

'The Anglo-Saxon disease'¹: a pilot study of the barriers to and facilitators of the use of randomised controlled trials of social programmes in an international context

Helen Roberts,¹ Mark Petticrew,² Kristin Liabo,³ Sally Macintyre⁴

¹General Adolescent and Paediatrics Unit, UCL Institute of Child Health, London, UK

²Faculty of Public Health and Policy, London School of Hygiene & Tropical Medicine, London, UK

³Social Science Research Unit, Institute of Education, London, UK

⁴MRC Social & Public Health Sciences Unit, Glasgow, UK

Correspondence to

Professor Helen Roberts, General Adolescent and Paediatrics Unit, UCL Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK; h.roberts@ucl.ac.uk

Accepted 28 January 2012

Published Online First
3 December 2012

ABSTRACT

Background There appears to be considerable variation between different national jurisdictions and between different sectors of public policy in the use of evidence and particularly the use of randomised controlled trials (RCTs) to evaluate non-healthcare sector programmes.

Methods As part of a wider study attempting to identify RCTs of public policy sector programmes and the reasons for variation between countries and sectors in their use, we carried out a pilot study which interviewed 10 policy makers and researchers in six countries to elicit views on barriers to and facilitators of the use of RCTs for social programmes.

Results While in common with earlier studies, those interviewed expressed a need for unambiguous findings, timely results and significant effect sizes, users could, in fact, be ambivalent about robust methods and robust answers about what works, does not work or makes no difference, particularly where investment or a policy announcement was planned. Different national and policy sector cultures varied in their use of and support for RCTs.

Conclusions In order to maximise the use of robust evaluations of public programmes across the world it would be useful to examine, systematically, cross-national and cross-sectoral variations in the use of different methods including RCTs and barriers to and facilitators of their use. Sound research methods, whatever their scientific value, are no guarantee that findings will be useful or used. 'Stories' have been shown to influence policy; those advocating the use of RCTs may need to provide convincing narratives to avoid repetition about their value.

BACKGROUND

Twenty-three researchers recently signed a paper in the *Lancet* arguing for the mandatory impact evaluation of public policies, pointing to the need for 'better use of research evidence to improve decisions about public programmes both internationally and nationally' and emphasising the lack of rigour of most evaluations.¹ In a similar vein, the House of Commons Health Select Committee noted the poor quality of much evaluation in social and public health policy in the UK:

¹This was a phrase used by the late Norman Glass in the course of his interview for this study. When asked to sign the consent form—he said breezily that he didn't go in for anonymity—he was happy to stand by his views. We therefore feel it appropriate to acknowledge him by name and to acknowledge the contribution he made more generally to this area, first in HM Treasury in the UK and subsequently as Chief Executive of the UK research agency NatCen.

"The most damning criticisms...we have heard in this enquiry [have been] of the Government's approach to designing and introducing new policies which make meaningful evaluation impossible...Even where evaluation is carried out, it is usually...little more than...asking those involved what they thought about them".²

Although the use of experimental designs such as randomised controlled trials (RCTs) is generally uncontroversial in medicine, this has not been the case in social policy circles in the UK. Arguments against RCTs of social programmes (eg, in the fields of transport, housing, criminal justice, education and early childhood development) have tended to focus on potential problems with feasibility, ethics, cost, public and professional acceptability and generalisability.³ While we do not advocate the use of RCTs for all programmes,^{4 5} we do think that many of these objections are overemphasised, particularly since some countries, including the USA, have a long history of using RCTs of social programmes.^{3 6} However, other countries have tended to avoid using controlled trials,⁷ and this raises questions about the extent to which different sectors, and different national jurisdictions, value and use different types of research. Given the considerable international variation in the use of social experiments, attempts to understand the cultural and practical barriers which policy makers and commissioners in different sectors face in the use of research evidence internationally may be useful. Lessons might be learnt from the implementation of RCTs in different national contexts.

As part of the International Collaboration for Complex Interventions (<http://www.intervention-research.ca/>), we conducted a pilot study to assess the extent to which it is possible to (a) identify how many RCTs have been undertaken of social policy programmes in different countries and (b) interview public policy makers and advisors in a range of different countries about the use of RCTs for social programmes. Our review of the prevalence of RCTs showed numerous examples across a wide range of social- and health-related programmes (eg, injury prevention, school feeding, day care for school-age children, delinquency prevention) but wide variation in their prevalence between sectors and particularly between nations.⁸

Here, we report on findings from the interviews, which were designed to collect qualitative data exploring the conditions under which RCTs may and may not be feasible: the barriers to and



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://jech.bmj.com/site/about/unlocked.xhtml>

facilitators of the development of new trials and the ways in which different kinds of evidence are valued within policy sectors including health, criminal justice, education and social welfare. These interviews were an extension of our previous work, which had examined policy makers and researchers' experiences of the use of different types of evidence in public health in the UK.^{9 10} While the debate about the place of RCTs in evaluating social policies is not new, we focus on the potential added value of RCTs as opposed to other forms of research and how this is perceived in different areas of public policy, and in different countries. Interviews with elites in this field are still relatively rare, as is exploration of differences in the acceptability of social experiments between sectors and between countries.

METHODS

For this pilot, we selected the USA, Canada, Australia, New Zealand, England and Scotland because the former two countries made most use of RCTs and the UK least and Australia and New Zealand were in the middle. Their political and welfare systems varied and we could easily identify key policy makers and researchers in these countries to interview. An open-ended interview topic guide was designed by KL, amended in discussion with the team and adapted at interview according to whether the respondent was a policy maker, commissioner of research or researcher. Further details can be found in our report.⁸

Policy is influenced not just by officials and politicians but also by researchers with the ear of policy makers and/or whose results have been useful or used in the past. Our sample was selected to include (a) those in a position to influence policy (including funding and research policies) and (b) individuals we considered to be familiar with the extent to which different research methods are (or in some cases are perceived to be) more or less appropriate to support decision making. Funders can have a substantial effect on evaluative methods: for instance European Union funding differs from North American funding in its emphasis on process evaluations as compared with trials.¹¹

We approached 15 individuals for an interview. None refused to participate, but a firm commitment was not forthcoming within our timescale from five interviewees. We interviewed 10 individuals from the six countries: six by telephone and four face to face. They were all professionals in the public sector and worked in a range of fields including criminal justice, education, public health and social care. Eight were involved with policy or research commissioning and two were senior researchers, one of whom had also been involved in commissioning (though not simultaneously). Interviews were audio taped and transcribed, with participant consent, and the transcripts were read by at least two of the researchers who agreed on emergent themes. In this pilot study, there were too few countries, policy sectors and policy advisory roles represented to undertake systematic comparisons between countries, policy sectors or roles. The main barriers and levers to the use of trials are described in table 1. Here, we restrict our findings to interviewees' observations on different policy sectors and the relative importance of different types of evidence.

To protect confidentiality, the quotations below do not give identifying details. However, they represent all 10 interviewees and all six countries. Identifying information is provided as appropriate on sector.

FINDINGS

The main finding from our pilot interviews was that it was possible and informative to interview relatively elite policy

advisers, in a range of countries, and that useful insights on barriers and facilitators to the use of RCTs could be obtained in this way.

What is the 'added value' of RCTs to users compared with other study designs?

All interviewees were asked about the extent to which the methodological robustness of RCTs is valued compared with other study designs and other types of information.

One policy advisor in education spoke in terms of a "Sliding scale; at the bottom end an unvalidated advocacy message...might instigate further research, but the scale goes from this to the...RCT...or systematic review or meta-analysis."

A senior researcher and user of research in public health, experienced in evaluating policy and in liaising with policy makers, spoke of the power of trials to influence policy makers: "I would tend naturally to have more confidence in the results,... assuming that it wasn't just the design but the implementation of it that was satisfactory...the advantage is that even politicians would tend to be influenced by something that was convincingly a controlled trial."

Another policy advisor in social care concurred, noting the advantages over observational methods: "...it's difficult to get the high quality of analysis other than through an RCT—where possible it should be an RCT."

A senior manager responsible for policy development in the field of education also claimed that well-controlled experiments: "...do tend to solve arguments.... People respect them." However, s/he went on to note that this did not always apply to researchers in government "Up until a wee while ago, our research division was very unsympathetic, if not downright antagonistic to, randomised trials." A related point was made by a social and public policy adviser, to the effect that researchers do not give policy makers clear advice on when to use particular research methods, and experiments were often downplayed:

"Policy makers are getting...rather muffled messages about when to do a trial or...when some other method will do...so it's hardly surprising that they are...quite happy to...go on using weaker methods,...they're not getting a clear steer...the choices they're given don't involve the option of running a trial. ...There are...quite influential papers about the methodology in evaluating social interventions which give policy makers a lot of rope to hang themselves."

However, in the experience of a senior education/health researcher, who had designed a number of innovative RCTs of social interventions, it was not trial methodology but instead the ability to attribute costs and benefits to interventions that mattered: "It's still the calculus of policy." Asked whether the findings of the RCT alone would have had the same weight, s/he continued "some, but the cost-benefit analysis was the big factor...a number of ministers have said it made a big difference to them being able to argue the case."

Sound methods ≠ useable findings

The senior official in education quoted above was someone who needed to use research but was critical of the sort of research knowledge s/he received: "A lot of research knowledge is not helpful from my perspective." While this might seem to imply that 'stronger' methods would be more useful, this was not necessarily the case. Even among vigorous advocates of trials, it was clear that sound methods did not necessarily lead to useable findings:

"As you move up that scale, there [are] more useful 'findings' in terms of their scientific validity. Whether they translate into a discourse that you can hold with policy makers is another question."

Table 1 Barriers to and facilitators of social policy trials identified with interviewees

Levers	Barriers
Personal contacts/researcher policy contact/serendipity	Poor communication by researchers; ambivalence/hostility by researchers and research brokers
Potential for good cost-benefit information	
Independence of evaluators from policy/politicians	Problems if policy initiative to which politicians have committed themselves shown not to work
Funding for new initiatives tied to good trial evidence/accountability	Cost/complexity of running a trial
Advocacy by those whose trials have shown an effect in other countries	Ambivalence/hostility by some researchers and research brokers
Good dissemination skills by trialists; willingness to avoid too many caveats when presenting results	Over-enthusiasm by some trial proponents
A lot of good research on what the problems are; less on what to do about them	Pejorative use of term 'experimentation'
Convincing trial welcome to politicians	Lack of high-quality trial applications
Support from key government departments (eg, Treasury)	Moral and ethical concerns/equipoise
	Lack of researcher experience in social policy trials
	Recruitment problems
	Timing (in relation to policy development)/political desire to get things up and running quickly
	RCTs more suited to clinical research
	The line of least resistance not to carry them out
	Culture of advocacy, case study, precedent and anecdote

RCTs, randomised controlled trials.

A policy advisor, experienced in using and commissioning RCTs in the field of healthcare, and an advocate of their wider use for social policy concurred: *"By and large, methodology is a weak influence in the sense that policy makers don't really tend to weigh up research evidence in terms of the strength of the source, it's much more the signal that they're interested in..."*

S/he felt that policy makers tended to prefer *"very small scale studies, pilots, rather informal evaluation evidence where it supports what they're interested in doing, and [they are]...quite resistant to the much stronger evidence where it doesn't support what they think."*

Another policy advisor, who had worked in several government departments and was a strong advocate of the greater use of social experiments, was similarly less than optimistic about the impact of trials:

"Certainly in [my country], the power of a story beats almost anything...If researchers would find a story to tell about their RCT, or personalise it...If you're dealing with...politicians, you have to...appeal with a story."

One reason for the underuse of RCTs which emerged from the interviews was that paradoxically, the straight answers described by some interviewees as useful in settling questions could be perceived as unhelpful—particularly when they show that favoured interventions do not 'work':

"...if the results tell you that your intervention isn't working then you're in trouble...to some extent, people would rather have vaguer information about processes, which...carries less risk...I mean, people like the idea of the process of continuous quality improvement with evaluation, contributing something to improve the way you implement your...new policy or your intervention, and I think, to some extent, that's preferred to evidence which...tell(s) you pretty starkly that you ought to stop and that you're wasting public money."

A government research commissioner in the area of social care observed: *"You can't necessarily say it's one type of research over another type, because the type of design depends on the kind of research question you're asking. Different types of research will be valuable in different contexts."*

Specific barriers to and facilitators of the use of RCTs, as opposed to evidence more generally, about which much has already been written^{12 13} were also identified. The perceived lack

of flexibility of RCTs, particularly in relation to the adaptability of programmes by practitioners, plus high costs and long timescales were referred to in several cases.

Interviewees suggested that RCTs are underused because users are more interested in evaluation being used strategically to demonstrate policy interest:

"There are quite complicated reasons for commissioning evaluations and they're not all about testing how things work. A lot of them are to do with...demonstrating that you are taking the issue seriously."

This political function of evaluation was highlighted by a former Treasury official who pointed to the importance of presentation and language:

"There was ministerial reluctance to appear to be just trying something out, or not to give something to one group who might be equally eligible. We did point out that...we did this kind of thing anyway in practice, but they were happier if we called it a pilot rather than an experiment or a trial."

One UK policy adviser made an interesting point about the intellectual background of politicians:

"versed in the law and advocacy and case study and precedent, rather than science..."

Asked for any examples of unsuccessful attempts to set up policy RCTs, an interviewee drew on the example of an early years intervention involving early education, childcare, health and family support where the evaluation funded was not a RCT: *"All the scientists were saying it should be an RCT, but...in this country, the service delivery people don't have the faith that it's important to evaluate things in a very rigorous way, and they felt it was more important to have services which could be adapted to local situations."*

The use of experiments by different sectors/countries

One policy maker who had worked across different sectors noted that:

"Health...says it takes—randomised trials much more seriously than other sectors. It certainly takes evidence more seriously...[Problem area] certainly doesn't rely on high quality [evidence], in this sector, we do"

things, futile things, inappropriate things, [more] than even the health sector or the housing sector or any other, because we're always in a rush. Nobody [in this area] prides themselves on being an expert [in the] evidence-based sector...It's a bizarre system...it's much more dysfunctional than health...the primacy of research is not there."

A commissioner and user of research suggested, however, that while there was apparently greater use of RCTs in health, this was accompanied by a range of other types of evidence:

"Public health professionals in particular are used to working without RCT evidence, for example when there is a disease outbreak, or some other crisis. There, decisions tend to be driven by theoretical constructs tested in related health issues, by basic biological evidence, plus aetiological evidence, and evidence of what's going on in the community, plus evidence on behaviour change—that is, a series of sources of information/evidence."

Reflecting on the lesser use made of RCTs in sectors other than health, another interviewee said that in [country], RCTs are used as a tool to cut funding, rather than to simply identify 'what works':

"The [country] Department of Education...led to a push for more RCTs in education...Now, RCTs and systematic reviews are used as a way of cutting programmes—where there is no good evidence from RCTs, it is used to justify a cut." S/he expanded this theme of cultural differences in use of experimentation, drawing comparisons between countries:

"[It] relates to levels of affluence and the degree of development of the scientific community but in Europe my impression is that Northern Europe has done a lot more in terms of social and health research of an organised type...In the Scandinavian countries you have a tradition of really being...organised and imposing quite high degrees of control over the population in terms of what people can and can't do, and gathering a lot of data on a large scale...whereas in Southern Europe, they have less of a history of social public health trials. The United States and Canada in terms of volume (not necessarily quality) is far ahead of the rest. Australia has done quite a bit given the size of the country as well."

S/he went on to explain why this may be so; in some countries, evaluation is important for public accountability:

"In [our country], we have a very strong tradition of evaluation research and population-based epidemiology. This is linked to the need for accountability for performance—as opposed to seeking harder evidence of effectiveness..."

Another interviewee noted that educational researchers' "methodological tool-kits tend to be in other areas" and that they assume that RCTs are only of value in medicine.

The single trialist we interviewed noted that many countries placed more weight on studies from abroad, and publication in a US journal was often more prestigious. S/he suggested that one argument posed against trials in the education setting was that it is not fair to deprive someone of an intervention which they perceive as being effective.

The US emphasis on trials was also underlined by a former policy maker:

"Very strong in the US and possibly in Canada—I have that impression. We're somewhere in the middle. The Europeans are nowhere. There's no interest in continental Europe. It's an Anglo-Saxon disease."

Training in appropriate skills was identified as a problem in several cases, for example, a UK Research Commissioner said:

"I think I find it quite dispiriting that in America, they will invest in these really rigorous studies, and yet in this country, we don't. There's a problem with research capability in this country because people don't develop the skills to do it."

Finally, interviewees reflected on what additional information is needed beyond RCTs. Suggestions included studies that permit comparisons between countries and studies which describe context: *"Useful information includes studies that illuminate the extent and the nature of the problem. [Our country] pays a lot of attention to the PISA study, by the OECD, which includes 40 countries...comparative studies are helpful."*

DISCUSSION

This small pilot suggested considerable diversity between countries and between sectors in the experience of and attitudes towards the use of RCTs for social programmes. Political cultures, both in the sense of national jurisdictions and particular disciplines and sectors, were seen to shape the perceived acceptability and desirability of, and responses to, RCTs.

Arguments against RCTs of social programmes have tended to focus on potential problems with feasibility, ethics, cost, public and professional acceptability and generalisability—arguments to which Macintyre,³ Oakley¹⁴ and others have responded. Our preliminary findings suggest that if RCTs are to be used more generally, additional concerns may need to be addressed.

McKee¹⁵ has described the influence of political ideologies on the conduct and use of RCTs in medicine, and we know that there are political impediments to robust evaluation in many healthcare areas.¹⁶ Our interviews suggest that there are similar influences on the conduct of social policy RCTs, with there being arguments in particular in education about appropriate evaluation methods. While not wanting to re-invent the wheel, sectors may prefer their *own* wheel. This may account in part for the common hostility to trials in other sectors, stemming from their view of RCTs as being 'over medical' (despite their early use in the social sciences⁶).

Our interviews pose challenges for advocates of robust evaluation methods. Much debate about evidence-based public policy and gaps in the evidence base in public health and elsewhere are predicated on the assumption that there is a supply-side problem: researchers have failed to do the right kind of evaluation in the past. This may well be true, but our interviews illustrate that there may also be significant problems on the demand side. Users do not always want robust methods because they do not always want robust answers about what works. Thus, the production of better evidence alone will not necessarily lead to its uptake (and in any case, political values and other factors legitimately play a role in policy decision making).

Previous studies have also pointed to the proliferation of terms to describe evaluation studies. Walker and colleagues¹⁷ in describing the implementation of a social experiment, the Employment Retention and Advancement (ERA) trial refer to a 'cacophony of names': pilots, pathfinders, experiments and so on, suggesting that this may be less to do with capturing the richness of evaluation methods than with obscuring the experimental (in the non-scientific sense of the word) nature of most public policy. According to a report on 'pilots' in UK policy-making, civil servants and ministers may themselves sometimes be confused about the distinctions between different policy-testing mechanisms. That report also described how a Minister had been given the option to choose a name that s/he liked best for a 'pilot' from a range of options.¹⁸

While the study we report here was planned only as a pilot, it provides a contribution in moving the debate beyond the need to produce robust evidence, important though that is to how its value to users may be enhanced and how the use of

What is already known on this subject

- ▶ Randomised controlled trials (RCTs) are more generally accepted in clinical medicine than in public health and much less often used for programmes in sectors such as social care, transport, criminal justice, housing and education which may influence health.
- ▶ There are cross-national, and policy sector, differences in the use of RCTs for evaluating social programmes.
- ▶ 'Stories' may influence policy-making more than methodologically robust evidence.

What this study adds

- ▶ It is possible to generate useful insights about cross-national and policy sector differences in the use of RCTs, which might help elucidate the barriers and facilitators for RCTs and the context in which they might be acceptable and useful.
- ▶ If advocating for RCTs for social programmes, account would need to be taken of political cultures in different jurisdictions and the cultures of particular disciplines and sectors.
- ▶ Clear results (often presented as a selling point for the use of RCTs) may, in fact, be a barrier, particularly where they cast doubt on a substantial investment or a policy announcement already made or planned.
- ▶ 'Stories' about the practical value of RCTs may be more convincing than detailed methodological arguments about their merits.

RCTs may be encouraged. Much previous research in this field has noted the impact of a 'good story'^{9 10 19 20} or 'killer facts'²¹ and the fact that those on the receiving end of policies and services may also prefer to present their data through stories.²² It may be that if the debate on the use of research evidence to inform policy is to move beyond the academic world, convincing 'stories' and 'killer facts' need to be provided by those advocating the use of RCTs to researchers, policy advisers, politicians and those on the receiving end of social programmes, rather than complex methodological arguments. An example would be the 'story' of the 'scared straight' programmes in the USA, where all the process and user reported information suggested that it was highly effective, but a meta-analysis of seven RCTs showed it to be counter productive.²³

This pilot suggests that it would be feasible and instructive to undertake a more extensive study, systematically comparing countries and sectors in relation to the use of RCTs and their barriers and facilitators.

Acknowledgements We are grateful to our interviewees for their time and thoughtfulness and to the reviewers for their constructive comments.

Contributors All authors were fully involved in the design of the study, the authors each conducted interviews and all were involved in the data analysis and preparation of this article.

Funding SM is funded by the UK Medical Research Council (MRC). This piece of work was funded by the MRC Social & Public Health Sciences Unit (Reference MC_US_A540_0070) and the International Collaboration on Complex Interventions (ICCI). ICCI was funded by the Canadian Institutes of Health Research.

Competing interests None.

Ethics approval Ethics approval was obtained from the Faculty of Children and Health Research Ethics Committee at the Institute of Education, University of London.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Oxman AD**, Bjørndal A, Becerra-Posada F, *et al*. A framework for mandatory impact evaluation to ensure well informed public policy decisions. *Lancet* 2010;**375**:427–31.
2. **House of Commons Health Committee**. *Health Inequalities (Third Report of Session 2008–09)*. London: The Stationery Office, 2009.
3. **Macintyre S**. Good intentions and received wisdom are not good enough: the need for controlled trials in public health. *J Epidemiol Community Health* 2011;**65**:564–7.
4. **Craig P**, Dieppe P, Macintyre S, *et al*. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655.
5. **Petticrew M**, Roberts H. Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Community Health* 2003;**57**:527–9.
6. **Oakley A**. Experimentation and social interventions: a forgotten but important history. *BMJ* 1998;**317**:1239–42.
7. **Macdonald G**, Roberts H. *What Works in the Early Years*. Ilford: Barnardos, 1996.
8. **Roberts H**, Petticrew M, Macintyre S, *et al*. *Randomised Controlled Trials of Social Interventions: Report of a Pilot Study of Barriers and Facilitators in an International Context*. MRC Social and Public Health Sciences Unit Occasional Paper. Glasgow: MRC Social and Public Health Sciences Unit, 2008.
9. **Petticrew M**, Whitehead M, Macintyre SJ, *et al*. Evidence for public health policy on inequalities: 1: the reality according to policymakers. *J Epidemiol Community Health* 2004;**58**:811–16.
10. **Whitehead M**, Petticrew M, Graham H, *et al*. Evidence for public health policy on inequalities: 2: assembling the evidence jigsaw. *J Epidemiol Community Health* 2004;**58**:817–21.
11. **Brady B**, O'Regan C. Meeting the challenges of doing an RCT evaluation of youth mentoring in Ireland: a journey in mixed methods. *J Mix Methods Res* 2009;**3**:265–80.
12. **Campbell S**, Benita S, Coates E, *et al*. Analysis for policy: evidence-based policy and practice, Government Social Research Unit, London: HM Treasury, Crown copyright, 2007. http://www.civilservice.gov.uk/wp-content/uploads/2011/09/Analysis-for-Policy-report_tcm6-4148.pdf
13. **Weiss C**. *Evaluation: Methods for Studying Programs and Policies*. Upper Saddle River, NJ: Prentice-Hall, 1998.
14. **Oakley A**. Experimentation in social science: the case of health promotion. *Soc Sci Health* 2000;**4**:73–89.
15. **McKee M**. Cochrane on Communism: the influence of ideology on the search for evidence. *Int J Epidemiol* 2007;**36**:269–73.
16. **Sowden SL**, Raine R. Running along parallel lines: how political reality impedes the evaluation of public health interventions. A case study of exercise referral schemes in England. *J Epidemiol Community Health* 2008;**62**:835–41.
17. **Walker R**, Hoggart L, Hamilton G. Observing the implementation of a social experiment. *Evid Policy* 2008;**4**:83–203.
18. **Jowell R**. Trying it out: the role of "pilots" in policy-making. *Government Chief Social Researchers Office*, Crown copyright, 2003. http://www.civilservice.gov.uk/wp-content/uploads/2011/09/Trying-it-Out_tcm6-36824.pdf
19. **Greenhalgh T**, Russell J. Reframing evidence synthesis as rhetorical action in the policy making drama. *Healthc Policy* 2006;**1**:34–42.
20. **Stevens A**. Telling policy stories: an ethnographic study of the use of evidence in policy-making in the UK. *J Soc Pol* 2011;**40**:237–55.
21. **Bowen S**, Zwi A, Sainsbury P, *et al*. Killer facts, politics and other influences: what evidence triggered early childhood intervention policies in Australia? *Evid Policy* 2009;**5**:5–32.
22. **Graham H**. Surveying through stories. In: Bell C, Roberts H, eds. *Social Researching: Policies, Problems and Practice*. London: Routledge and Kegan Paul, 1984:104–24.
23. **Petrosino A**, Turpin-Petrosino C, Finckenauer J. Well-meaning programs can have harmful effects! Lessons from experiments of programs such as scared straight. *Crime Delinq* 2000;**46**:354–79.