle GS, Kiene H. The powerful placebo effect: fact or fiction? J Clin Epidemiol. 1997; 50:1311–18. [PubMed: 9449934]

6. Levine JD, Gordon NC, Fields HL. The mechanisms of placebo analgesia. Lancet. 1978; 2:654–7. [PubMed: 80579]

7. Sauro MD, Greenberg RP. Endogenous opiates and the placebo effect. A meta-analysic review. J Psychosom Res. 2005; 58:115–20. [PubMed: 15820838]

8. Benedetti F, Pollo A, Lopiano L, Lanotte M, Vighetti S, Rainero I. Conscious expectations and unconsious conditioning in analgesic, motor, and hormonal placebo/nocebo responses. J Neurosci. 2003; 24:4215–23.

9. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Biological, clinical and ethical advances in placebo effects. Lancet. 2010; 374:686–95. [PubMed: 20171404]

10. Hróbjartson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Eng J Med. 2001; 344:1594–602.

11. [accessed March 17 2010]. www.nytimes.com/2001/05/24/us/placebo-effect-is-more-myth-than-science-study-says.html

12. Brody H, Weismantel D. A challenge to core beliefs. Adv Mind Body Med. 2001; 17:296–8. [PubMed: 11931054]

13. Miller FG, Colloca L, Kaptchuk TJ. The placebo effect: illness and interpersonal healing. Perspectives Bio Med. 2009; 52:518–39.

14. Hróbjartson A, Gotzsche PC. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. J Intern Med. 2004; 256:91–100. [PubMed: 15257721]

15. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev. 2010; (1):CD003974. [PubMed: 20091554]

16. Sackett DL. Bias in analytic research. J Chron Dis. 1979; 32:51–63. [PubMed: 447779]

17. Kaptchuk TJ. Effect of interpretive bias on research evidence. BMJ. 2009; 326:1453–5. [PubMed: 12829562]

18. Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials. Cochrane Database Syst Rev. 2007; 18(2):MR000012. [PubMed: 17443633]

19. Allan LG, Siegel S. A signal detection theory analysis of the placebo effect. Eval Health Prof. 2002; 25:410–20. [PubMed: 12449084]

20. Orne MT. On the social psychology of the psychological experiment. Am Psych. 1962; 17:776–83.

21. Cleophas TJM. The importance of placebo effects. JAMA. 1995; 273:283. [PubMed: 7815648]

22. Knight LJ, Barbaree HE, Boland FJ. Alcohol and the balanced-placebo design: the role of experimenter demands in expectancy. J Abnormal Psych. 1986; 35:335.

23. Kaptchuk TJ, Shaw J, Kerr CE, Conboy LA, Kelley JM, Lembo AJ, Csordas TJ, Jacobson EE. "Maybe I made up the whole thing:" Placebos and patients experiences in a randomized controlled trial. Culture Med Psych. 2009; 33:382–412.

24. Colloca L, Lopiano L, Lanotte M, Benedetti F. Overt versus covert treatment for pain, anxiety, and Parkinson's disease. Lancet Neurol. 2004; 3:679–84. [PubMed: 15488461]

25. Amanzio M, Pollo A, Maggi G, Benedetti F. Response variablity to analgesics: a role for non-specific activation of endogenous opioids. Pain. 2001; 90:205–15. [PubMed: 11207392]

26. Benedetti F, Maggi G, Lopiano L, et al. Open versus hidden medical treatments: the patient's knowleddge about a therapy affects the therapy of outcome. Prevention Treatment. 2003; 6 ArtID 1a.

27. Wager TD, Rilling JK, Smith EE, et al. Placebo-induced changes in fMRI in anticipation and experience of pain. Science. 2004; 303:1162–66. [PubMed: 14976306]

28. Price DD, Craggs J, Verne GN, Perlstein WM, Robinson ME. Placebo analgesia is accompanied by large reduction in pain-related brain activity in irritable bowel syndrome. Pain. 2007; 17:63–77. [PubMed: 16963184]

29. Eippert F, Finsterbusch J, Bingel U, Cuchel C. Direct evidence for spinal cord involvement in placebo analgesia. Science. 2009; 326:404. [PubMed: 19833962]

30. Eippert F, Bingel U, Schoell EC, Yacubian J, Klinger R, Lorenz J, Buchel C. Activation of the opioidergic descending pain control system underlies placebo analagesia. Neuron. 2009; 63:533–43. [PubMed: 19709634]

31. Madsen MV, Gøtzsche PC, Hróbjartsson A. Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. BMJ. 2009; 338:a3115. [PubMed: 19174438]

32. Linde K, Streng A, Jürgens S, et al. Acupuncture for patients with migraine. A randomized controlled trial. JAMA. 2005; 293:2118–25. [PubMed: 15870415]

33. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: Review of publication and presentation. PloS. 2008; 5:e217.

34. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent effficacy. N Eng J Med. 2008; 358:252–60.

35. Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004; 291:2457–65. [PubMed: 15161896]

36. Grammer LC, Shaughnessy MA, Shaughnessy JJ, Patterson R. Asthma as a variable in a study of immunotherapy for allergic rhinitis. The Journal of Allergy and Clinical Immunology. 1984; 73(5 part 1):557–60. [PubMed: 6715723]

37. Sim I, Chan AW, Gülmezoglu AM, Evans T, Pang T. Clinical trial registration: transparency is the watchword. Lancet. 2006; 367:1631–3. [PubMed: 16714166]

38. Miller FG, Kaptchuk TJ. Acupuncture trials and informed consent. Journal of Medical Ethics. 2007; 33:43–44. [PubMed: 17209110]

39. Bergmann JF, Chassany O, Gandiol J, Deblois P, Kanis JA, Segrestaa JM, Caulin C, Dahan R. A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. Clin Trials Metaanal. 1994; 29:41–7. [PubMed: 10150184]

40. Kaptchuk TJ, Kelley JM, Conboy LA, Davis RB, Kerr CE, Jacobson EE, Kirsch I, Schyner RN, Nam BY, Nguyen LT, Park M, Rivers AL, McManus C, Kokkotou E, Drossman DA, Goldman P, Lembo AJ. Components of the placebo effect: a randomized controlled trial in irritable bowel syndrome. BMJ. 2008; 336:999–1003. [PubMed: 18390493]

41. Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, Singer JP, Kowalczykowski M, Miller FG, Kirsch I, Lembo AJ. Placebos without deception: a randomized controlled trial in irritable bowel syndrome. PLoS One. 2010; 5:e15591. [PubMed: 21203519]

Table

Main types of challenges to the reliability and generalizability of randomised trials and experiments assessing the effect of placebo

Challenges	Characterisation	Mechanism	Likely impact
Selection bias[18]	Selection of patients or experimental subjects with different prognosis into compared groups	Patients included in the compared groups differ at baseline due to either random events or preferred selection of one type of subjects to the experimental group	Overestimation of the effect of the placebo in studies only involving placebo vs. no-treatment. Unclear impact on the estimated effect of placebo in studies involving active vs placebo vs no-treatment.
Response bias[19,20]	The tendency for patients or experimental subjects to report their symptoms in a way they feel is socially acceptable or desirable	Patients or experimental subjects in the placebo group may report symptoms more optimistically than in the no-treatment group	Overestimate placebo effects of patient reported outcomes, for example pain and nausea
Co-intervention bias[31]	The tendency for patients or experimental subjects to seek out and get treatment that is not part of the trial or the experiment	Patients or experimental subjects in the no-treatment group may be more inclined to seek out non-protocolised interventions	Underestimate placebo effects when the non-protocolised intervention has a clinical effect, either due to a placebo effect or a non-placebo effect
Attrition bias[32]	The tendency for patients or experimental subjects to drop out of the trial or the experiment	Patients or experimental subjects in the no-treatment group may me more inclined to drop out	Unclear. The degree of bias and its direction depend on whether those leaving the no-treatment group had better or worse outcomes than those who stayed.
Outcome reporting bias[35]	The tendency in scientific publications for statistically significant outcomes to be selected for reporting more frequently than outcomes with insignificant results	The authors of scientific publications often report only a subset of the outcomes studied, and tend to select those with statistically significant results	Overestimate placebo effects in articles aimed at studying placebo. Unclear impact on articles aimed at studying an active intervention (typically active vs placebo vs no-treatment)
Publication bias[34]	The tendency for scientific publications with a statistically significant result to be published more frequently than studies with an insignificant result	Published scientific studies often reflect only a subset of the studies conducted, and those published tend to report statistically significant results	Overestimate placebo effects in articles aimed at studying placebo. Unclear impact on articles aimed at studying an active intervention (typically active vs placebo vs no-treatment)
Causal indeterminateness in clinical trials	A placebo intervention will often serve as a 'surrogate' causal factor for the largely indetermined true causal factors	The causal factors of the placebo effect are not typically imbedded in the placebo intervention per se, but in the patient-provider interaction	Competing interpretations of which causal factors are most important in a study finding large effects of placebo would typically have very different clinical implications
Non-clinical settings in laboratory experiments	A laboratory experiment will differ from the typical clinical situation in important ways	Non-clinical experimental studies on placebo tend to be of very short duration and may involve healthy volunteers	Provide valuable insight into the neurobiology and mechanisms of placebo effect, but results cannot reliably be extrapolated to a clinical setting
Informed consent and randomisation[39]	The trial or experiment may interact with the patients included	Informing patients about being part of a trial or experiment may alter preconceptions and beliefs	May underestimate or overestimate placebo effects. Beliefs in the effect of an interventions may be less pronounced compared with a clinical situation

18–20;31–32,34–35,39 Numbers refer to references to articles in manuscript text