

Clinical Trials in Psychiatry and Clinical Psychology: Science or Product Testing?

Robert L. Woolfolk

Department of Psychology, Rutgers University, United states

Corresponding author:

Robert L. Woolfolk

✉ rwool@rci.rutgers.edu

Department of Psychology, Rutgers University, United states.

Tel: 848-445-2008

Citation: Wilson RL. Clinical Trials in Psychiatry and Clinical Psychology: Science or Product Testing? *Acta Psychopathol.* 2015, 1:2.

Abstract

During the recent era of evidence-based medicine, the randomized controlled trial (RCT) has been regarded as the most authoritative method of evaluating interventions. The methodology is utilized not only in medicine, but in other fields such as economics, education, and agriculture. In psychiatry and clinical psychology, RCTs have been utilized extensively in conjunction with the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM) [1]. This RCT/DSM combination has produced somewhat limited progress, both in identifying efficacious treatments and in facilitating progress in better understanding the scientific foundations of clinical intervention in these fields. This unhappy circumstance results not from limitations of the RCT as a tool of inductive logic, but rather its use with data that are neither theoretically grounded nor psychometrically sound, under background conditions in which publication bias and economic interest converge to distort the rational, impartial use of the RCT. Until the biases due to human interests are reduced and the fields of psychiatry and clinical psychology are more scientifically advanced, the RCT will be of limited use.

Keywords: Clinical Trials; RCT; DSM; Methodology

Abbreviations: RCT- Randomized Controlled Trial; DSM- Diagnostic and Statistical Manual of the American Psychiatric Association

Received: July 06, 2015; **Accepted:** September 21, 2015; **Published:** September 30, 2015

Main Text

The randomized controlled trial (RCT) has become the “gold standard” in outcome research evaluating both somatic and psychosocial treatments for psychopathology. The literature abounds with expositions of the RCT’s strengths as a methodology, and I would concur with many of those. What follows, however, is a discussion of its limitations as an explanatory device when specifically applied within psychiatry and clinical psychology and an explanation of why those limitations are particularly problematic for those fields.

The typical state of the art therapy outcome study in psychiatry involves patients who have been categorized via the various DSMs and randomly assigned to an experimental condition in which they will receive one of several possible interventions: a treatment being evaluated, a specified comparator treatment, some generic and loosely specified “treatment as usual” or a

placebo; or they may receive no treatment at all. Preferably, the study is double-blinded, such that neither the patients nor study personnel should be aware of the treatment group to which the patients have been assigned. The RCT was the cornerstone of what appeared to many during the late twentieth century to be a leap of scientific progress in the treatment of psychopathology. Putatively efficacious pharmacological therapies for mental illness, tested in RCTs and, as a consequence, approved by the Food and Drug Administration (FDA), were plentiful and burgeoning.

But the progress of a clinical science based on the DSM/RCT framework, until recently, seems to have been greatly overstated and is now being subjected to a revisionist analysis by numerous authorities. Former NIMH director Stephen Hyman refers to a biomedical psychiatry “revolution stalled” [2], and current director, Thomas Insel, questions whether the recent era of biomedical psychiatry has actually produced more than minimal

benefits for psychiatric patients or contributed significantly to a clinical science of psychopathology [3]. Concurrently a sizable portion of the pharmaceutical industry has shut down its psychiatric drug research and development, proclaiming that promising molecular targets have not emerged from the kind of research that evaluates therapeutic efficacy using the DSM/RCT structure [4, 5]. Thomas Insel [3] laments that despite widespread use of psychotropic medications vetted in RCTs, “these last few decades have not seen reductions in morbidity or mortality for people with serious mental illness (p. 1).” Although the recent regime of biomedical psychiatry was a financial smash hit for the pharmaceutical industry, it is beginning to look like a flop as either science or technology. How could this be, with so many investigations having reported statistically significant findings and what were thought to be clinically meaningful effect sizes?

Part of the answer lies in the fit between the RCT and the scientific and societal contexts in which it has been applied. Within the fields of psychiatry and clinical psychology, as they are presently constituted and operate within their cultures, the RCT is problematic as a tool of scientific inquiry. This is not due to any fundamental defect within the RCT, assuming it is applied to phenomena for which measurement is unproblematic and when all non-trivial causal variables can be controlled or systematically manipulated. When it is utilized under the appropriate circumstances the RCT is impeccable as a tool of inductive logic in identifying cause and effect, or covariation among variables that is not due to chance.

The contemporary RCT employed in medicine, is a slightly more refined version of forebearer factorial experimental designs, employed in education, psychology, and agriculture. Factorial designs were created by Sir Ronald Fisher [6] decades before the first published study that we today would label a randomized clinical trial, which was a test of streptomycin in treating tuberculosis, published in 1948 [7]. Fisher designed the experimental structure that evolved into the RCT for the purpose of studying crop yields on those famous plots of land at the Rothamsted Experimental Station in Hertfordshire, England. There Fisher married the mathematics of probability and the inductive logic of John Stuart Mill [8]. Fisher’s development of experimental designs and statistical analyses was a necessary to validly identify cause and affect relationships among phenomena that are highly variable. When dealing with the effectively invariant entities and processes of the “hard sciences,” such as gravitation or the formation of chemical bonds, the methodological tools of the biobehavioral sciences are mostly irrelevant. Living things, however, are characterized by variability, complexity, and context dependency that makes the more straightforward observation methods of the natural sciences inadequate.

When the RCT is utilized in somatic medicine to treat a well-understood disease, e.g., tuberculosis, via some treatment whose mechanism of action is understood and outcomes are physical changes that can be measured objectively, the RCT can work well, if properly implemented. Even under such ideal conditions, we can be led astray by the published findings of RCTs. This is because editors, reviewers, and investigators are human beings have extra-scientific interests and proclivities that can conflict with the

canons of scientific rationality. We know that findings from the most prestigious medical journals (RCTs and less well-controlled studies) more than infrequently either are not replicated or are contradicted by subsequent studies [9]. Data from RCTs is more than occasionally selectively reported [10]. Some failed RCTs are never submitted for publication and journals in the biomedical sciences tend to favor novel, positive findings to a greater degree than fields such as chemistry and physics [11].

When the RCT is applied to treatment outcome research in psychiatry and clinical psychology, as opposed to somatic medicine, various circumstances converge to make RCTs much more unlikely to produce sound inferences. This unhappy circumstance results not from limitations of the RCT as a tool of inductive logic, but rather its use with data that are neither theoretically comprehended nor psychometrically unimpeachable, under background conditions in which publication bias and economic interest converge to distort the rational use of RCTs and the interpretations of their findings. In psychiatry and clinical psychology we have an unfortunate confluence of several factors the militate against the effective use of RCTs: 1) lack of understanding of mechanisms underlying treatment 2) lack of understanding of mechanisms involved in the etiology of mental disorders 3) a diagnostic system based on symptom clusters and no validated theories of psychopathology 4) the absence of any established biomarker, making assessment of outcome highly subjective and vulnerable to various biases 5) powerful extra-scientific factors that can control or influence investigators, reviewers, editors, and funders.

A large part of the problem that has been widely recognized is that the system of classification (DSM) lacks validity in the sense that both the systems of diagnostic classification and the psychiatric RCT outcome measures are at best indirect and highly inferential indicators that may not map onto any fundamental brain process. There is no biological test that confirms or disconfirms a DSM diagnosis or that of any other system of psychopathology classification. The structured clinical interview, really a guided patient self-report with some clinician inference alloyed, is the gold-standard method of diagnosis. Clinical ratings of patient status based on patient self-report are the primary outcome measures in many psychiatric studies. These measures are highly subjective. Asserting that little substantive scientific understanding of underlying mechanisms of illness or cure has been generated by the last few decades of treatment outcome research, the NIMH has publically critiqued the symptom cluster approach underlying the DSMs and is searching for a system underlain by basic biological science [12]. NIMH currently is undertaking an effort to create a diagnostic system based on neuroscience, the Research Diagnostic Criteria (RDoC) but the work is in its infancy.

The hope that differential response to treatments would validate diagnoses (pharmacological dissection) remains unfulfilled. Drugs with very different mechanisms of action have comparable effects on a single disorder, while singular treatments seem to produce effects in trials that are very broad and efficacious for a variety of maladies. Rather than a multitude of treatments that are differentially effective for particular disorders, many

treatments are found to be transdiagnostically efficacious, producing a kind of conceptual muddle, if one is a scientist seeking systematic relationships. For example, if we look at perhaps the most frequently investigated mental disorder, depression, once confidently being proclaimed by leading psychiatrists to result from a deficiency in serotonin, the current consensus conclusion enunciated by Stephen Hyman (2) is that, "Despite the resource investment, this [pharmacological] research has not substantially clarified the pathogenesis or pathophysiology of depression or other phenotypes characterized by negative affectivity, or the complex and interesting actions of serotonin in the human brain (p. 2)."

We should, of course, remember that research in psychology and psychiatry is fraught with difficulties not present in other sciences. Capturing and assessing mental life never involves placing calipers on the thing itself. In psychometrics, the field that has arisen to cope with the additional measurement issues raised by a science of psychology, all measurement is essentially indirect and, therefore, uncertain and approximate. In prior years, much of the appeal of behaviorism (and its focus on "objective" publicly observable data) was the possibility of sidestepping or explaining away a fundamental difficulty that confronts our field. How are we to measure the mind and in so doing, what compromises and concessions is it scientifically acceptable to make?

In RCTs the objectivity required of scientific investigation is difficult to achieve when the dependent variables are, in essence, subjective data provided by participants or raters, data that cannot be validated against any objective standard of measurement. The experimental control offered by true blinding is difficult to achieve among raters or patients. In drug trials medication side effects often break the blind; in psychotherapy studies, complete blinding simply cannot occur because the patient has a particular awareness of the participation in the treatment. Market forces can have critical influence on study design and the reporting of results. Biased findings are more likely when psychiatric trials are sponsored by industries that will profit only if the trial is successful [13, 14]. Much of the data that we analyze turns out to be, in essence, stories told by one person to another person.

Theoretically speaking, the RCT is a logically flawless method when applied under ideal conditions, but as it has been employed in psychiatry and clinical psychology it has manifested some shortcomings. The field has clearly expected too much of it. To quote the document *Developing and Evaluating Complex Interventions* prepared by the U.K.'s Medical Research Council, before testing an intervention one should ask the question, "Does your intervention have a coherent theoretical basis?" (p. 4). And in many trials there has been an abject absence of adequate theory. When the field makes more scientific progress and understands the mechanisms of disorder and treatment, all methods of research will likely be of greater use. In the meantime we need not solely rely on the RCT. Methodologist Alan Kazdin has decried "overreliance" on the RCT and suggested that the field should look for convergences among findings of RCTs, single-subject designs of the sort employed in the applied behavior analysis tradition, qualitative research that can provide a rich account of the phenomenological aspects of response to treatment, and case studies conducted in a disciplined fashion that allow for sound processes of inductive inference [15]. Also we do not need to hamstring the RCT by tying it to DSM-5. Barlow's transdiagnostic approach [16] and recent symptom-focused approaches to psychosis [17] have untethered the RCT from the DSM. Given the developments at NIMH, this trend may continue. Also efforts to make sure that research using RCTs is adequately powered and controlled do seem to improve the rate of replicability [18]. The use of meta-analysis can make us less reliant upon findings from a single high profile, multi-site trial and it biases it may contain.

Until we better understand the mind/brain, our intervention RCT's will be, for the most part, analogous to industrial product testing or educational program evaluation. They are pragmatic tests of the practical effects of treatment methods rather than true scientific experiments that are capable of extending and deepening our knowledge beyond the "People like Coke better than Pepsi," kind of factoid. In conjunction with assessment of both statistical and clinical significance such knowledge may be of great practical utility and of social benefit, but limited in the fundamental scientific advance that it can promote.

References

- 1 American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (5th ed) Washington.
- 2 Hyman SE (2012) Revolution stalled. *Sci. Transl. Med.* 4: 155cm11.
- 3 Insel TR (2012) Next-generation treatments for mental disorders. *Sci. Transl. Med* 4: 1-9.
- 4 G. Miller (2010) Is pharma running out of brainy ideas? *Science* 329: 502-504.
- 5 Abbott A (2011) Novartis to shut brain research facility: Drug giant redirects psychiatric efforts to genetics. *Nature* 480: 161-162.
- 6 Fisher RA (1971) Collected papers of R. A. Fisher (Ed.), J.H. Bennett., Adelaide: The University of Adelaide Press.
- 7 Meldrum ML (2000) A brief history of the randomized controlled trial: From oranges and lemons to the gold standard. *Hematol Oncol Clin North Am* 14: 745-760.
- 8 Mill JS (1843) System of logic: Ratiocinative and inductive. In J. M. Robson (Ed.), *The collected works of John Stuart Mill* (Vols. 7 & 8). Toronto, Ontario, Canada: University of Toronto Press (1973)-Original work published.
- 9 Ioannidis JP (2005) Contradicted and initially stronger effects in highly cited clinical research. *Journal of the American Medical Association* 294: 218-228.
- 10 Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P (2009) Comparison of registered and published primary outcomes in randomized controlled trials. *Journal of the American Medical Association* 302: 977-984.
- 11 Fanelli D (2011) Negative results are disappearing from most disciplines and countries. *Scientometrics* 90: 891-904.
- 12 Insel TR (2013) Director's blog: Transforming diagnosis. Retrieved October 15, 2013, from the National Institute of Mental Health site.
- 13 Mathieu S, Boutron I, Moher D, Altman DG & Ravaud P (2009) Comparison of registered and published primary outcomes in randomized controlled trials. *Journal of the American Medical Association* 302: 977-984.
- 14 Vedula SS, Bero L, Scherer RW, Dickersin K (2009) Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *N Engl J Med* 361: 1963-1971.
- 15 Kazdin AE (2006) Assessment and evaluation in clinical practice. In Goodheart CD, Kazdin AE, Sternberg RJ (Eds.), *Evidence-based psychotherapy: Where practice and research meet*. Washington, DC: American Psychological Association 153-177.
- 16 Barlow DH, Allen LB, Choate ML (2004) Toward a unified treatment for emotional disorders. *Behaviour and Research Therapy* 35: 205-230.
- 17 Freeman D (2011) Improving cognitive treatments for delusions. *Schizophr Res* 132: 135-139.
- 18 Ioannidis JP (2005) Why most published research findings are false. *PLoS Medicine* 2: 696-701.