

## Why randomised controlled trials of psychological treatments are still essential



The movement towards evidence-based practice in psychotherapy has recently provoked a backlash. In the past few years, some authors have raised questions concerning the validity of randomised controlled trials (RCTs), which have long been regarded as a crucial linchpin for evidence-based practice. These critics have argued that data from RCTs should not be prioritised above data derived from alternative evidentiary sources, such as observational studies, uncontrolled case studies, and clinical experience.<sup>1,2</sup> Although many of the criticisms have been aimed at RCT methodology in general, some sceptics have extended them to raise questions concerning the role of RCTs in the evaluation of psychological treatments.

For example, several scholars have recently contended that RCTs of psychotherapy are marked by limited external validity.<sup>3</sup> Others have noted that RCTs can tell us whether, but not why, interventions work.<sup>3,4</sup> In a 2017 blog post, Jonathan Shedler encouraged psychotherapists to ignore the new practice guidelines for post-traumatic stress disorder released by the American Psychological Association because they relied inordinately on RCTs when formulating recommendations. Similarly, the New Jersey Psychological Association recently maintained that these practice guidelines gave RCTs undue priority at the expense of other forms of evidence. The former Editor-in-Chief of *Science News*, Tom Siegfried, has also joined the fray. Drawing on arguments from a 2015 doctoral thesis by Christopher Blunt of the London School of Economics, Siegfried argued that methodological flaws, such as the frequent failure of random assignment to precisely equalise treatment and control groups on confounding variables, often affect RCTs. Echoing Blunt, Siegfried maintained that because RCTs are group designs, they do not permit practitioners to apply treatment recommendations to individual patients.

These criticisms of RCTs are not entirely without merit. Nevertheless, they provide less than convincing grounds for downgrading the status of RCTs within the evidence-based practice hierarchy relevant to psychotherapy practice.

Notably, many of the criticisms are largely ahistorical. Psychiatry's past is replete with a long line of harmful

interventions, including spinning, blistering, purging, leeching, prefrontal lobotomy, insulin coma therapy, and surgical removal of organs, all of which were justified on the basis of uncontrolled observations and clinical experience.<sup>5,6</sup> More recently, several psychological treatments defended on these grounds, such as crisis debriefing for trauma victims or Scared Straight programmes for conduct disorder, were found to be ineffective or iatrogenic in subsequent controlled studies, including RCTs.<sup>6</sup>

Critics rightly observe that designs other than RCTs can provide useful evidence for psychological treatment outcomes, and that well conducted observational studies can sometimes be more informative than poorly conducted RCTs.<sup>1,2,4</sup> Additionally, rigorous observational studies do not invariably overestimate the magnitudes of effects derived from RCTs.<sup>7</sup> Nonetheless, observational studies cannot support strong causal inferences, and no responsible practitioner would administer a psychological treatment that does not produce causal effects. All else being equal, RCTs provide less biased estimates of psychotherapy outcomes than do other designs.<sup>8</sup> By virtue of their minimisation of between-group differences via randomisation, RCTs afford enhanced control over various causes of spurious therapeutic effectiveness, which can fool observers into believing that ineffective psychotherapies are effective. Causes of spurious therapeutic effectiveness include regression to the mean, spontaneous remission, selection bias, maturation, patients' concurrent use of adjunctive interventions, and other internal validity threats.<sup>6</sup>

Given their better control over causes of spurious therapeutic effectiveness, RCTs are probably superior to other designs with respect to a key scientific criterion: replicability. Indeed, an analysis of highly cited medical trials revealed that RCT designs generate substantially more replicable results than do non-randomised designs.<sup>9</sup>

RCT critics have at times fallen prey to logical errors. Some critics, such as the members of the UK Council of Psychotherapy, have conflated methodological limitations of specific RCTs with RCT methodology per se.<sup>2</sup> For example, they and others have observed legitimately that RCTs have sometimes focused



Maire Fermatillo/  
Science Photo Library

*Lancet Psychiatry* 2018

Published Online  
March 27, 2018  
[http://dx.doi.org/10.1016/S2215-0366\(18\)30045-2](http://dx.doi.org/10.1016/S2215-0366(18)30045-2)

For more on the 2017 blog post by Jonathan Shedler see <https://www.psychologytoday.com/blog/psychologically-minded/201711/selling-bad-therapy-trauma-victims>

For more on the APA Practice Guidelines see [http://www.psychologynj.org/index.php?option=com\\_content&view=article&id=366:friday-update-12-1-17&catid=20:site-content](http://www.psychologynj.org/index.php?option=com_content&view=article&id=366:friday-update-12-1-17&catid=20:site-content)

For more on Tom Siegfried's blog post see <https://www.sciencenews.org/blog/context/critique-medical-evidence-hierarchies>

exclusively on symptomatic changes at the expense of broad changes in psychological functioning, including quality of life.<sup>10</sup> Several critics have also noted that certain RCTs of specific psychiatric disorders have been restricted to patients who do not meet criteria for overlapping (comorbid) conditions, often yielding conclusions that are limited to atypical patient groups.<sup>2,10</sup> Nevertheless, such shortcomings are not inherent flaws of RCT designs, and can be rectified by incorporating measures of everyday functioning into RCTs or conducting RCTs with more heterogeneous samples, or both.

Additionally, some critics apparently assume that ecological validity equates to external validity.<sup>1,10</sup> The concept of ecological validity captures the extent to which conditions mirror real-world settings, whereas external validity captures the extent to which findings generalise to other settings. Although many observational studies possess higher ecological validity than RCTs, their external validity might be lower because of less control over causes of spurious therapeutic effectiveness. Moreover, critics have sometimes presumed a typical or even necessary trade-off between internal and external validity.<sup>1,11</sup> By contrast, Campbell's principle of the "primacy of internal validity" implies that internal validity is a prerequisite for external validity—namely, one cannot meaningfully generalise from a finding unless it is genuine.<sup>12</sup> Finally, several RCT critics<sup>2,3,10</sup> appear to have committed the nirvana fallacy: the error of arguing that because a method is imperfect, it is no better than other methods. A research design can be flawed but still preferable to more flawed designs. Furthermore, standard criticisms of RCTs, such as incomplete blinding and absence of information about treatment mechanisms, apply a fortiori to observational designs. To paraphrase Winston Churchill's quip regarding democracy, RCTs are the worst outcome designs except for all other designs that have been tried.

RCTs are neither panaceas nor gold standards. Nevertheless, they are essential safeguards against inferential errors that can contribute to suboptimal patient care. Recent criticisms notwithstanding, clinicians and researchers should continue to assign greater weight to well conducted RCTs than to other designs in the selection of psychological treatments.

\*Scott O Lilienfeld, Dean McKay, Steven D Hollon

Department of Psychology, Emory University, Atlanta, GA 30322, USA (SOL); Melbourne School of Psychological Sciences, University of Melbourne, Melbourne VIC, Australia (SOL); Department of Psychology, Fordham University, New York, NY, USA (DM); and Department of Psychology, Vanderbilt University, Nashville, TN, USA (SDH)  
slilien@emory.edu

We declare no competing interests.

- 1 Clay R. More than one way to measure. *Am Psychol* 2010; **41**: 52.
- 2 Guy A, Thomas R, Stephenson SL, Loewenthal D. NICE under scrutiny: the impact of the National Institute for Health and Clinical Excellence guidelines on the provision of psychotherapy in UK. *J Psychol Therap Primary Care* 2012; **1**: 77–102.
- 3 Carey TA, Stiles WB. Some problems with randomized controlled trials and some viable alternatives. *Clin Psychol Psychother* 2016; **23**: 87–95.
- 4 Cartwright N, Deaton A. Understanding and misunderstanding randomized controlled trials. Working paper. Durham: Centre for Humanities Engaging Science and Society and National Bureau of Economic Research, 2016.
- 5 Grove WM, Meehl PE. Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychol Public Policy Law* 1996; **2**: 293–323.
- 6 Lilienfeld SO, Ritschel LA, Lynn SJ, Cautin RL, Lutzman RD. Why ineffective psychotherapies appear to work: a taxonomy of causes of spurious therapeutic effectiveness. *Perspect Psychol Sci*, 2014; **9**: 355–87.
- 7 Concato J, Shah N, Horwitz RJ. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000; **342**: 1887–92.
- 8 La Caze A. Evidence based medicine must be.... *J Med Philos* 2009; **34**: 509–27.
- 9 Ioannidis JA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2015; **294**: 218–28.
- 10 Shean G. Limitations of randomized control designs in psychotherapy research. *Adv Psychiatr Treat* 2014; **2014**: 561452.
- 11 Duncan BL, Reese RJ. Empirically supported treatments, evidence based treatments, and evidence based practice. In: Stricker G, Widiger T, Weiner I, eds. *Handbook of Psychology*, 2nd edn. Hoboken: John Wiley & Sons, 2012: 977–1023.
- 12 Cook TD, Campbell DT, Peracchio L. Quasi-experimentation. In: Dunnette MD, Hough LM, eds. *Handbook of industrial and organizational psychology*, 2nd edn. Palo Alto: Consulting Psychologists Press, 1990: 491–576.

For more on the problems with randomised controlled trials see <http://andrewgelman.com/2017/01/10/30829/>