Policy Review

Test, Learn, Adapt

Penelope J. Teoh[a], Christian F. Camm BA (Hons.)[b]

[a]University of Southampton Medical School, [b]New College, University of Oxford, ■■■
Correspondence to: Penelope J. Teoh, 1 Sirdar Mews, 20 Sirdar Road, Southampton, SO17 3JS, Tel: +44 7411 551 505,
Email: penelopejt@gmail.com

Article history

Received: 11 November 2012
Accepted: 2 December 2012
Available online: 1 February 2012

Provenance and Peer Review

Uncommissioned, editorial review

Keywords

Clinical Trials
Government Policy
Guideline

A recent paper[1] has highlighted the use of randomised controlled trials (RCTs) as a basis for creating and implementing policies by the United Kingdom government. However, RCTs are potentially difficult, costly, and complex. Therefore, this paper explores what public policy-makers can learn from medicine's extensive experience with RCTs so that public policy may be both cost-effective and efficacious.

The first RCT in the health sector is often accredited to Sir A Bradford Hill, who randomised patients to a control arm and to a streptomycin arm to treat tuberculosis in 1948.[2] However, history is pitted with earlier examples of controlled trials including the biblical Book of Daniel and James Lind's scurvy trial of 1747.[2] Since then, RCTs have become a staple research tool in the health sector. They provide high quality evidence by using a randomised control group to compare to the intervention group being tested, removing bias and ensuring the intervention itself, and not other factors, exerts an effect upon the results.

RCTs aim to establish a causal link between interventions and outcomes. As such, they are used extensively to assess the usefulness of drugs, interventions, and changes in practice. As they have aided the health sector, business, and many other areas of society, should they also be used systematically to aid policy-makers?

Randomised Controlled Trials

RCTs compare two groups which are randomly assigned to different interventions. This, combined with blinding and standardized measurement of outcomes, allows for effective comparisons of the control and the intervention, or two different interventions.

The use of RCTs outside of health care is not entirely novel. Examples in the public- and private-sector include:

- In business, consumers using Netflix were randomised to four different forms of a new service to determine which users watched more movies.[1]
- In international development, some schools in India were provided with a low-cost tutor to help students falling behind and were compared to schools without the extra tutor.[1]

One example of the use of an RCT in public policy, as opposed to the health sector, involves sending message reminders to those with unpaid court fines. The Court Service and Behavioural Insights Team randomly allocated individuals to a control group, to whom no text message were sent, and to other groups sent text message reminders with or without personalised information. This RCT demonstrated that sending a reminder improved payments. In particular, sending a reminder with the individual's name improved responses to over 30%. If implemented, this intervention would yield 3 million pounds and prevent 150,000 bailiff visits annually. Furthermore, as outcome data was already being collected, costs were minimal.

Table 1 shows the nine steps created by the Cabinet Office to produce an RCT in the public sector in order to help policy-makers.[1] Their advice echoes principles followed to hold a good RCT in the health sector.[3]

The Benefits of Using RCTs

The decision process between different solutions to social issues can take different forms. It is suggested that progress in society is primarily made by trial and error.[4] This implies that a variety of options are consecutively attempted in an unscientific manner until something seems to work. Depending on your political persuasion, the free market may be seen as an environment in which different solutions can develop and market pressures will decide which should prevail. The utilisation of RCTs avoids the need for trial and error or market forces to determine public policy. Policies can be developed and analysed in a scientific manner with clearly defined outcomes.

Determining the Efficacy of a Policy

RCTs provide evidence superior to the general thoughts and predictions of policymakers. In medicine, practitioners often thought their experiences were sufficient in predicting the efficacy of a treatment, however RCTs are now widely used, and are seen as far more effective. For example, the Cardiac Arrythmia Suppression Trial (CAST) was designed

Table 1 Nine steps to conduct an RCT[1]

Test	
1	Identify two or more policy interventions to compare (e.g. old vs new policy; different variations of a policy).
2	Determine the outcome that the policy is intended to influence and how it will be measured in the trial.
3	Decide on the randomisation unit: whether to randomise to intervention and control groups at the level of individuals, institutions (e.g. schools), or geographical areas (e.g. local authorities).
4	Determine how many units, people, institutions, or areas are required for robust results.
5	Assign each unit to one of the policy interventions using a robust randomisation method.
6	Introduce the policy interventions to the assigned groups.
Learn	
7	Measure the results and determine the impact of the policy interventions.
Adapt	
8	Adapt your policy intervention to reflect your findings.
9	Return to Step 1 to continually improve your understanding of what works

Table 2 Possible errors in RCTs, adapted from Keirse *et al.*[9]

Systematic Errors	Random Errors
Selection bias (allocation bias, exclusion bias)	Type 1 Error
Observer bias	Type 2 Error
Co-intervention bias or the Hawthorne effect	
Contamination bias	

to determine the effect of antiarrhythmic therapy on sudden death in survivors of myocardial infarctions (MI).[5] Survivors of an MI often have ventricular premature depolarization, and it was supposed that arrhythmia suppression may reduce risk of sudden death. However, patients treated with antiarrhythmics ultimately had a higher rate of death from arrhythmias. This RCT has prevented the administration of harmful antiarrhytmic drugs to MI survivors. RCTs provide a high level of evidence in evidence-based medicine; doctors can deliver one of their most valuable tools to policy-makers.[6] RCTs also demonstrate the effects of no intervention at all. These results can determine whether a policy should be scrapped or used more widely.

Current practice to determine cost-effectiveness of a policy in the public sector, involves the piloting of an intervention in one geographical location. This has clear drawbacks compared to an RCT. It has a poor comparator; the prior system is in a different geographical area, with less interest in its implementation and no blinding to the results. No randomisation means there is the distinct possibility of bias in selection of the geographical area, and in addition that location may not be representative of the full area the results will be applied to. The validity of results may be poor due to the small, but highly motivated, sample size. Due to the reasons stated, governments may be inclined to see the intervention as effective and, rather than testing feasibility, pilots may just be a stepping block on the road to rolling an intervention out nationwide.

The Disadvantages of Using RCTs

Given the potential errors (Table 2), good methodology is critical and we would advise that criteria similar to those extolled by Jadad *et al.* are followed.[7] Furthermore, if effective RCTs are not correctly reported, discussions and future policies may be invalid. It is imperative reporting is done well and in line with the CONSORT guidelines,[8] Policy-makers should follow the same guidelines in order to present their RCTs and facilitate discussion.

Cost

As belts tighten, and public sector cuts become more numerous, policies need to be as cost-effective as ever. However, RCTs can be very costly. Costs are incurred duplicating the infrastructure in order to implement two interventions, whilst bureaucracy can also inflate costs. However, such costs depend upon the RCT design. In the public sector, RCTs may be more easily implemented due to lack of expensive drugs, surgery, or recruitment of participants not normally accessed. When testing the efficacy of current policies, some interventions are already in place and there is no cost associated with the implementation of the policy- just with setting the RCT up. In addition, outcomes may already be monitored by the public sector. Disregarding finances, the report released by the Cabinet Office asks, "What are the costs of not doing an RCT?"

RCTs can be low cost in the long-term and can determine which policies are most cost-effective and where savings can be made. Not only do RCTs determine efficacy of a policy, they demonstrate if such benefits outweigh the policy's cost. They can direct whether new policies are implemented and if existing ones are still effective. RCTs can also identify what areas of policies are most effective and if they should be modified to become more valuable. In the public sector, the most cost-effective policies can be implemented in more areas and least cost-effective policies may be halted. The financial savings can be utilized on other policies, known as rational disinvestment.[10]

Complexities

RCTs can be difficult to set up and run. In the health sector, the help of contract research organisations (CRO's) are often sought to perform RCTs, to avoid poor allocation concealment and other difficulties. In the public sector, the Cabinet's report refers policy makers instead to the Behavioural Insights Team[11] for support from experienced academics.

Gold standard?

Although RCTs are used extensively in the health sector, some argue that within economics other trial designs can give equivalent evidence.[12] Cartwright states that the reason RCTs are thought to be gold standard is their deductive method.[11] She claims other forms of evidence, for example the econometric method, may be an alternative form of deductive evidence. In addition, RCTs may have a narrow scope to balance internal and external validity.

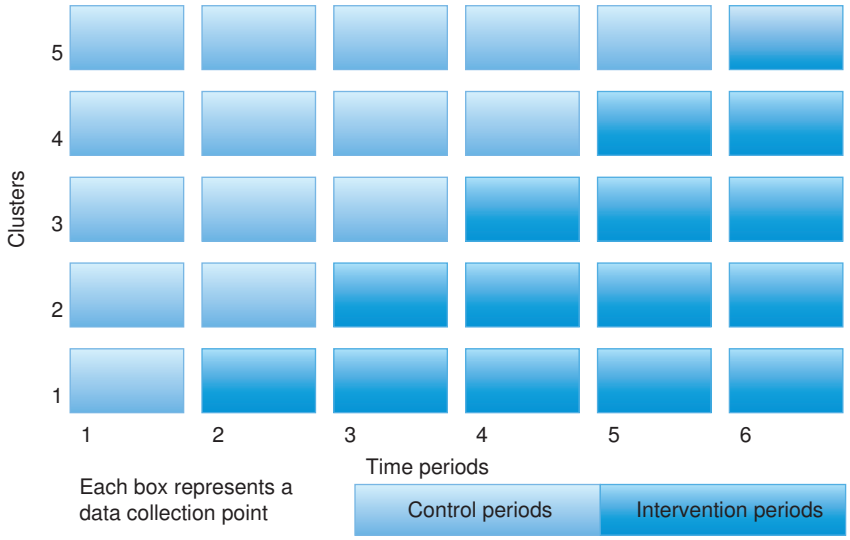

Fig 1 A stepped wedge trial design, adapted from Brown *et al*.[14]

What Else Can Be Learnt from RCTs in the Health Sector?

Peer review and open access

Results of such RCTs should also be easily accessible to society and the public-sector, allowing both assessment of its validity and discussion of other alternative, more effective, interventions. In medicine, this open access stimulates and educates the field. Furthermore, others can replicate the same policies with knowledge of its cost-effectiveness. Allowing the public to see the work of their own government, and the evidence and reasoning behind certain policies improves public support for such policies. The recent white paper released by the cabinet describes how, 'transparency is at the heart of [their] agenda for government' in order to 'reform public services'.[13]

Policy-makers can also learn from medical sciences that an RCT demonstrating that an intervention is not effective is just as successful as the converse. Both types of result can benefit policies.

Frequent review

In medicine, RCTs may be repeated on different subjects, with slightly different doses or other variations. Policy-makers should be reminded of this- that RCTs can be altered to create increasingly effective policies. This continuity allows a policy to be improved upon in a systematic way.

Multi-disciplinary working

Just as those in the health sector require help from academics, statisticians and other professionals, policy-makers will also.

Ethics and confidentiality

In medicine, we know it is unethical to perform an RCT that involves a control group not being treated with an intervention that you know would be of benefit to them. Thus, groups in RCTs may compare an intervention plus baseline treatment to the control group treated with a baseline treatment alone. Similarly, in the public sector, it may be unethical to withhold a policy to a group of participants that will benefit from it. However, as demonstrated by the CAST RCT, some interventions that were thought to be beneficial may in fact be detrimental. For example, the "Scared Straight" US intervention was designed to frighten children at risk of becoming delinquents in to avoiding crime. Unfortunately, an RCT revealed the programme increased crime and as a result had a high social, as well as financial, cost.

The impact of such a trial could have been mitigated by a 'stepped wedge' approach, often necessitated due to logistical or financial constraints.[14] In medicine, 'stepped wedge' trials are also performed, for example in quality improvement exercises.[15] It is necessary to point out that the ethical issue does not exist if there is true equipoise between the two positions. Figure 1 demonstrates this process of policies being sequentially implemented to clusters or individuals over time, in a random order until all groups are allocated the intervention.

It is imperative policy-makers maintain the confidentiality of those they are trialling, especially when working in public policy. In medical RCTs, patients are provided with information about the trial and may then decide on whether to enrol. However, in public policy, this may be more complex and unavailable if the policy in question relates to educational infrastructure for example. In the vast majority of situations it may be logistically impossible or unnecessary to seek consent.

Conclusion

RCTs can provide clear benefits to policy-makers in their decision-making. However, the RCTs in question must be well set-up and run, and analysed with the help of experienced academics. RCTs should be viewed as a continuous tool rather than a one-off measure. If these requirements are fulfilled, RCTs will improve the efficacy of policies, identify weaknesses and cut costs. Policy-makers may look to the experience of those in the health sector for past mistakes to learn from, and for useful tools to adopt.

Ethical approval

No ethical approval required for this publication.

Conflict of interest

No conflicts of interest have been declared by the authors.

Author contributions

PJT: Writing, Critical Revision, Final Approval.
CFC: Concept, Critical Revision, Final Approval.

Funding

No funding source declared by authors.

1 Cabinet Office, Behavioural Insights Team. Test, Learn, Adapt Developing Public Policy with Randomised Controlled Trials. [insert year of publication here] [online]. http://www.cabinetoffice.gov.uk/resource-library/test-learn-adapt-developing-public-policy-randomised-controlled-trials (Accessed: 10th August 2012).
2 West A, Spring B. Randomised Controlled Trials. [insert year of publication here] [online]. http://www.ebbp.org/course_outlines/randomized_controlled_trials/ (Accessed: 14th August 2012).
3 Stanley K. Statistical Primer for Cardiovascular Research Design of Randomized Controlled Trials. Circulation 2007;115:1164–1169.
4 Harford T. Adapt: Why successs always starts with failure. London: Little, Brown, 2011.
5 Ruskin JN. The cardiac arrhythmia suppression trial (CAST). NEJM 1989;321(6):386–388.
6 Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM. 2nd ed. Edinburgh, United Kingdom: Churchill Livingstone, 2000.
7 Moher D, Jadad A R, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomised controlled trials: An annotated bibliography of scales and checklists. Control Clin Trials 1995;16:62–73.
8 Consort Group. The CONSORT Statement. [insert year of publication] [online]. http://www.consort-statement.org/consort-statement/ (Accessed: 18th August 2012).
9 Keirse M, Hanssens M. Control of error in randomized clinical trials. European Journal of Obstetrics & Gynaecology and Reproductive Biology 2000;92(1):67–74.
10 Donaldson C, Bate A, Mitton C, Dionne F, Ruta D. [title of article?] QJM 2010;103(10):801–807.
11 Cabinet Office. The Behavioural Insights Team. [insert date of publication] [online]. http://www.cabinetoffice.gov.uk/behavioural-insights-team (Accessed: 15th August 2012).
12 Cartwright N. Are RCTs the Gold Standard? BioSocieties 2007;2:11–20.
13 Cabinet Office. Open Data White Paper: Unleashing the Potential, CM8353. London: Stationery Office; 2012.
14 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol 2006;6:54.
15 Van de Steeg L, Langelaan M, Likema R, Wagner C. The effect of a complementary e-learning course on implementation of a quality improvement project regarding care for elderly patients: a stepped wedge trial. Implement Sci 2012;7:13.