Title

Where are randomized trials necessary: Are smoking and parachutes good counterexamples?

Journal

European Journal of Clinical Investigation, 52(5)

ISSN

0014-2972

Authors

Powell, Kerrington
Prasad, Vinay

Publication Date

2022-05-01

DOI

10.1111/eci.13730

Received: 13 November 2021 | Revised: 7 December 2021 | Accepted: 9 December 2021

DOI: 10.1111/eci.13730

COMMENTARY

WILEY

Where are randomized trials necessary: Are smoking and parachutes good counterexamples?

Kerrington Powell[1] | **Vinay Prasad**[2]

[1]College of Medicine, Texas A&M Health Science Center, Bryan, Texas, USA

[2]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

Correspondence
Vinay Prasad, Department of Epidemiology and Biostatistics, University of California San Francisco, 550 16th St, 2nd Fl, San Francisco, CA 94158, USA.
Email: vinayak.prasad@ucsf.edu

Funding information
This study was funded by Arnold Ventures.

Keywords: epidemiology, evidence-based medicine, health policy, randomized control trials

As we consider the appropriate role for randomized control trials (RCTs) in biomedicine, it is important to clarify that RCTs are generally conducted for interventions thought to be beneficial, but at best, offer effect sizes that are modest or marginal. These two prerequisites are present among nearly all RCTs and help clarify when RCTs are essential and when they are unnecessary.

Consider a thought experiment. Imagine a spectrum of benefit to harm that we have depicted in Figure 1. From left to right, we move from the worst thing that a person can experience (an intervention that is likely fatal) to the best thing a person can experience (an intervention that is almost always lifesaving).

On the leftmost edge, we see, for example, a gunshot wound (GSW) to the heart has a near 100% fatality probability, and on the rightmost, we see that pushing someone out of the path of a speeding vehicle has a near 100% success rate. In the centre is a line of neutrality, which encompasses all of the things we do in life that are insignificant.

Things on the extreme left of the graph are considered universally fatal despite never having been tested in a randomized manner nor examined in epidemiological research, since their harms are so self-evident. Things on the right side of the graph are interventions with dramatic benefit, to the point that the necessity for randomization is nonsensical owing to their obviousness. For example, there are no randomized trials of pulling someone out of the way of a speeding vehicle, since its benefit is unquestionable and does not need sophisticated statistical analysis to demonstrate the effect size.

Now, consider interventions of more modest effect size. These are certain acts whose consequences or benefits are not immediately apparent. Smoking is an illustration of this, which necessitated epidemiological research to reveal a 20:1 risk ratio for smokers getting lung cancer, even when cigarettes were largely regarded as innocuous at the time.[1] Because these impact sizes are smaller than GSWs, we often conduct risk factor epidemiology to identify the consistency and magnitude of risk. This is true not just for smoking, but also for vanillin chloride compounds associated with bladder cancer, talcum powder associated with ovarian cancer, and various deficiencies such as poor nutritional exposure—although some of these associations remain hotly contested.[2]

When one discusses the beneficial (right) side of the Figure 1, which often includes medical treatments, a problem emerges. Too often, medical treatments are compared to those on the extreme right of the graph, such as parachutes, an intervention with a 99.99% absolute risk reduction in all-cause mortality.[3] This analogy originates from Smith and Pell's 2003 satirical article in the *The BMJ's* Christmas edition, in which they performed a systematic review of randomized trials evaluating the parachute, of which none existed at the time.[4] The primary takeaway

was a critique of the most ardent proponents of evidence-based medicine, an approach that requires robust randomized evidence before the adoption of medical therapies, and boils down to the fact that certain medical interventions with clear benefit are analogous to parachutes and do not require randomized evidence prior to adoption.

However, this metaphor overlooks the reality that the majority of medical treatments have a modest to minimal risk reduction in harm, making the contrast to a parachute a straw-man comparator. To demonstrate this, we examined all articles referencing the original 2003 *BMJ* article claiming that randomized trials are unnecessary for treatments with obvious benefit. Among these 822 articles, the greatest absolute risk reduction (ARR) was 30.8% in reported findings,[5] which is an unrepresentative comparison to the 99.99% ARR of a parachute.

Additional evidence supporting this claim comes from Pereira and colleagues, who found only one intervention among 80,000 practices consistently had a large effect (defined as an odds ratio of ≥5) on mortality in their search of Cochrane reviews, which was a 40% reduction in the risk of death associated with extracorporeal oxygenation for severe neonatal respiratory failure.[6] Although these risk differences of up to 40% are massive, a 99% absolute risk difference has yet to be discovered in medicine, tempering the parachute analogy and bolstering the need for randomized evidence.

In a response paper to the original *BMJ* article, Yeh and colleagues randomly assigned individuals to jump with or without a parachute–although from a height of just 0.6 metres.[7] The study discovered that use of a parachute had no significant effect on mortality or serious injury, however, the message of the paper is more subtle. A commonly cited reason for the failure of many pivotal trials is that researchers were reluctant to randomize the sickest patients, and had they not been, the outcome may have been different. The connection to their parachute RCT is that when randomization is performed, researchers are not willing to randomize from the highest heights, but only 2 feet above the ground. As the authors emphasize in their accompanying opinion article, they favour randomized trials; but, since randomization may exclude the sickest patients as providers deem it unethical to randomize them, a negative study may not categorically prove that the intervention is ineffective.[8] In other words, the results of these low altitude RCTs would lack external validity, rendering them inapplicable to real-world patient populations. Three rebuttals to this reasoning would be as follows: (1) We are still unsure if the intervention benefits patients. (2) Positive studies are still required. (3) This is a problem with the design and conduct of randomized trials, not with the principle of randomization itself.

This leads us to the question: when are randomized trials necessary? RCTs gained attraction when evident pathophysiology, logical mechanisms of action, and a community belief that an intervention was likely to succeed were turned on its head by rigorous trial methodology.[9] These were times when empiricism triumphed over rationalism and can be attributed to randomization's elimination of issues with confounding, immortal time bias and multiplicity. Because of these advantages, in addition to superiority studies, RCTs are often used in noninferiority and safety trials to examine therapeutic toxicity reduction, foster market competition and develop alternative treatment options. However, RCTs do have limitations, which are usually related to trial design rather than the approach itself (e.g., randomization). If not addressed, issues with faulty comparators, insufficient crossover, improper drug dosage and high noninferiority margins may result in misleading findings.[10]

Even with their limitations, RCTs remain the gold standard of evidence; they are needed to answer the critical question: does this intervention work *under some circumstances?* Without RCTs, it may be difficult to distinguish between ambitious thinking and real effects. However, is a randomized trial necessary for *every* intervention? Glasziou and colleagues tackled this question by developing a model to assess when well-designed observational evidence for treatments is sufficient enough to eliminate the need for randomized trials, such as insulin for diabetes or liver transplantation in end-stage liver disease.[11] While this may be

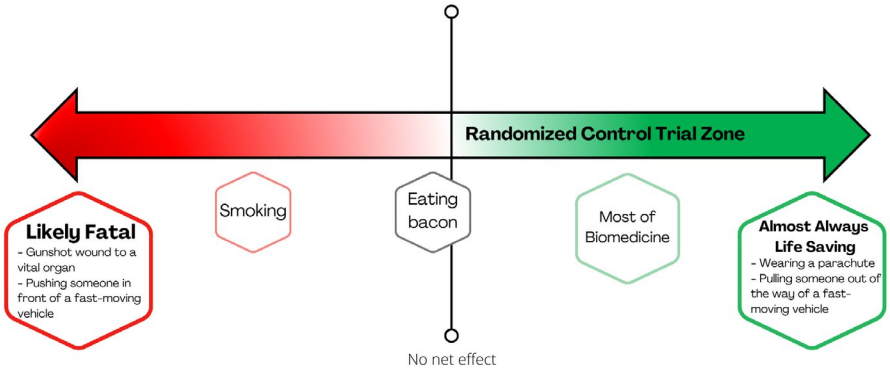

FIGURE 1 Schematic depicting when randomized control trials are necessary

true for specific medical interventions, this approach does not fully eliminate bias, which is why the evidence should be evaluated in aggregate, and Glasziou et al.'s signal-to-noise model used as a tool and not a replacement. Most treatments, ideally, would have randomized data generated in conjunction with observational data, however, others claim that randomization is not feasible for many interventions owing to a lack of equipoise or trial design infeasibility. The issue with the latter point is that *most* interventions are eligible for randomization. Not long ago, many predicted that randomized studies of appendectomy vs high-dose antibiotics for appendicitis would never be conducted. And, although there is considerable leeway in interpreting the findings, there is no doubt that we have conducted at least four such randomized studies.[12]

Often when therapies are accepted without randomization, they are later shown to be in error.[13] One instance is the use of hormone therapy (HT) to reduce the risk of cardiovascular events in postmenopausal women, a bioplausible intervention that was supported by preclinical science as well as cohort data from the Nurses' Health Study.[14] This large public health study was the key impetus for the adoption of postmenopausal HT throughout the nation, as well as its ensuing US Preventive Services Task Force (USPSTF) grade B classification.[15] Concerns regarding HT arose after the end the Heart and Estrogen/progestin Replacement Study (HERS), which found an elevated risk of heart disease among women with a history of the condition.[16] Because of the HERS' slightly different patient cohort, it wasn't until the Women's Health Initiative (WHI), a randomized trial comparing HT to placebo, that we understood how harmful postmenopausal HT may be. The study's findings indicated that postmenopausal women who received hormone therapy were developing maladies at a higher rate (including cardiovascular disease) than those who received a placebo, suggesting that the risks of the therapy outweighed the benefit.[17] Not only did this error affect people's health, but also their confidence and trust in the healthcare system. When women are actively marketed to in accordance with national healthcare guidelines, and these recommendations turn out to be not only ineffective, but also harmful, a significant loss of faith in the healthcare system occurs. The lawsuits surrounding Wyeth Pharmaceutical's hormone therapies, Premarin and Prempro, are emblematic of this issue, as the corporation failed to adequately disclose the dangers associated with its medicines, resulting in patient harm, a loss of faith in physicians prescribing them and concern about national healthcare recommendations supporting their use.[18] Elsewhere, we have detailed hundreds of instances of medical reversal and its associated harms in recent decades, reversals that may have been averted had rigorous randomized data been cultivated prior to widespread adoption.[9,13,19,20]

Why, then, do people often assert that RCTs are unfeasible, despite the danger associated with adopting therapies without randomized data? How is it appropriate to use smoking and parachutes as counterexamples? Medicine cannot be likened to a parachute; our patients do not leap from planes, our treatments are not as successful as hitting-the-silk while falling from the sky, and demonstrating effectiveness in a patient group is much more complex than pulling a rip cord. As a result, almost everything in biomedicine can be randomized; in most cases, there is equipoise. A persistent bias in biomedicine is that more expensive, intrusive and novel treatments *must* improve outcomes; nevertheless, the only way to remove this bias is to confront it using empiricism. Ignoring evidence and relying on heuristics and personal judgement in the face of empiricism may result in a loss of credibility, a stalling of innovation and a loss of public confidence in our medical initiatives.

CONFLICT OF INTEREST

Vinayak Prasad's Disclosures. (Research funding) Arnold Ventures (Royalties) Johns Hopkins Press, Medscape, MedPage (Consulting) UnitedHealthcare. (Speaking fees) Evicore. New Century Health (Other) Plenary Session podcast has Patreon backers. All other authors have no financial or nonfinancial conflicts of interest to report.

AUTHOR CONTRIBUTIONS

VP conceptualized study design; KP reviewed literature; VP reviewed and confirmed abstracted data; KP wrote the first draft of the manuscript; and all authors reviewed and revised subsequent and finalized draft of the manuscript.

ORCID

Kerrington Powell https://orcid.org/0000-0001-7067-3559
Vinay Prasad https://orcid.org/0000-0002-6110-8221

REFERENCES

1. Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. Featured Executive Branch Publications 1964; Public Health Service Publication No. 1103.
2. Schoenfeld JD, Ioannidis JP. Is everything we eat associated with cancer? A systematic cookbook review. *Am J Clin Nutr*. 2013;97(1):127-134. doi:10.3945/ajcn.112.047142
3. Ellitsgaard N. Parachuting injuries: a study of 110,000 sports jumps. *Br J Sports Med*. 1987;21:13-17.
4. Smith GCS, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ*. 2003;327:1459. doi:10.1136/bmj.327.7429.1459

5. Hayes MJ, Kaestner V, Mailankody S, Prasad V. Most medical practices are not parachutes: a citation analysis of practices felt by biomedical authors to be analogous to parachutes. *CMAJ Open*. 2018;6(1):E31-E38. doi:10.9778/cmajo.20170088
6. Pereira TV, Horwitz RI, Ioannidis JPA. Empirical evaluation of very large treatment effects of medical interventions. *JAMA*. 2012;308(16):1676-1684. doi:10.1001/jama.2012.13444
7. Yeh RW, Valsdottir LR, Yeh MW, et al. Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. *BMJ*. 2018;363:k5094. doi:10.1136/bmj.k5094
8. Yeh RW, Kasi D, Valsdottir LR, Nallamothu BK. We jumped from planes without parachutes (and lived to tell the tale). *The BMJ*. Accessed October 17, 2021. https://blogs.bmj.com/bmj/2018/12/13/we-jumped-from-planes-without-parachutes-and-lived-to-tell-the-tale/ Published December 14, 2018.
9. Prasad V, Vandross A, Toomey C, et al. A decade of reversal: an analysis of 146 contradicted medical practices. *Mayo Clin Proc*. 2013;88(8):790-798. doi:10.1016/j.mayocp.2013.05.012
10. Haslam A, Prasad V. When is crossover desirable in cancer drug trials and when is it problematic? *Ann Oncol*. 2018;29(5):1079-1081. doi:10.1093/annonc/mdy116
11. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. *BMJ*. 2007;334:349. doi:10.1136/bmj.39070.527986.68
12. Varadhan KK, Neal KR, Lobo DN. Safety and efficacy of antibiotics compared with appendicectomy for treatment of uncomplicated acute appendicitis: meta-analysis of randomised controlled trials. *BMJ*. 2012;344:e2156. doi:10.1136/bmj.e2156
13. Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Intern Med*. 2011;171(18):1675-1676. doi:10.1001/archinternmed.2011.295
14. Stampfer MJ, Colditz GA, Willett WC, et al. Postmenopausal estrogen therapy and car-diovascular disease: ten-year follow-up from the nurses' health study. *N Engl J Med*. 1999;325:756-762.
15. US Preventive Services Task Force. *Guide to Clinical Preventive Services: Report of the U.S. Preventive Services Task Force*. 2nd ed. Williams & Wilkins; 1996. 68, Postmenopausal Hormone Prophylaxis. Available from: https://www.ncbi.nlm.nih.gov/books/NBK15462/
16. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for second- ary prevention of coronary heart disease in postmenopausal women. Heart and Estro- gen/progestin Replacement Study (HERS) Research Group. *JAMA*. 1998;280:605-613.
17. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the women's health initiative randomized controlled trial. *JAMA*. 2002;288(3):321-333. doi:10.1001/jama.288.3.321
18. Press TA. Jury awards $134.5 million in Wyeth Hormone Drug Case (published 2007). The New York Times. Accessed December 6, 2021. https://www.nytimes.com/2007/10/12/business/12wyeth.html Published October 12, 2007
19. Herrera-Perez D, Haslam A, Crain T, et al. A comprehensive review of randomized clinical trials in three medical journals reveals 396 medical reversals. *Elife*. 2019;8:e45183. Published 2019 Jun 11. doi:10.7554/eLife.45183
20. Haslam A, Gill J, Crain T, et al. The frequency of medical reversals in a cross-sectional analysis of high-impact oncology journals, 2009-2018. *BMC Cancer*. 2021;21(1):889. Published 2021 Aug 3. doi:10.1186/s12885-021-08632-8

How to cite this article: Powell K, Prasad V. Where are randomized trials necessary: Are smoking and parachutes good counterexamples? *Eur J Clin Invest*. 2022;52:e13730. doi:10.1111/eci.13730