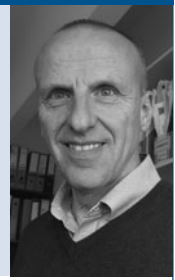


Editorial

A realist approach to the evaluation of complex mental health interventions

Craig Duncan, Scott Weich, Sarah-Jane Fenton, Liz Twigg, Graham Moon, Jason Madan, Swaran P. Singh, David Crepaz-Keay, Helen Parsons and Kamaldeep Bhui

**Summary**

Conventional approaches to evidence that prioritise randomised controlled trials appear increasingly inadequate for the evaluation of complex mental health interventions. By focusing on causal mechanisms and understanding the complex interactions between interventions, patients and contexts, realist approaches offer a productive alternative. Although the approaches might be combined, substantial barriers remain.

Declaration of interest

All authors had financial support from the National Institute for Health Research Health Services and Delivery Research Programme while completing this work. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the National Health Service, the National

Institute for Health Research, the Medical Research Council, Central Commissioning Facility, National Institute for Health Research Evaluation, Trials and Studies Coordinating Centre, the Health Services and Delivery Research Programme or the Department of Health. S.P.S. is part funded by Collaboration for Leadership in Applied Health Research and Care West Midlands. K.B. is editor of the British Journal of Psychiatry.

Keywords

Randomised controlled trial; complex interventions; realist evaluation.

Copyright and usage

© The Royal College of Psychiatrists 2018.

Craig Duncan (pictured) is a Senior Research Fellow in the Department of Geography at the University of Portsmouth. His research focuses on the role of context in health outcomes and behaviours. **Scott Weich** is a Professor of Mental Health in the School of Health and Related Research at the University of Sheffield. He is head of the Mental Health Research Unit and has been an Honorary Consultant Psychiatrist for over 20 years.

Sarah-Jane Fenton is a Lecturer in Mental Health Policy in the Institute for Mental Health at the University of Birmingham. She has particular expertise in youth, adolescence, mental health policy and realist methods. **Liz Twigg** is a Professor of Human Geography in the Department of Geography at the University of Portsmouth. She has over 25 years' experience of using quantitative models to investigate place effects on mental and physical health outcomes and health risk behaviours. **Graham Moon** is a Professor of Spatial Analysis in Human Geography in the School of Geography and Environment at the University of Southampton. He researches psychiatric health service provision and place effects on health-related behaviours. His work is supported by NIHR and ESRC.

Jason Madan is a Professor in Health Economics in the Warwick Clinical Trials Unit at the University of Warwick. He is also Director of Postgraduate Research at Warwick Medical School. **Swaran P. Singh** is Head of Mental Health and Wellbeing at Warwick University, Associate Medical Director for Research and Innovation at Coventry and Warwickshire Partnership Trust and Commissioner for Equality and Human Rights Commission, UK.

David Crepaz-Keay is Head of Empowerment and Social Inclusion at the Mental Health Foundation. **Helen Parsons** is a Senior Research Fellow at Warwick Medical School. She is a medical statistician and is based at Warwick Clinical Trials Unit. **Kamaldeep Bhui** is a Professor of Cultural Psychiatry and Epidemiology at the Centre for Psychiatry, Queen Mary University of London and Honorary Consultant Psychiatrist at East London NHS Foundation Trust.

Evidence-based policy and practise is firmly established in mental healthcare, bringing with it a need for evidence to be critically appraised. The conventional biomedical view sees randomised controlled trials (RCTs) as the gold standard against which all other forms of evidence should be judged. The suitability of RCTs for evaluating complex mental health interventions has long been questioned. An alternative approach to evidence, realist evaluation, has been developed and advocated across other policy areas. Realism challenges the foundations on which scientific knowledge is created and represents a paradigm shift. We outline the advantages of realist evaluation for mental health research and explore what this might mean for the current hierarchy of evidence on which policy and commissioning decisions are based.

Randomised controlled trials

RCTs have played a vital role in the development of evidence-based mental healthcare. When applied in their ideal, or close to ideal form, their ability to control for confounding, reduce bias and elucidate the direction of causation has enabled robust evaluation (and quantification) of the effectiveness of interventions. Estimates of 'overall' net effect sizes derived from trial results have allowed policy makers and commissioners to estimate the health economic consequences of delivering interventions to target populations.

However, work across a range of disciplines has suggested that the strengths of RCTs may not be sufficient to warrant their inherent place as the gold standard.¹ Many of the criteria for a well-conducted RCT are often not achieved in practise because of attrition, lack of blinding and other biases post-randomisation. Moreover, RCTs are intrinsically better suited to some areas and research questions than others. Mental healthcare, other than that limited to pharmaceutical interventions, is one area where RCTs fit less well.² Psychological treatments rely on human agency (and specifically interpersonal interactions) and are therefore harder to manualise than drug treatments.³ Unlike pharmaceutical interventions, psychological and behavioural interventions typically change during treatment. The skill of therapy, after all, lies in the ability to understand and meet patients' individual needs and to respond to therapeutic progress or stasis. Complex mental health interventions, based on multiple interactive components and centred on interpersonal and informational activities, are problematic within an RCT paradigm.

Recent moves towards personalised medicine represent a further challenge to the predominant position of RCTs. Average treatment effects, even if estimated on the basis of well-conducted trials, apply to groups rather than individuals, and hence do not apply equally to everyone. This is especially so when treatments are multifaceted and target individual human agency and interactions. This also applies to social and other contexts: what works in one setting may not work in another. The social, economic and cultural contexts in which patients and complex interventions are embedded will, therefore, have an effect at many levels, and although data arising from RCTs

may be used to test for subgroup effects (i.e. interactions between allocation group and factors such as patient characteristics or setting), these tests are limited in number and often lack statistical power.

Realist evaluation

There is a need for alternative approaches to the evaluation of complex mental health interventions. Crucially, approaches are needed that recognise, examine and evaluate variation and sources of variation in patient outcomes, with particular reference to the role of context.

Realist evaluation is one such approach. Arising from the work of scientific realists, most particularly Pawson and Tilley, the central formulation in realist evaluation concerns context, mechanism and outcome.⁴ Causation that generates outcomes is established through identifying mechanisms that may be activated or lie dormant, depending on context. Pawson and Tilley emphasise that interventions, in and of themselves, do not 'work' but rest instead on the choices, intentions and behaviours of social actors as they engage with the mechanisms that interventions contain. Thus, outcomes of interventions derive from the operation of specific mechanisms in particular contexts; using an often-quoted analogy from the physical world, sparks lead to gunpowder exploding but only under specific conditions (the presence of oxygen, a dry atmosphere, etc.). The mechanism is based on the chemical composition of gunpowder whereas the conditions derive from the context. The context-dependent operation of mechanisms leads, in turn, to the development of so-called 'CMO configurations' – specific instances of contexts and mechanisms that are hypothesised to produce particular patterns of outcomes.

Realist evaluation and networks of evidence

According to realist theory, evaluation cannot simply be concerned with estimating net effect sizes, i.e. in finding out whether a complex intervention works overall or 'on average'. Instead, researchers should seek to establish *what works, for whom, under what circumstances and why*. Crucially, it is the careful thinking about and exploration of mechanisms – which pragmatic RCTs eschew – that allows researchers to uncover why interventions work and not simply what interventions work overall for populations of trial participants.

Given this formulation, realist evaluation is concerned with building 'programme theories' that aim to capture the mechanisms underlying complex interventions in the contexts in which they work. These theories are developed and refined through an iterative process centred on the findings of empirical research (networks of evidence), which may include but does not prioritise the results of RCTs. In this way, realist evaluation does not favour any particular research design or methods of analysis. Certainly, intensive (qualitative) methods have a role to play, especially those that move beyond simply recording the perceptions and views of those involved in interventions. Equally, quantitative observational research can capture variation in outcomes that may be overlooked in trials that are concerned with evaluating overall effectiveness, and are typically based on far smaller and more homogenous (and hence less representative) samples than studies using routine clinical data-sets.

A more open and diverse approach to evidence does not, however, constitute an 'anything goes' attitude or a relinquishing of scientific values. Rather, as Pawson puts it, evaluating interventions in complex social worlds becomes the scientific task of 'scavenging' amongst a range of evidence developed through a plurality of methods and subjecting each part of it to 'organised scepticism'.⁵ Methodological guidelines, training resources and quality and reporting standards have been produced for this task by the Realist And Meta-narrative Evidence Syntheses: Evolving Standards (RAMESES) project, funded by the National Institute for Health Research.⁶

Can realist evaluation and RCTs ever coexist?

It has been suggested that adopting a realist perspective could lead to significant improvements in the way RCTs are conducted. Triallists have also made their own advances with the basic parallel group design evolving to more advanced designs such as stepped wedge and patient preference designs, which are better suited to evaluating complex interventions.

Realism may bring additional gains to the conduct of RCTs. The emphasis on underlying causal mechanisms highlights the importance of elucidating theories around 'mechanisms of action' before the actual running of trials. Paradoxically, although this type of work-up is often conducted in relation to trials of specific clinical interventions (e.g. pharmaceutical and psychotherapeutic interventions), it is often missing from the development of more complex psychosocial interventions centred on mental health service delivery and patient behaviour.

Realist evaluation might also be seen as a means of improving parallel process evaluations of RCTs of complex interventions. For example, the work of Byng and colleagues on the Mental Health Link in London provided information that enabled the results of the accompanying RCT to be interpreted, as well as providing key insights into the core functions of the intervention.⁷

Realist evaluation principles have also been used to shape the design and implementation of RCTs themselves, leading to the creation of 'realist RCTs'.⁸ Although similar to realist-based qualitative process evaluation, realist RCTs also consist of quantitative mediation and moderation analyses that aim to capture variations in intervention outcomes between contexts and subgroups. Such analyses have shown how a Welsh intervention aimed at increasing physical activity worked for patients at risk of coronary heart disease but not for patients referred for mental health reasons.⁹

This approach has, however, been roundly dismissed as antithetical to realism.¹⁰ According to realists, the fundamental characteristics and ambitions of RCTs mean that they are wholly unsuitable to (and in principle, designed to obviate rather than embrace) variability in the interactions between interventions, agents and contexts, seeking where possible to minimise the effects of all but the first of these, and thereafter to reduce the other two to variables to be measured. On the other hand, realists, despite framing programme theories as testable hypotheses, might be accused of avoiding refutation by insisting on the primacy of (further) theorising.

Although scientific psychiatry has always been proudly pragmatic and eclectic, the question we now face is whether these two epistemologies (positivism and realism) can ever be reconciled and harnessed in a common evaluative approach. With some justification, proponents of realism object to this approach being subverted by the positivism and emphasis on outcomes simplified to average effects inherent in RCTs. Realist evaluation, by contrast, is about understanding mechanisms of change based in infinitely complex (and changing) interactions between agents and contexts, and for many, never the twain shall meet.

Conclusion

RCTs have significant limitations when it comes to evaluating complex and context-dependent interventions, of the type that are common in mental healthcare. By focusing on causal mechanisms and adopting a more pragmatic attitude to evidence, approaches based on realism offer potential for creating forms of evidence that are beyond the scope of positivist clinical trials. Realist RCTs are not, it seems, for everyone, and whether this approach is

worth pursuing (and investing public research funds) will ultimately come down to a question of utility (and value of information) in the generation of evidence to inform policy and healthcare commissioning decisions.

Craig Duncan, PhD, Senior Research Fellow, Department of Geography, University of Portsmouth, UK; **Scott Weich**, MSc, MD, FRCPsych, Head of the Mental Health Research Unit, Honorary Consultant Psychiatrist and Professor of Mental Health, School of Health and Related Research, University of Sheffield, UK; **Sarah-Jane Fenton**, PhD, Lecturer in Mental Health Policy, Institute for Mental Health, University of Birmingham, UK; **Liz Twigg**, PhD, Professor of Human Geography, Department of Geography, University of Portsmouth, UK; **Graham Moon**, PhD, HonMFPH, FACSS, Professor of Spatial Analysis in Human Geography, School of Geography and Environment, University of Southampton, UK; **Jason Madan**, PhD, Professor in Health Economics, Warwick Clinical Trials Unit and Director of Postgraduate Research, Warwick Medical School, University of Warwick, UK; **Swaran P. Singh**, MBBS, MD, DM, FRCPsych, Head of Mental Health and Wellbeing, Warwick Medical School, University of Warwick, UK and Associate Medical Director for Research and Innovation, Coventry and Warwickshire Partnership Trust, UK and Commissioner for Equality and Human Rights Commission, UK; **David Crepaz-Keay**, DProf, Head of Empowerment and Social Inclusion, Mental Health Foundation, UK; **Helen Parsons**, MSc, PhD, Senior Research Fellow, Warwick Medical School, University of Warwick, UK and Medical Statistician, Warwick Clinical Trials Unit, UK; **Kamaldeep Bhui**, MD, FRCPsych, Professor of Cultural Psychiatry and Epidemiology, Centre for Psychiatry, Barts and The London School of Medicine & Dentistry, Queen Mary University of London, UK and Honorary Consultant Psychiatrist, East London NHS Foundation Trust, UK

Correspondence: Craig Duncan, PhD, Department of Geography, Buckingham Building, Lion Terrace, University of Portsmouth, Portsmouth PO1 3HE, UK. Email: craig.duncan@port.ac.uk

First received 25 Jan 2018, final revision 27 Mar 2018, accepted 26 Apr 2018

References

- 1 Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med* 2017, in press.
- 2 Priebe S, Slade M. *Evidence in Mental Health Care*. Brunner-Routledge, 2002.
- 3 Carey TA, Stiles WB. Some problems with randomized controlled trials and some viable alternatives. *Clin Psychol Psychother* 2016; **23**(1): 87–95.
- 4 Pawson R, Tilley N. *Realistic Evaluation*. Sage, 1997.
- 5 Pawson R. *The Science of Evaluation*. Sage, 2013.
- 6 Greenhalgh T, Wong G, Westhorp G, Pawson R. Protocol-realist and meta-narrative evidence synthesis: evolving standards (RAMESES). *BMC Med Res Methodol* 2011; **11**(1): 115.
- 7 Byng R, Norman I, Redfern S, Jones R. Exposing the key functions of a complex intervention for shared care in mental health: case study of a process evaluation. *BMC Health Serv Res* 2008; **8**(1): 274.
- 8 Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Soc Sci Med* 2012; **75**(12): 2299–306.
- 9 Murphy SM, Edwards RT, Williams N, Raisanen L, Moore G, Linck P, et al. An evaluation of the effectiveness and cost effectiveness of the National Exercise Referral Scheme in Wales, UK: a randomised controlled trial of a public health policy initiative. *J Epidemiol Community Health* 2012; **66**(8): 745–53.
- 10 Van Belle S, Wong G, Westhorp G, Pearson M, Emmel N, Manzano A, et al. Can 'realist' randomised controlled trials be genuinely realist? *Trials* 2016; **17**(1): 313.