Replication crisis and placebo studies: rebooting the bioethical debate

Charlotte Blease,[1,2] Ben Colagiuri,[3] Cosima Locher[4,5]

[1]General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
[2]Digital Psychiatry, Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
[3]School of Psychology, University of Sydney, Sydney, New South Wales, Australia
[4]Department of Consultation-Liaison Psychiatry and Psychosomatic Medicine, University Hospital Zurich, Zurich, Switzerland
[5]Psychology, Plymouth University, Plymouth, UK

Correspondence to
Dr Charlotte Blease, General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215, USA;
charlotteblease@gmail.com

Received 28 September 2022
Accepted 30 December 2022

To cite: Blease C, Colagiuri B, Locher C. *J Med Ethics* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jme-2022-108672

ABSTRACT

A growing body of cross-cultural survey research shows high percentages of clinicians report using placebos in clinical settings. One motivation for clinicians using placebos is to help patients by capitalising on the placebo effect's reported health benefits. This is not surprising, given that placebo studies are burgeoning, with increasing calls by researchers to ethically harness placebo effects among patients. These calls propose placebos/placebo effects offer clinically significant benefits to patients. In this paper, we argue many findings in this highly cited and 'hot' field have not been independently replicated. Evaluating the ethicality of placebo use in clinical practice involves first understanding whether placebos are efficacious clinically. Therefore, it is crucial to consider placebo research in the context of the replication crisis and what can be learnt to advance evidence-based knowledge of placebos/placebo effects and their clinical relevance (or lack thereof). In doing so, our goal in this paper is to motivate both increased awareness of replication issues and to help pave the way for advances in scientific research in the field of placebo studies to better inform ethical evidence-based practice. We argue that, only by developing a rigorous evidence base can we better understand how, if at all, placebos/placebo effects can be harnessed ethically in clinical settings.

INTRODUCTION

Placebos have been used by doctors for centuries.[1] Placebo research, however, is a relatively new and emerging field with potentially enormous implications for clinical research and practice. The placebo effect occurs when the treatment context generates expectancies that trigger therapeutically beneficial outcomes.[2] In the archetypical example, a sugar pill administered under the guise of an analgesic produces pain relief. Hundreds of studies report placebo effects for conditions ranging from experimentally induced pain to Parkinson's disease.[3 4] Deception was traditionally considered necessary to elicit a placebo effect. However, recent research on the open-label placebo (OLP) approach, where placebo treatments are openly administered, challenges that idea.[5 6] A growing number of studies have reported that OLPs can be effective for various conditions, including in chronic pain, mental disorders, healthy individuals and in other physical complaints.[5] This has led to increasing calls from within the field to harness placebo effects clinically.[2 7]

In parallel, many studies across the world indicate that general practitioners commonly use placebos in clinical practice. In 2018, a systematic review and meta-analysis with data drawn across 12 countries found that 53% to 89% of primary care physicians reported using placebos in their clinical practice at least monthly, and 16% to 75% at least weekly.[8] People with medically unexplained symptoms, chronic primary pain or fatigue-related conditions, anxiety and those with conditions perceived as psychosomatic are more likely to be prescribed placebos.[9 10] Many doctors prescribe placebos in order to harness therapeutic benefits to patients, via placebo effects.[10] Notably, this is an approach that appears to be supported by the American Medical Association's code of ethics, which approves placebo use if the patient is aware the doctor would like to prescribe a placebo and the patient can provide informed consent.[11]

Against this growing interest in and use of placebos, some researchers—including those outside the field—have raised concerns about the reliability of placebo research on the basis that many findings are likely overstated.[12–14] Some of those concerns have been effectively challenged; for example, previous comprehensive meta-analyses suggesting no evidence of a placebo effect have been criticised for overincluding conditions and symptoms for which placebo effects are not believed to be relevant[15] However, other identified concerns, such as the conflation of placebo responses with placebo effects,[15–17] do not always appear to have been consistently considered or attended to within the literature.[18] The term 'placebo response' refers to an undifferentiated amalgam of changes that can arise after administration of placebos or treatments in clinical trials and encompass: participant response biases, natural history, Hawthorne effects (the potential for participants to change their behaviour when monitored), methodological biases and placebo effects proper. Importantly, conflating placebo responses with placebo effects means researchers are likely to inflate the size of placebo effects.[15 18]

Advancing the debate about the clinical evidence for or against placebos, we build on a key concern that has been discussed within medical and psychological research but has been overlooked in placebo studies: namely, 'replication.' This is not a new idea: publication has long required that researchers carefully present their methods to facilitate independent replication by other scientists. Yet, in many scientific fields, including medicine there is concern that a large proportion of findings in the literature may not be replicable. This is dubbed the 'replication crisis', which refers to the discovery that often classic—that is, well regarded and highly cited—findings in the literature may not be reproduced in subsequent studies.[19 20] This finding, in turn, has led researchers to raise questions about the methodological integrity of research.

Stanford University Professor John Ioannidis helped identify the scale of the problem. He set out to investigate the reliability of novel treatments. Ioannidis probed highly cited studies published in three top medical journals between 1990 and 2003.[21] He found fewer than half of the original treatments he investigated were supported by further research, a quarter were never tested again, 16% were found to be less effective than first thought, with a further 16% found to be completely ineffective. Prasad and Cifu delved deeper to gain some traction on the frequency with which treatments or interventions are later contradicted.[22] In an article published in 2011, they reviewed every article published in 2009 in *The New England Medical Journal*[23] identifying 35 articles that probed the evidence for current medical practices. These studies revealed that nearly 50% of accepted practices were ineffective. In a later article, Prasad, Cifu and a team of colleagues sifted through all the original research papers published between 2001 and 2010 in *The New England Medical Journal*. In total, 363 papers tested an established medical practice; among them, 40% contradicted earlier published findings with 38% confirming earlier studies.[22]

Prasad and Cifu call practices that were once considered unimpeachable and later discovered to be ineffective or harmful 'medical reversals'. Medical reversals encompass medications, surgeries and even public health programmes.[22] In their 2015 book on the topic, they offer a range of examples including atenolol, which was one of the first drugs prescribed in the treatment of high blood pressure. In a major study conducted in 2004, researchers found it was no better than taking a placebo. On the other hand, aprotinin, a drug widely prescribed after cardiac surgery was later found to increase the risk of death. Similarly, used to treat patients with angina or those who have recently suffered a heart attack, clinical trials now show stents—small metal tubes surgically implanted into arteries—neither reduce chest pain nor curb the risk of further heart attacks.

Medical reversals could be considered a key feature of progressive, evidence-based medicine. However, this relies on scrutinising previous findings, even ones which seem incontrovertible. Replication, then, is a central tenet of science and evidence-based medicine. Any single finding can be subject to many biases, including systemic (eg, publication bias[24]) and researcher biases (eg, questionable research practices[25]). Replication by independent researchers strengthens confidence in findings. These are concerns that have not yet been scrutinised in placebo research but are crucial to evaluate the clinical effectiveness and ethicality of placebo treatments.

The paper begins by situating discussion about the potential benefits of placebos/placebo effects, in the bioethical literature on placebo use in clinical practice. Next, we turn to the evidence, and identify several problems arising from the replication crisis, and other questionable research practices, indicating how they pertain to placebo studies. Following this, we offer recommendations and advice for future directions in the field of placebo studies to address both empirical and ethical concerns emerging from lack of adequate replication.

PLACEBOS AND CLINICAL ETHICS

Brief background on the bioethical debate

For many years now, an extensive bioethics literature has examined intricate concerns about clinicians prescribing deceptive placebos.[1] This scholarship has overwhelmingly focused on the conflict that arises between patient autonomy and clinical beneficence when placebo treatments are deceptively prescribed to patients.[26–28] This bioethical literature pivots on questions about whether it is ethical to deceive patients if doing so alleviates symptoms or promotes their well-being, balanced with the potential to undermine patient autonomy—or, indeed, whether any such infringements on autonomy are morally significant.[27 29]

A more recent proposal, aimed at resolving this perceived dilemma is the proposal of OLPs to harness placebo effects.[30] Although less explored by bioethical researchers, it has recently received some attention.[18 31–33] While OLPs appear to offer transparency combined with symptom relief, some scholars question whether, if routinely offered, such treatments might be more likely to be prescribed to patients whose symptoms are already incorrectly psychologised by clinicians.[31] Others worry whether prescribing OLPs leads to secondary harms from stigmatisation, with the understanding that the clinician has discredited their symptoms as 'all in their head'.[16 18] Furthermore, there are concerns about the acceptability of the approach among patients and clinicians.[10 34–36]

Although current bioethical explorations are invaluable to evaluate the ethical use of placebos by clinicians, a different but perhaps more neglected assumption in medical ethics lies with the underlying premise that deceptive placebos and OLPs are clinically effective.[1 18 26–28 37] If the clinical benefits of placebos, or placebo effects, are overestimated or considerably less valuable than is often stated, then the justification for placebo use requires additional, ethical argumentation. Whether such argumentation can be persuasively mounted is beyond the remit of this paper. Our focus, which we will shortly address, is with the more elementary concern: the quality of evidence for the clinical effectiveness of placebos.

Ethical evidence-based practice

Further impetus for exploring the strength of placebo research is medicine's embrace of evidence-based practice. This approach carries repercussions for clinicians' duty of professional competence—what have been described as 'epistemic duties'—associated with the responsible acquisition of empirical knowledge and its application in clinical practice.[38] Professional competence—the clinician's ability to accurately assess the patient's problems, make diagnoses and recommend treatments—depends on the ability to keep up to date with, and critically evaluate evidence. If clinicians use placebos—whether deceptively or openly—it needs to be determined that these interventions are effective in clinical practice. Placebo studies is a particularly complex field (and shortly, we will identify some of the challenges researchers face); nonetheless, upholding evidence-based practice, we argue, requires clinicians to recognise the strengths and limitations associated with this research. As such, we hope that by identifying methodological issues concerning placebo studies we can help arm clinicians—who are already pressed for time—to better evaluate placebo research to help inform ethical evidence-based practice.

The duty of professional competence also carries consequences for the duty to respect the patient's autonomy. Particularly in the case of OLPs, the quality of disclosures about the effectiveness of placebos and of placebo effects, will be informed by how adequately clinician knowledge reflects the state of evidence in placebo studies. Patients may also assume that because a treatment is offered it is therefore acceptable; for example, speaking in the case of psychotherapy and clinical psychology, O'Donohue and Henderson note, 'Consumers can make false knowledge assumptions regarding what a particular mental health professional or the profession as a whole actually has' and 'health professionals can make false representations or passively accept false client assumptions regarding their

knowledge'.[39] These same considerations may also arise with placebo use. For example, illustrating the implicit trust patients have in clinicians' knowledge of the effectiveness of placebos, a recent qualitative study exploring patients views about both deceptive and OLPs reported patients implicitly trusted the competence of their doctor, tending to assume that, regardless of whether the placebo was open or deceptive, if such treatments were offered they must be effective: or as one participant summed it up, 'It's not my greengrocer [prescribing the treatment], it's someone from the medical profession.'[36] Bridging the relationship between evidence and ethics, next we turn our attention to empirical concerns associated with placebo studies.

REPLICATION AND QUESTIONABLE RESEARCH PRACTICES

Many problems arising from the replication crisis and questionable research practices across science have not been examined in placebo studies. This has led some to argue that placebos/placebo effects are overstated and likely have limited clinical relevance.[13 14] The only way to adjudicate between these views is to develop a rigorous evidence base that either supports or does not support the clinical efficacy of placebos/the placebo effect. Drawing on, and adapting a list of systemic and methodological problems that have been identified by other medical scholars,[19 40] we describe eight general problems and provide examples of how they may apply to placebo research. We emphasise that many of these problems do not suggest mal intention on the part of researchers—we identify ourselves as placebo researchers, and we too have fallen prey to many of these problems. Nonetheless, we suggest that to advance the field, it will be important to engage with these concerns more explicitly and pro-actively.

Systemic issues

Publication bias and 'canonisation' of findings

Publication bias is a well-known phenomenon whereby positive findings are more likely to be published.[24] Likewise, some meta-analyses in the field of placebo reveal an asymmetrical distribution among the included studies or an unknown level of reporting bias (eg,[41]) suggesting that a publication bias might be present. Greater awareness has led to increased efforts to publish null findings in traditional journals or open access repositories. However, one effect of historical publication bias is that some unsubstantiated findings may become 'canonised'. Canonisation occurs when a finding becomes widely accepted and is rarely scrutinised further. Resource-wise, it makes sense that findings with sufficient evidence become canonised—why invest resources testing a hypothesis known to be true? The problem is that multiple replication attempts may have failed but these have not been published due to publication bias, the 'file drawer' problem.[24] Some findings in placebo research appear to have reached canonisation but, to the best of our knowledge, have not been empirically scrutinised.

For example, based on a single study in duodenal ulcers,[42] it is widely believed that the more placebo pills administered, the stronger the placebo effect.[43] Similarly, it is frequently reported that the coloration of pills can influence placebo effects with red pills prompting greater pain relief than other colours.[44] Again, this appears to be based on only one single finding published in 1974 which enrolled a limited sample size of 22 participants, 5 of whom received red pills.[45] It is also commonly reported, based on a single source,[46] that blue pills prompt soporific placebo effects a claim that has not been subjected to robust clinical investigation. Subsequent, though only limited recently published studies report conflicting associations between colour and placebo effects.[47 48] Finally, the size or shape of pills is often assumed relevant to the size of placebo effects; however, once again, this appears to be based on a single early study which examined only perceived efficacy of medications and did not directly investigate placebo effects.[49] We do not imply these claims are false but simply that they have reached canonisation without a commensurate number of published replications.

Hotness of the field and grant culture

A paradoxical finding is that the hotter a scientific field the less likely published findings are to be true.[19] In the last two decades, publications in placebo research have increased tenfold,[50] with many appearing in the most revered and highly cited medical journals. Placebo research has also gained considerable international media attention (eg, *BBC, New York Times*). This 'hotness' could signal the need for caution and replicability. Relatedly, artefacts of the grant culture, such as hyperspecialisation, emphasis on programmatic research, and desirability of newsworthy findings may drive 'hotness'. Specialised, programmatic research is necessary to generate focused studies, but may carry significant downsides by increasing the risk of confirmation bias, creating disincentives for contradictory approaches and exclusion of 'big picture thinkers'.[40] In turn, grant culture may cause a positive feedback loop perpetuating research canonisation, dissemination of exaggerated claims and replication disincentives.

Conflicts of interest

Considerable evidence shows that conflicts of interest ('COI')—both financial and non-financial—can influence clinical decisions and research outcomes.[51] We are not aware of any financial COIs that might influence clinician decisions to administer placebos. However, non-financial COIs may still exist.[52] As Ioannidis notes,[19] researchers may be drawn to a particular field because it aligns with their prior beliefs, and this may prejudice research findings. Relatedly, research shows non-financial COIs and 'allegiance to the field' can foster biases via peer and grant review, selective citations or in the worst-case scenario, suppression of research findings.[19 51 53] We are not aware of any studies exploring the existence of such biases in placebo research. However, it would be quite optimistic to assume that placebo research was the outlier in which non-financial COIs and researcher allegiance play no role.

Methodological problems

Identification and measurement of effect size

Even before the need for replication, many results in placebo studies may be questionable. As indicated in the Introduction section, problems arise if placebo responses are not differentiated from placebo effects.[16 54 55] Furthermore, outside of pharmacological clinical trials there is ongoing debate about how to conceive of and measure placebo effects in complex treatments both in conventional (eg, psychotherapy, psychological approaches[56–58]), and complementary and alternative medicines (eg, acupuncture,[59] mindfulness treatments[60]). Placebo effects are also sometimes interchangeably conceived as synonymous with 'contextual factors,' or 'non-specific' treatment effects.[61–64] These considerations may lead to unintentional inflation of placebo effects (see also: Researcher Degrees of Freedom).

Control groups in placebo research

Double-blind randomised placebo-controlled trials (RCTs) are considered the 'gold standard' of medical research because they intend to control participant and researcher biases via blinding and other features. Similarly, control conditions are considered

Table 1 Identified replication problems and recommendations for placebo studies

Problem	Description	Placebo research	Recommendations
Systematic issues			
Publication Bias and 'canonisation' of findings	Publication Bias: Positive findings are more likely to be published. Canonisation: A finding becomes widely accepted and is rarely scrutinised further.	Some findings in placebo research appear to have reached canonisation but have not been empirically scrutinised. For example, based on limited studies, it is widely believed that the more placebo pills administered, the shape of pills, their coloration and the no of pills taken per day can influence the size of the placebo effect.	We recommend placebo research should strive to identify and conduct replications on early landmark studies that may have become canonised unjustifiably
Hotness of the field and grant culture	The hotter a scientific field, the less likely that published findings are true.	In the last two decades, publications in placebo research have increased 10-fold[50] gaining considerable media attention. This 'hotness' could signal the need for caution and replicability.	Encouragingly, some national science foundations have released funding specifically for replications, for example, the Netherlands Science Foundation. Greater funding should be allocated to incentivise replication.
Conflicts of interest (COI)	'COI'—both financial and nonfinancial—can influence clinical decisions and research outcomes.	We are not aware of any financial COIs that might influence clinician decisions to administer placebos. However, nonfinancial COIs such as researcher allegiance to 'mind-body' healing may exist.	Strategies to mitigate potential nonfinancial COIs should be implemented (eg, collaborations between researchers with opposing theoretical alignments). Also, researcher allegiance could be disclosed in placebo trials.
Methodological problems			
Identification and measurement of effect size	Conceptual confusions may incline researchers to unintentionally inflate effect sizes when designing studies and measuring effect sizes.	Even prior to replications, placebo responses are often conflated with placebo effects leading to inflated reporting. Similar problems arise in psychotherapy and clinical psychology settings, where the term placebo effect is often used interchangeably with 'contextual factors' or 'nonspecific' treatment effects.	We recommend researchers take care to differentiate and explicitly define key terminology including 'placebo response', 'placebo effect' and to be clear about how they are measuring effects in different domains.
Control groups	Double-blind randomised placebo-controlled trials are considered the 'gold standard' of medical research because they intend to control participant and researcher biases.	Persistent misunderstandings about how to conceive of 'placebos' as methodological tools (compared with placebos as treatments) is an impediment to determining the most appropriate control condition in placebo research. This is further complicated by the inherent difficulty of blinding participants and researchers to placebo interventions.	We recommend placebo researchers should be more explicit in their interpretation of the term 'placebo' within clinical settings and pay attention to structural equivalence of controls, comparator trials, and include the blinding researchers, assessors and data analysts.
Researcher degrees of freedom	Researchers must make many choices when collecting and analysing data, which can introduce biases that artificially inflate positive findings.	To our knowledge, no studies have explored evidence of p-hacking or other researcher degrees of freedom biases in placebo research.	We recommend that basic placebo research adopt preregistration as standard practice (eg, AsPredicted.org) to help combat this problem.
Low-powered studies	Researchers consider a study to be adequately powered if it has ≥80% chance of detecting a significant effect. Underpowered studies can lead to problems in any field.	We are not aware of any research assessing the typical power in placebo studies. However, many studies in the field have groups of 15–20 participants. As such, it seems likely that many placebo studies would have <80% power.	We recommend placebo studies should be powered to at least 80% in general and could aim to be powered to 90% when attempting to establish clinical efficacy.
Self-selection biases	Clinical and experimental research can be influenced by sampling biases and under-recruitment from diverse populations.	Placebo studies disproportionately rely on self-selection among recruits interested in 'mind-body' healing. Only recent studies have recruited individuals from ethnic and racial minorities to explore placebo and placebo effects in different populations.	We propose that researchers should strive for greater diversity in participant recruitment wherever possible and to conduct replications of results derived from more diverse samples.

necessary to establish evidence of a placebos/placebo effects when these are considered as potential clinical interventions.[2] The challenge then is to adequately differentiate two distinctive but very different meanings of the term placebo: one as a potential therapeutic treatment and one as a methodological tool (a control) for measuring treatment effect sizes in clinical trials. Again, even before replications of placebo treatments (including honestly prescribed or so-called OLPs) can fruitfully get underway, as a minimum, as we have already noted, the placebo treatment must be compared with natural history to rule out biases such as regression to the mean, response biases and Hawthorne effects.[65] However, persistent misunderstandings about how to conceive of 'placebos' in research contexts,[18 54 58 66] and the challenges of addressing this problem, are impediments to determining the most appropriate control condition to evaluating these treatments.[18]

The challenge of lack of robust controls is further complicated by the inherent difficulty of blinding participants in placebo research. Most commonly, participants receiving placebo treatment are very intentionally led to believe they are receiving active treatment while their controls knowingly receive no treatment or are put on a waiting list. Waiting list control groups, in turn, have been associated with nocebo effects— adverse health effects that may arise as a result of negative expectations about treatments.[67] This opens the door for demand characteristics and other participant biases in placebo research.[65] Placebo researchers also often tend to be aware of participants' allocation because they typically administer the intervention, introducing the possibility of experimenter effects.[68] Also, clinicians are often aware about whether a pill is a placebo or not, based on participants' experiences of side effects that are known to be specific for the pharmacological intervention under investigation.[69] Identifying control conditions that adequately equate to those in typical double-blind RCTs is inherently challenging in placebo studies.[70]

Researcher degrees of freedom

A major methodological concern identified in the replication crisis is 'researcher degrees of freedom'. Researchers must make many choices when collecting and analysing data, which can introduce biases that artificially inflate positive findings.[71] P-hacking is a prominent example whereby researchers may collect or select data (eg, outcomes) or statistical analyses until a statistically non-significant results becomes significant.[72] To our knowledge, no studies have explored evidence of p-hacking or other researcher degrees of freedom biases in placebo research, but it would be surprising if the field was immune given its prevalence in other domains of research.

Low-powered studies

In heavily relied on null hypothesis significance testing, power analysis defines the probability a study will reject the null hypothesis when it is, in fact, false and determines the probability of finding a true-positive result.[73] Typically, researchers consider a study to be adequately powered if it has ≥80% chance of detecting a significant effect.[74] Besides being an arbitrary value and higher power always being preferable,[75] underpowered studies can lead to problems in any field since small studies usually report stronger intervention effects than larger studies.[76] Perhaps most relevant is the fact that low statistical power generally undermines and reduces the chance of detecting a true effect.[77] We are not aware of any research assessing the typical power in placebo research. However, many studies in the field recruit groups of 15–20 participants. As such, it seems likely that many placebo studies would have <80% power and be categorised as 'small'. This is certainly the case relative to clinical trials where there is an expectation of obtaining ≥90% power.

Self-selection biases

Clinical and experimental research can be influenced by sampling biases and under-recruitment from diverse populations. For example, psychological research disproportionately relies on people from Western, educated, industrialised, rich and democratic societies (the so-called 'WEIRD' problem).[78] This is problematic when attempting to generalise research findings to populations beyond those studied. This makes it difficult to know whether potential benefits of interventions to harness the placebo effect also apply to people located in other countries or regions of the world. Moreover, even within 'WEIRD' societies, we are not aware of any research specifically seeking to understand placebo effects in people with low incomes, or persons with comorbidities; only recent studies have begun to explore placebo effects among people who identify as non-white.[79–81]

There are also some specific self-selection biases that may apply to placebo research. For example, in the case of emerging OLP research, trials are often advertised as investigating 'mind-body' effects and may appeal to individuals who have a desire to demonstrate the power of the mind. This could lead to participant bias in such studies. Unfortunately, even high-impact journal articles have failed to include information about the wording embedded in recruitment advertising materials.[82]

FUTURE DIRECTIONS

To take this empirical research agenda further, we offer recommendations for how researchers might address some of the challenges we have reviewed (see table 1). We emphasise that our intention is to raise awareness of these problems to advance progress in the field to be better able to determine the clinical relevance of placebos, thereby helping inform ethical clinical use. As our discussion of replication shows, the field of placebo studies is particularly complex. Upholding ethical evidence-based practice will therefore require clinicians to recognise the strengths and limitations associated with this research, and we urge that this will not be an easy task. In light of the prevalence of placebo use, and the media attention given to placebo studies, we argue that greater understanding about this complex field is nonetheless important for both research and clinical practice. Should further evidence demonstrate the clinical effectiveness of placebos/placebo effects, further work will be needed to explore how best to communicate this evidence effectively including to a wide range of patient populations. Such disclosures may also need to be carefully formulated in ways that do not cause secondary harms such as self-stigmatisation.

Finally, although not discussed in detail in this article, we underline that our arguments about the need for replication also apply to nocebo effects—the so-called 'evil twin' of placebo effects. Nocebo effects is the term for adverse health effects which are believed to arise as a result of negative expectancies about treatments.[83] This phenomenon is even less investigated than placebo effects, and we caution that critical attention should be given to the challenges of identifying nocebo effects, and of embarking on replicability in this field too. This is particularly significant because there may be a temptation to translate findings from the field of placebo studies to nocebo studies.

CONCLUSIONS

Placebo research has expanded rapidly, with increasing calls to use placebos or harness placebo effects clinically.[2 7 84] More broadly, debate on the ethical use of clinical placebos pivots on the assumption placebos/placebo effects offer genuinely salubrious benefits to patients. In the case of deceptively prescribed placebos, scholars question whether purported benefits are justified balanced against the risk of compromising clinician honesty, thereby infringing on patient autonomy over their treatment decisions. OLPs have been proposed as a potential salve for this traditionally framed ethical dilemma: the salubrious benefits of placebos, it is argued, can be harnessed honestly. However, the ethicality of either approach assumes these interventions lead to benefit, which some scholars have questioned based on lack of replication and other concerns arising from questionable research practices.[18] To address this, we identified concerns relevant to placebo research and mapped preliminary recommendations for how the field might address them. In doing so, we hope to stimulate further research that will help establish the evidence base for whether placebos do provide reliable benefits to patients, thereby informing bioethical discussion about their current and future use by clinicians.

Contributors CB, BC and CL developed the concept for the paper, discussed and identified content to be included. CB wrote the first draft with BC and CL significantly contributing to content. All authors revised and approved the final version of the manuscript. CB is the guarantor of the article.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Not applicable.

REFERENCES

1 Annoni M. Better than nothing: a historical account of placebos and placebo effects from modern to contemporary medicine. *Int Rev Neurobiol* 2020;153:3–26.
2 Evers AWM, Colloca L, Blease C, *et al*. Implications of placebo and nocebo effects for clinical practice: expert consensus. *Psychother Psychosom* 2018;87(4):204–10.
3 Colagiuri B, Schenk LA, Kessler MD, *et al*. The placebo effect: from concepts to genes. *Neuroscience* 2015;307:171–90.
4 Finniss DG, Kaptchuk TJ, Miller F, *et al*. Biological, clinical, and ethical advances of placebo effects. *The Lancet* 2010;375(9715):686–95.
5 Charlesworth JEG, Petkovic G, Kelley JM, *et al*. Effects of placebos without deception compared with no treatment: a systematic review and meta-analysis. *J Evid Based Med* 2017;10(2):97–107.
6 von Wernsdorff M, Loef M, Tuschen-Caffier B, *et al*. Effects of open-label placebos in clinical trials: a systematic review and meta-analysis. *Sci Rep* 2021;11(1):1–14.
7 Evers AWM, Colloca L, Blease C, *et al*. What should clinicians tell patients about placebo and nocebo effects? Practical considerations based on expert consensus. *Psychother Psychosom* 2021;90(1):49–56.
8 Linde K, Atmann O, Meissner K, *et al*. How often do general practitioners use placebos and non-specific interventions? systematic review and meta-analysis of surveys. *PLoS One* 2018;13(8):e0202211.
9 Faasse K, Colagiuri B. Placebos in Australian general practice: a national survey of physician use, beliefs and attitudes. *Aust J Gen Pract* 2019;48(12):876–82.
10 Bernstein MH, Locher C, Stewart-Ferrer S, *et al*. Primary care providers' use of and attitudes towards placebos: an exploratory focus group study with us physicians. *Br J Health Psychol* 2020;25(3):596–614.
11 American Medical Association. Use of placebo in clinical practice: code of medical ethics opinion 2.1.4. Delivering care, 2022. Available: https://www.ama-assn.org/delivering-care/ethics/use-placebo-clinical-practice [Accessed 27 Sep 2022].
12 Hróbjartsson A, Gøtzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* 2001;344(21):1594–602.
13 Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2004;(3).
14 Maher CG, Traeger AC, Abdel Shaheed C, *et al*. Placebos in clinical care: a suggestion beyond the evidence. *Med J Aust* 2021;215(6):252–3.
15 Kirsch I. The placebo effect revisited: lessons learned to date. *Complement Ther Med* 2013;21(2):102–4.
16 Blease CR. The role of placebos in family medicine: implications of evidence and ethics for general practitioners. *Aust J Gen Pract* 2019;48(10):700–5.
17 Hróbjartsson A, Kaptchuk TJ, Miller FG. Placebo effect studies are susceptible to response bias and to other types of biases. *J Clin Epidemiol* 2011;64(11):1223–9.
18 Blease CR, Bernstein MH, Locher C. Open-label placebo clinical trials: is it the rationale, the interaction or the pill? *BMJ Evid Based Med* 2020;25(5):159–65.
19 Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005;2(8):e124.
20 Open Science Collaboration. Psychology. Estimating the reproducibility of psychological science. *Science* 2015;349(6251):aac4716.
21 Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005;294(2):218–28.
22 Prasad VK, Cifu AS. *Ending medical reversal: improving outcomes, saving lives*. JHU Press, 2015.
23 Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Intern Med* 2011;171(18):1675–6.
24 Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979;86(3):638–41.
25 John LK, Loewenstein G, Prelec D. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychol Sci* 2012;23(5):524–32.
26 Bok S. The ethics of giving placebos. *Sci Am* 1974;231(5):17–23.
27 O'Neill O, O'Neill O. Paternalism and partial autonomy. *J Med Ethics* 1984;10(4):173–8.
28 Foddy B. A duty to deceive: placebos in clinical practice. *Am J Bioeth* 2009;9(12):4–12.
29 Blease C. Deception as treatment: the case of depression. *J Med Ethics* 2011;37(1):13–16.
30 Kaptchuk TJ. Open-label placebo: reflections on a research agenda. *Perspect Biol Med* 2018;61(3):311–34.
31 Specker Sullivan L, Sullivan LS. More than consent for ethical open-label placebo research. *J Med Ethics* 2021;47(12):e7.
32 Blease C, Colloca L, Kaptchuk TJ. Are open-label placebos ethical? Informed consent and ethical equivocations. *Bioethics* 2016;30(6):407–14.
33 Justman S. Placebo: the lie that comes true? *J Med Ethics* 2013;39(4):243–8.
34 Locher C, Buergler S, Frey Nascimento A, *et al*. Lay perspectives of the open-label placebo rationale: a qualitative study of participants in an experimental trial. *BMJ Open* 2021;11(8):e053346.
35 Annoni M, Buergler S, Stewart-Ferrer S, *et al*. Placebo studies and patient care: where are the nurses? *Front Psychiatry* 2021;12.
36 Druart L, Vauthrin O, Pinsault N, *et al*. 'It's not my greengrocer, it's someone from the medical profession': a qualitative study regarding acceptability of deceptive and open-label placebo prescribing in France. *Br J Health Psychol* 2022. doi:10.1111/bjhp.12624. [Epub ahead of print: 09 Sep 2022].
37 Annoni M. The ethics of placebo effects in clinical practice and research. *Int Rev Neurobiol* 2018;139:463–84.
38 Blease CR, Lilienfeld SO, Kelley JM. Evidence-based practice and psychological treatments: the imperatives of informed consent. *Front Psychol* 2016;7:1170.
39 O'Donohue W, Henderson D. Epistemic and ethical duties in clinical decision-making. *Behaviour Change* 1999;16(1):10–19.
40 Lilienfeld SO. Psychology's replication crisis and the grant culture: righting the SHIP. *Perspect Psychol Sci* 2017;12(4):660–4.
41 Zunhammer M, Bingel U, Wager TD, *et al*. Placebo effects on the neurologic pain signature: a meta-analysis of individual participant functional magnetic resonance imaging data. *JAMA Neurol* 2018;75(11):1321–30.
42 de Craen AJ, Moerman DE, Heisterkamp SH, *et al*. Placebo effect in the treatment of duodenal ulcer. *Br J Clin Pharmacol* 1999;48(6):853–60.
43 Meissner K, Linde K. Are blue pills better than green? how treatment features modulate placebo effects. *Int Rev Neurobiol* 2018;139:357–78.
44 de Craen AJ, Kaptchuk TJ, Tijssen JG, *et al*. Placebos and placebo effects in medicine: historical overview. *J R Soc Med* 1999;92(10):511–5.
45 Huskisson EC. Simple analgesics for arthritis. *Br Med J* 1974;4(5938):196–200.
46 Moerman DE. *Meaning, Medicine, and the" placebo Effect"*. Cambridge University Press Cambridge, 2002.
47 Świder K, Bąbel P. The effect of the type and colour of placebo stimuli on placebo effects induced by observational learning. *PLoS One* 2016;11(6):e0158363.
48 Wiercioch-Kuzianik K, Bąbel P. Color hurts. The effect of color on pain perception. *Pain Med* 2019;20(10):1955–62.
49 Buckalew LW, Coffield KE. An investigation of drug expectancy as a function of capsule color and size and preparation form. *J Clin Psychopharmacol* 1982;2(4):245???248–8.
50 Enck P, Klosterhalfen S, Weimer K. Unsolved, forgotten, and ignored features of the placebo response in medicine. *Clin Ther* 2017;39(3):458–68.
51 Ioannidis JPA. Materializing research promises: opportunities, priorities and conflicts in translational medicine. *J Transl Med* 2004;2(1):5–6.
52 Cuijpers P, Cristea IA. How to prove that your therapy is effective, even when it is not: a guideline. *Epidemiol Psychiatr Sci* 2016;25(5):428–35.
53 Tatsioni A, Bonitsis NG, Ioannidis JPA. Persistence of contradicted claims in the literature. *JAMA* 2007;298(21):2517–26.
54 Blease C, Annoni M. Overcoming disagreement: a roadmap for placebo studies. *Biol Philos* 2019;34(2).
55 Kirsch I. Placebo effect in the treatment of depression and anxiety. *Front Psychiatry* 2019;10.
56 Blease C, Kirsch I. The placebo effect and psychotherapy: implications for theory, research, and practice. *Psychol Conscious: Theory Res Pract* 2016;3(2):105–7.
57 Blease CR. Psychotherapy and placebos: manifesto for conceptual clarity. *Front Psychiatry* 2018;9.
58 Locher C, Gaab J, Blease C. When a placebo is not a placebo: problems and solutions to the gold standard in psychotherapy research. *Front Psychol* 2018;9:2317.
59 Colagiuri B, Smith CA. A systematic review of the effect of expectancy on treatment responses to acupuncture. *Evid Based Complement Alternat Med* 2012;2012:1–12.
60 Davies JN, Sharpe L, Day MA, *et al*. How do placebo effects contribute to mindfulness-based analgesia? Probing acute pain effects and interactions using a randomized balanced placebo design. *Pain* 2022;163(10):1967–77.
61 Jopling DA. *Talking cures and placebo effects*. Oxford University Press, 2008.
62 Wampold BE, Imel ZE. *The great psychotherapy debate: the evidence for what makes psychotherapy work*. Routledge, 2015.
63 Kirsch I, Wampold B, Kelley JM. Controlling for the placebo effect in psychotherapy: noble quest or tilting at windmills? *Psychol Conscious: Theory Res Pract* 2016;3(2):121–31.
64 Locher C, Hasler S, Gaab J. When do placebos in psychotherapeutic research work? A systematic review on the example of systematic desensitization. *Verhaltenstherapie* 2016;26.
65 Colagiuri B, Lovibond PF. Psychological processes that can bias responses to placebo treatment for pain. In: *Placebo and pain*, 2013: 175–82.
66 Howick J, Hoffmann T. How placebo characteristics can influence estimates of intervention effects in trials. *CMAJ* 2018;190(30):E908–11.
67 Furukawa TA, Noma H, Caldwell DM, *et al*. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatr Scand* 2014;130(3):181–92.
68 Rosenthal R. *Experimenter effects in behavioral research*, 1976.
69 Cochrane Methods. Rob 2: a revised Cochrane risk-of-bias tool for randomized trials, 2022. Available: https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials [Accessed 27 Sep 2022].
70 Colagiuri B, Sharpe L, Scott A. The blind leading the not-so-blind: a meta-analysis of blinding in pharmacological trials for chronic pain. *J Pain* 2019;20(5):489–500.
71 Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 2016.
72 Head ML, Holman L, Lanfear R, *et al*. The extent and consequences of p-hacking in science. *PLoS Biol* 2015;13(3):e1002106.
73 Lenth RV. Some practical guidelines for effective sample size determination. *Am Stat* 2001;55(3):187–93.
74 Cohen J. *Statistical power analysis for the behavioral sciences*. Routledge, 2013.
75 Halpern SD, Karlawish JHT, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *JAMA* 2002;288(3):358–62.

76 Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;53(11):1119–29.
77 Button KS, Ioannidis JPA, Mokrysz C, *et al*. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013;14(5):365–76.
78 Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci* 2010;33(2-3):61–83.
79 Howe LC, Hardebeck EJ, Eberhardt JL, *et al*. White patients' physical responses to healthcare treatments are influenced by provider race and gender. *Proc Natl Acad Sci U S A* 2022;119(27):e2007717119.
80 Okusogu C, Wang Y, Akintola T, *et al*. Placebo hypoalgesia: racial differences. *Pain* 2020;161(8):161.
81 Letzen JE, Dildine TC, Mun CJ, *et al*. Ethnic differences in experimental pain responses following a paired verbal suggestion with saline infusion: a quasiexperimental study. *Ann Behav Med* 2021;55(1):55–64.
82 Lembo A, Kelley JM, Nee J, *et al*. Open-label placebo vs double-blind placebo for irritable bowel syndrome: a randomized clinical trial. *Pain* 2021;162(9):2428–35.
83 Rooney T, Sharpe L, Todd J, *et al*. The relationship between expectancy, anxiety, and the nocebo effect: a systematic review and meta-analysis with recommendations for future research. *Health Psychol Rev* 2022:1–28.
84 Mitsikostas DD, Blease C, Carlino E, *et al*. European headache federation recommendations for placebo and nocebo terminology. *J Headache Pain* 2020;21(1):1–7.