The biases and limits of randomised controlled trials in medicine

Alexander Krauss
London School of Economics
University College London

Evidence from randomised controlled trials (RCTs) is generally thought of as the most reliable form of evidence within the medical sciences. The RCT method is viewed by most medical researchers as having revolutionised the medical sciences in the second half of the 20th century, and is often viewed as the only means to generate rigorous 'causal' knowledge about what works and what does not work. Medical researchers also often think of RCTs as being free from overly complex theory and strong methodological assumptions and biases that unavoidably affect other methods.

To better understand this leading scientific method and its limits, philosophers of science have been increasingly studying the theory behind RCTs. But philosophers largely use abstract reasoning to discuss issues related to randomisation, statistical probabilities, deductive reasoning, ethical implications and the like. Worrall (2007; 2007a) for example assesses the function and limitation of randomisation and outlines some of its ethical constraints. Cartwright (2007; 2010) reasons about statistical probabilities, external validity and the conditions under which causal conclusions follow deductively in the ideal RCT. Clarke, Gillies, Illari, Russo and Williamson (2014) argue for a rethinking about hierarchies of evidence and the position of RCTs within them, and for a greater focus on mechanisms and not just correlations as evidence for causal claims. These leading philosophers, while making important contributions in improving our understanding about the RCT method, largely take a theoretical perspective.

This paper instead takes a broader and more applied approach by combining philosophical, methodological and scientific perspectives and by providing concrete examples from the ten most cited RCT studies that together can significantly improve our understanding of the RCT method. The paper thereby outlines a large number of new and important theoretical assumptions, methodological biases and empirical limitations not yet discussed in the medical or philosophical literature that emerge when designing, implementing and analysing trials in the real world. These assumptions include that participants' background traits that affect outcomes would not change between trial groups during trial implementation (i.e. not just no differences at baseline but also at endline); that the particular time points for the baseline and endline would be chosen to adequately reflect the average (or greatest possible) treatment outcome; that randomisation can ensure participants are evenly distributed between trial groups along measurable, non-measurable and unknown background influencers, among many others.

This paper thereby provides a much more comprehensive overview of the range of issues and problems facing RCTs (which is needed to assess an RCT's overall robustness) than in the existing literature that tends to focus on specific issues. Epistemologically, the paper shows that RCTs generally have some degree of bias in their results – as illustrated by assessing the ten most cited RCT studies worldwide. A central and important epistemic topic underlying the paper is thus whether or not RCTs can, despite the range of issues and problems associated with them, provide sufficiently rigorous evidence that would allow us to be confident in their reported causal claims. The paper's implications include that we must not overly rely on any single research method but always combine methods.