DOI:10.5281/zenodo.3886579; CC BY 4.0

Nonplussed by placebos: A response to Colloca and Barsky (2020)

June 06, 2020

Darren Dahly. Principal Statistician, HRB Clinical Research Facility - Cork. School of Public Health, University College Cork.

Zad Rafi, Statistical Analyst, Department of Population Health, NYU Langone Health.

In a recent review published in the New England Journal of Medicine (*Placebo and Nocebo Effects)*, Colloca and Barsky[1] concluded that "Placebo and nocebo effects are powerful, pervasive, and common in clinical practice" and that "Strategies to promote placebo effects and to prevent nocebo effects can improve therapeutic outcomes…" The review cites a substantial amount of impressive neurological, mechanistic research supporting this assertion, much of it led by the authors. However, the review also included a number of problematic statements about the design and interpretation of clinical trials that substantially undercut the authors' conclusions.

One problem was the authors' assertion that evidence for therapeutic placebo effects could be observed by comparing outcome improvements *within* placebo arms to those seen *within* treatment arms (*"In many double-blind clinical trials of treatments for pain or psychiatric disorders, for example, the responses to placebo are similar to the responses to active treatment"*). The fault in this reasoning stems from a phenomenon called regression towards the mean[2–4], which has been well understood for over a century[5]. It results from the fact that we tend to recruit patients into trials when they are at their worst, experiencing relatively extreme values for one or more outcomes (e.g. pain). For many illnesses, we might reasonably expect such extreme outcomes to resolve without treatment (even if only to return to problematic levels later on), so that a group of such patients, if followed up over time, will appear to have improved on average, even in the absence of an effective intervention. This will be especially true when the outcomes are subjective and measured with error. Thus if we were to intervene in such a group of patients, it would be a mistake to interpret this improvement in outcomes as the effect of our intervention (for further explanation see Senn[4]). This is one of the main reasons trialists employ concurrent control arms, and then limit inferences about treatment effects to the *differences observed between them*, since they will both be impacted by regression to the mean, but only one of them can be impacted by the intervention being tested. Then, if we don't observe a difference in outcomes between the groups, we typically conclude that the intervention didn't work, not that the placebo worked just as well as the treatment, as the Colloca and Barsky have done here. Statisticians have attempted to quantify how much of the so-called placebo effect seen in clinical trials can be explained by regression to the mean, and found that older clinical trials with design deficiencies were more likely to show improvements in patients taking placebos, whereas newer studies often showed no such improvement[2,6]. They also found

that regression to the mean can yield notable improvements even in serial biochemical measurements.

We also object to the authors' argument that placebos are likely to exert therapeutically relevant effects since some studies report a substantial prevalence of "side-effects" in placebo arms (*"...up to 19% of adults and 26% of elderly persons taking placebos report side effects. Furthermore, as many as one quarter of patients receiving placebo in clinical trials discontinue it because of side effects, suggesting that a nocebo effect may contribute to discontinuation of or lack of adherence to active treatments."*). Importantly, two of the three reviews cited to support this argument[7,8] report on *adverse events*, which is a technical term that specifically refers to any and all adverse events recorded during the course of a trial without consideration of any causal link to a treatment, placebo or otherwise[9]. For example, a patient enrolled in a drug trial who fell down some stairs resulting in injury would likely be recorded as having experienced an adverse event. This is in contrast to the term *side-effects*, which would surely suggest to many or most readers that the effect was the *result* of a treatment. Making this leap from adverse events to side-effects (or more formally, *adverse effects or adverse reactions)* requires expert adjudication, so it is concerning to see the authors shortcut this important step. The third review they cite[10] presents data on adverse effects that were "possibly drug related" (including placebo) but these were largely compiled from Physicians' Desk References (known for frequently citing outdated data from questionable sources), and it doesn't appear possible to determine whether the placebo data actually refers to study arms where patients received standard care plus placebo (vs placebo alone). We would also like to point out a recent review, uncited by the authors, where adverse events rates were compared in placebo vs no treatment arms (i.e. where the patients did not even receive a placebo), finding that the rates were similar between the two groups[11], further illustrating the nature of adverse event reporting in clinical trials.

Perhaps the most problematic aspect of Colloca and Barsky's review is their citation of a large review by Hróbjartsson and Gøtzsche[12] that identified studies that included both placebo and no treatment arms. They cited this study to support their view that clinical trials should do exactly this - to include arms with no treatment - so that we could better understand placebo effects. What they omitted however were the conclusions of this review, which were that the appearance of placebo effects were rare, that placebo effects were small when they were reported, and that they were limited to subjective outcomes and smaller studies. This of course is completely consistent with regression towards the mean and measurement error as a major drivers of so-called placebo effects, and completely contradicts Colloca and Barsky's own conclusions, which was that placebo effects were "powerful" and "pervasive". Given that the point of a scientific literature review is to provide a complete picture of existing research on a given topic, we find this omission rather astonishing, particularly since there is no way to argue they weren't aware this important, contradictory evidence existed, given that they cited it. In a personal communication, Professor Colloca suggested that the editor requested that the review focus on mechanistic research[13] (which we interpreted as a request to ignore applied therapeutic research) and that placebo effects have been "proven" to be different in lab studies vs human RCTs[14]. The former statement gives pause, given that a substantial portion of the review was on the "implications for placebo and nocebo effects in research and clinical practice"; while the

latter one stands out since there are of course other plausible, though perhaps less compelling, explanations for effects that can be demonstrated in the lab but not in clinical trials.

It is for these reasons that we strongly disagree with the authors' conclusion that "Strategies to promote placebo effects and to prevent nocebo effects can improve therapeutic outcomes and minimize the unintended exacerbation of symptoms in clinical practice and clinical trials." It is hard to imagine a stronger conclusion, even if the quality of the evidence underpinning it was pristine. Instead, the evidence presented seems to support the plausibility of therapeutically useful placebo effects based on indirect evidence (an argument placebo researchers have been making for decades now), and some limited empirical evidence of them (systematic effects beyond regression to the mean and measurement error) in a very limited number of areas. To be clear we have no problem accepting that the nature of interactions between clinicians and patients can be important for improving outcomes. Nor do we diminish the importance of pain as an outcome for patients, or the value of even small reductions in pain (or other important, patient-reported outcomes). However, to suggest that placebos have potent, widely-applicable therapeutic effects is clearly a substantial overreach, and in our opinion the suggestion that we should explore subgroups of patients who are most likely to benefit from such therapeutic effects will result in much more noise than signal.

Citations

1. Colloca, L. & Barsky, A. J. Placebo and Nocebo Effects. *N. Engl. J. Med.* **382**, 554–561 (2020).
2. McDonald, C. J., Mazzuca, S. A. & McCabe, G. P. How much of the placebo 'effect' is really statistical regression? *Stat. Med.* **2**, 417–427 (1983).
3. Bland, J. M. & Altman, D. G. Statistic Notes: Regression towards the mean. *BMJ* **308**, 1499–1499 (1994).
4. Senn, S. Francis Galton and regression to the mean. *Significance* **8**, 124–126 (2011).
5. Galton, F. Regression Towards Mediocrity in Hereditary Stature. *J. Anthropol. Inst. G. B. Irel.* **15**, 246 (1886).
6. Senn, S. J. How much of the placebo 'effect' is really statistical regression? *Stat. Med.* **7**, 1203–1203 (1988).
7. Häuser, W., Sarzi-Puttini, P., Tölle, T. R. & Wolfe, F. Placebo and nocebo responses in

randomised controlled trials of drugs applying for approval for fibromyalgia syndrome treatment: systematic review and meta-analysis. *Clin. Exp. Rheumatol.* **30**, 78–87 (2012).

8. Rosenzweig, P., Brohier, S. & Zipfel, A. The placebo effect in healthy volunteers: influence of experimental conditions on the adverse events profile during phase I studies. *Clin. Pharmacol. Ther.* **54**, 578–583 (1993).
9. ICH E2A Clinical safety data management: definitions and standards for expedited reporting.
10. Rief, W., Avorn, J. & Barsky, A. J. Medication-Attributed Adverse Effects in Placebo Groups: Implications for Assessment of Adverse Effects. *Arch. Intern. Med.* **166**, 155 (2006).
11. Howick, J., Webster, R., Kirby, N. & Hood, K. Rapid overview of systematic reviews of nocebo effects reported by patients taking placebos in clinical trials. *Trials* **19**, 674 (2018).
12. Hróbjartsson, A. & Gøtzsche, P. C. Is the Placebo Powerless?: An Analysis of Clinical Trials Comparing Placebo with No Treatment. *N. Engl. J. Med.* **344**, 1594–1602 (2001).
13. https://twitter.com/Colloca_Luana/status/1228775987380473856.
14. https://twitter.com/Colloca_Luana/status/1228777003660300293.

Additional reading

15. Rafi, Zad. We May Not Understand Control Groups. *LessLikely* https://lesslikely.com/statistics/control-group-effects/ (2019).
16. Barrowman, Nick. The Myth of the Placebo Effect. *The New Atlantis* https://www.thenewatlantis.com/publications/the-myth-of-the-placebo-effect (2014).
17. Colquhoun, David. Placebo effects are weak: regression to the mean is the main reason ineffective treatments appear to work. *DC's Improbable Science*

http://www.dcscience.net/2015/12/11/placebo-effects-are-weak-regression-to-the-mean-is-the-main-reason-ineffective-treatments-appear-to-work/ (2015).