

What does randomisation achieve?

Adam La Caze,¹ Benjamin Djulbegovic,² Stephen Senn³

10.1136/ebm.2011.100061

¹School of Pharmacy, The University of Queensland, St Lucia, Queensland, Australia

²Center for Evidence-based Medicine & Health Outcome Research, University of South Florida & H Lee Moffitt Cancer Center & Research Institute, Tampa, Florida, USA

³School of Mathematics and Statistics, University of Glasgow, Glasgow, UK

Correspondence to

Adam La Caze

School of Pharmacy, The University of Queensland, St Lucia, QLD 4072, Australia; a.lacaze@uq.edu.au

What are the benefits of random allocation in clinical studies? John Worrall, a philosopher of science, recently questioned whether evidence-based medicine's advice to base therapeutic decisions on the results of randomised controlled trials (RCTs) could be justified.^{1,2} Here we provide a response to Worrall and others who challenge the epistemological value of RCTs. Worrall's primary target is the view that RCTs are the only reliable source of evidence for medicine. But in arguing against this strong view, he posits a similarly strong counterposition. Worrall argues that randomisation offers no advantage over balanced systematic designs in which experimental and control groups are carefully matched according to known confounders.

The best we can do (as ever) is test our theories against rivals that seem plausible in the light of background knowledge. Once we have eliminated other explanations that we know are possible (by suitable deliberate, or *post hoc*, control) we have done as much as we can epistemologically.²

Focusing on this claim, we first discuss Worrall's arguments and then provide reasons to reject this view. There are good reasons to randomise studies of therapeutic interventions; principally, RCTs have the capacity to avoid a form of selection bias that cannot be avoided in observational studies.

Worrall draws the conclusion that randomised trials provide no epistemological benefit from two key arguments:

- 1 He criticises the notion promoted in the literature that random allocation controls for known and unknown confounders at baseline. Since the number of unknown (possible) causes are 'innumerable', "it would clearly be a miracle if all of those factors just happened to be balanced in the two groups. . . on a single random division."² Indeed, the probability that any single confounder is unevenly distributed in a given RCT ranges from zero to one.³ Hence, it is possible that some unknown confounder is distributed unevenly between treatment and control and correlated to the effect of intervention that was tested in the RCT.
- 2 Randomisation is not essential for controlling selection bias or concealing allocation. Worrall and others⁴ maintain that these sources of bias can be equally well eliminated in studies that are not randomised.

Neither '1' nor '2', however, undermines the epistemological benefits of RCTs over observational studies in testing the efficacy of therapeutic interventions. Worrall's first argument is only relevant if the point he makes is not understood (however, '1' is well understood by trial analysts) and '2' suffers from serious problems.

Worrall's '1', while not always explicitly discussed, is nonetheless well recognised in the statistical and epidemiological literature. The crucial point is that it is not necessary that randomisation control for all known and unknown confounders at baseline to make valid statistical

inferences. It is not necessary for the purpose of issuing a probability statement to know the value of unmeasured covariates; it is sufficient to know their distribution in probability, which randomisation is designed to provide.⁵ Similar knowledge regarding unmeasured covariates is not possible in observational designs. Analysis of observational studies rely on additional assumptions which cannot be easily verified. In short, Worrall's '1' does not undermine standard frequentist statistical tests.

Another epistemological benefit of RCTs over observational studies is the capacity of RCTs to avoid confounding by indication (choice of treatment bias). Confounding by indication, in which treatment assignment is a function of the risk of future health outcomes (prognosis), is a particular problem for observational studies. Even if we match for a number of known factors, it is difficult in observational designs to rule out or account for all the factors which may influence the physicians' treatment assignment or the patients' reasons to accept or decline the intervention (see, for instance, Collins and MacMahon⁶). RCTs are prone to other biases but not this one.

Eliminating confounding by indication is a strong argument for interventional studies over observational studies. Worrall's '2' claims that it is possible to use alternatives to randomisation to ensure experimental groups are equally balanced for known confounders and conceal allocation. Worrall does not provide details on these alternatives. Presumably, he is focusing on interventional studies and expressing the idea that it is conceptually possible to devise a method (other than randomisation) that stratifies patients according to relevant prognostic information and allocates them to treatment or control that is independent of the investigator. A number of problems can be raised against this suggestion.

First, alternative methods of allocation are not possible for case-control or cohort studies – in these study designs, the patient decides to (or the patient's circumstances lead them to) take or not take the intervention under investigation. Worrall and Borgerson fail to give sufficient weight to the importance of confounding by indication as a justification for interventional studies (and RCTs in particular) over observational studies.

Second, where randomisation is possible and easily carried out, methods that claim to do better than randomisation should be critically examined. Randomising provides a distinct (and uncontroversial) experimental distribution on which to base statistical inferences; this is the case for frequentist^{7,8} and Bayesian^{9,10} approaches to statistical inference. Alternative methods will require a more complicated statistical model and additional assumptions of the data. This makes the appropriate choice of statistical model and analysis more cumbersome and typically more difficult to defend. These problems vitiate the proposed advantages of a non-random methods of allocation in interventional studies.

It is important to note that none of the previous discussion invalidates the importance of observational

studies. RCTs are neither necessary nor sufficient to conclusively establish therapeutic claims: not necessary because alternative methods may be appropriate when the effect size is large relative to bias or random error that may possibly obscure it,¹¹ and not sufficient because there is more to assessing a therapeutic claim than success in a randomised trial (eg, assessing if the results from a RCT are generalisable). Furthermore, observational studies may be the best available method for particular questions (eg, assessing rare adverse effects of a medication). The question we are addressing here is: do RCTs offer unique epistemological value above and beyond observational studies?

It should be no surprise that a positive result from a well-conducted RCT falls short of providing conclusive proof of the efficacy of an intervention. A number of assumptions need to hold for a causal conclusion to be valid. Different accounts of causation and different approaches to statistical inference spell out the necessary assumptions in different ways. Nancy Cartwright¹² and Dan Steel¹³ discuss the assumptions that need to hold on two probabilistic accounts of causation. Senn,⁸ with a focus on clinical trials, discusses assumptions within the frequentist approach. The need for these assumptions does not undermine the importance of RCTs for assessing the efficacy of medical interventions.

RCTs offer unique epistemological advantages that cannot be realised via observational studies. Neither does this mean that the observational studies are unimportant nor does this mean that RCTs are the best method for all questions and in all circumstances (as is reflected in the Oxford Centre for Evidence-based Medicine's Levels of Evidence).¹⁴ Rather, (therapeutic) clinical research should be understood as the means to respond to uncertainties about treatment effects of competing interventions, and the design of clinical study (observational vs RCTs) should be matched to the question/uncertainties at hand.¹⁵

Acknowledgements This article began as a discussion on the Evidence-Based Health email discussion list initiated by BD. The authors would like to acknowledge the

members of the EBM discussion group, whose opinion and free sharing of ideas has informed this work.

Competing interests ALC and BD declare no competing interests. SS notes that he is a consultant to the pharmaceutical industry and an academic whose career is furthered by publishing.

References

1. Worrall J. What evidence in evidence-based medicine? *Philos Sci* 2002;69:S316–30.
2. Worrall J. Why there's no cause to randomize. *Br J Philos Sci* 2007;58:451–88.
3. Howson C, Urbach P. *Scientific Reasoning: The Bayesian Approach*. 3rd edition. Chicago, IL: Open Court 2006:195–6.
4. Borgerson K. Valuing evidence: bias and the evidence hierarchy of evidence-based medicine. *Perspect Biol Med* 2009;52:218–33.
5. Senn S. Baseline balance and valid statistical analyses: common misunderstandings. *Appl Clin Trials* 2005;14:24–7.
6. Collins R, MacMahon S. Reliable assessment of the effects of treatments on mortality and major morbidity. In: Rothwell PM, ed. *Treating Individuals: From Randomised Trials to Personalised Medicine*. Philadelphia, PA: Elsevier 2007.
7. Senn S. Fisher's game with the devil. *Stat Med* 1994;13:217–30.
8. Senn S. Controversies concerning randomization and additivity in clinical trials. *Stat Med* 2004;23:3729–53.
9. Lindley DV. The role of randomization in inference. Proceedings of the Biennial Meeting of the Philosophy of Science Association, 30 October. 1982:431–46.
10. Berry SM, Kadane JB. Optimal Bayesian randomization. *J Roy Stat Soc B* 1997;59:813–19.
11. Glasziou P, Chalmers I, Rawlins M, et al. When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007;334:349–51.
12. Cartwright N. What are randomised controlled trials good for? *Philos Stud* 2010;147:59–70.
13. Steel D. Causal inference and medical experiments. In: Gifford F, ed. *Handbook of the Philosophy of Science Volume: Philosophy of Medicine*. Vol 16. Elsevier: North Holland 2011:1–26.
14. OCEBM Levels of Evidence Working Group. The Oxford 2011 Levels of Evidence. 2011. Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653> (accessed May 2011).
15. Djulbegovic B. Articulating and responding to uncertainties in clinical research. *J Med Philos* 2007;32:79–98.