# What counts as evidence? (Some reflections on) Trials and other sources of evidence in public health

Mark Petticrew
Public and Environmental Health Research Unit
Dept. of Public Health & Policy
LSHTM

- The role of evidence, and RCTs, and other types of evidence in public health; why evaluating effectiveness can be difficult
- This talk takes a “wider public health” perspective – i.e. focussed on interventions addressing the social determinants of health (e.g., housing, transport, employment, education)
- How trials are seen in other sectors
- Some of the barriers to “social” RCTs
- A bit of a rant about the persistence of simplistic views of trials and the choice of appropriate study designs

- Lack of RCTs in public health is often rehearsed (e.g. Wanless)
- The reasons are often said to be:
- Practical (e.g., difficulty in randomising large scale interventions; finding enough units of randomisation; apparent lack of flexibility; expense; contamination)
- Political and ethical (e.g., withholding an apparently beneficial intervention from a control group)

- Other reasons exist; one is that the concept of “effectiveness” can be more difficult to apply to some social interventions

- “What Works” is only part of the story
- How it works, who it affects, and how
- “Outcome evaluation thinking” can be difficult to apply to complex social interventions (e.g. social policies)
- Such interventions do not simply “work” or “not work”
- Identifying a primary outcome can be difficult - there are multiple equally valid “primary” outcomes – valued by different stakeholders

## An example from social housing improvement: the SHARP study

Assessed health **BEFORE** ***AND*** **AFTER** move

**1 year**

- Assess change in self-reported health at baseline and 1 & 2 years after house move
  - » in the same residents
  - » Matched control group (n=330)

## Evaluating housing and health

- Does housing “work”? – the question only makes sense with respect to an outcome
- In public health the question might be: Is housing renewal effective in improving health?
- Or, (better) is social housing improvement (moving tenants from old, damp homes into warm, dry, new build homes) effective in improving respiratory health?
- A good, clear question about effectiveness, amenable to an RCT or similar design

## It might be a good research question, but it misses the point

- What if better housing didn't improve health? We would not advocate disinvestment in better housing
- Still valuable information, but housing is not built to improve health - there are other (more important) outcomes. E.g., "amenity," social justice
- But most public health housing research has until recently been health-focussed (respiratory symptoms, injuries)
- What sort of evidence is valued in housing policy circles?

- In housing and other areas of research, non-experimental methods including case studies and cross-sectional surveys predominate
- The level of evidence required to justify action differs from say the health sector
- In the case of demolition and rebuilding poor, damp social housing, the primary outcome is provision of new housing - not health improvement
- So there are relatively few “true” experiments

**Toryglen, Glasgow**

# This makes sense

- You don’t need an RCT to tell whether a house, or a housing estate has been built, or not
- You don’t even need research
- You just need good audit procedures to make sure that the money has been spent properly by the local council (and not salted away in a dodgy Icelandic bank)

## Producing “better” public health evidence…

- Requires a better understanding of how, and why evidence is used in different policy sectors
- And an understanding that different sectors are at different stages in the development of the evidence base; in some, observational and other forms of evidence predominate
- The public health/housing evidence base is becoming slowly more experimental; there are now three RCTs

# NZ Housing, Insulation & Mould Study *

- Single-blind cluster RCT of retrofitting insulation in homes with an occupant with a respiratory illness
- N=1400 homes, 7 locations in NZ, 4415 residents, 49% Maori, 25% Pacific
- Trained teams retrofitted ceiling and underfloor insulation, polythene ground cover and draft stopping around windows and doors -$2000 per house

*** Howden-Chapman P et al. Effect of insulating houses on health inequality: cluster randomised study in the community. *BMJ* 2007; 334:460**

# The findings of the NZ RCT

- Small increase in bedroom temperatures during the winter (0.5°C) and decreased relative humidity (–2.3%) - despite energy consumption being reduced to 81% of that in uninsulated houses.

- Bedroom temperatures were below 10°C for 1.7 fewer hours each day in insulated homes

- Reduced odds of:
- - Fair/ poor self rated health (AOR 0.50, 95% CI 0.38 to 0.68)
- - Self reports of wheezing in the past three months (0.57, 0.47 to 0.70), - Self reports of children taking a day off school (0.49, 0.31 to 0.80)
- - Self reports of adults taking a day off work (0.62, 0.46 to 0.83).
- - Visits to general practitioners (0.73, 0.62 to 0.87).

- CBA: The health and energy benefits outweighed the costs by about 2, when calculated at a 5% real discount rate over 30 years

- Although the evidence base is weak in terms of outcome evaluations (e.g., Millward, Kelly et. al., 2003), it is improving, but slowly

- But we can't afford to dismiss the existing evidence base as "methodologically weak", and wait until more RCTs are published.

- In the case of housing, at the current rate of publication of 1 RCT every 20 years, we'll have to wait until about 2209 to have enough for an informative meta-analysis (about 10 studies)

## What public health policymakers say they don't like:

- "Policy-free evidence" - research that does not answer clear, or policy relevant questions

- "Researcher naïveté" of the policy environment (which militates against knowledge transfer between science and policy )

- Evidence from far down the causal chain, (often concerned with health behaviours and clinical issues, not with broader social determinants of health)

***J Epidemiol Community Health*** **2004 58(10):811-6**

- “Researchers are overly concerned with critically appraising internal validity (bias)”

- “Not concerned enough with assessing whether research evidence is generalisable”

# Except . . .

- There is a problem with the signal/noise ratio. In many fields relevant to public health there are few trials and indeed few evaluative studies, but a range of other types of information which decisionmakers need to use

- This information is of varying relevance or value to different stakeholders

- It can be difficult to sift, prioritise and integrate this information and produce something relevant to decision-making

- We have to get better at using and integrating the experimental and non-experimental evidence we *do* have, rather than rehearsing the bad news about what we *don't* have

- We also need to communicate better how to identify when RCTs are feasible and appropriate, when other forms of research are informative

- Recognise the complexity of evidence use in different sectors

- Accept that the issue is not simply a lack of RCTs; it is a lack of outcome evaluation *of any type*

- …a better understanding of the types of evidence used to justify action in different sectors

# Why is this important?

- Because however much research money is made available, there will never be “enough” evidence (and we have no definition of “enough”)
- There will never (for all the practical, ethical, political reasons outlined previously) be “enough” RCTs of public health interventions
- One further challenge is “social” RCTs are often seen as non-generalisable, while the same criticism is not applied to other types of research
- Developing methods of assessing the generalisability of different types of public health research is a priority

House of Commons

Health Committee

# Health Inequalities

## Third Report of Session 2008–09

*Volume I*

*Report, together with formal minutes*

*Ordered by the House of Commons*
*to be printed 26 February 2009*

# Lost opportunities: Conclusions from HoC Select Committee 3rd report

"The most damning criticisms of Government policies we have heard in this inquiry have not been of the policies themselves, but rather of the Government's approach to designing and introducing new policies which make meaningful evaluation impossible. As one witness described, "there is a continual procession of area-based initiatives and that in itself is quite disruptive. Nothing is given time to really bed in and function" (MW)

House of Commons
Health Committee

“Even where evaluation is carried out, it is usually “soft”, amounting to little more than examining processes and asking those involved what they thought about them. All too often Governments rush in with insufficient thought, do not collect adequate data at the beginning about the health of the population which will be affected by the policies, do not have clear objectives, make numerous changes to the policies and its objectives and do not maintain the policy long enough to know whether it has worked. As a result, in the words of one witness, ‘we have wasted huge opportunities to learn’.”(KJ)

“Governments have spent large sums of money on social experiments to reduce health inequalities, but we do not know whether these experiments have worked or whether the money has been well spent… All the reforms we have discussed are experiments on the public and can be as damaging (in terms of unintended effects and opportunity cost) as unevaluated new drugs or surgical procedures. Such wanton large-scale experimentation is unethical, and needs to be superseded by a more rigorous culture of piloting, evaluating and using the results to inform policy.”


Roberts et al. 2008
Occasional paper
(No.19)

**www.sphsu.mrc.ac.uk**

Interviews with researchers and policy advisors in UK and other countries on the use of RCTs, and how they are viewed by policymakers; whether they are valued or if not, what types of evidence holds value instead

# Politically unhelpful evaluation

- “… there are costs to having really good evaluations not just the financial cost, they do cost more obviously but you know, if … it’s really good and the results you know tell you that your intervention isn’t working then you’re in trouble, and I think to some extent …people would rather have you know vaguer information about processes, which …carries less risk of being hostages to fortune to some extent…

- “…I mean, people like the idea of the process of continuous quality improvement with evaluation, you know, 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.” (Policy advisor, UK)

- “Certainly in British politics, the power of a story beats almost anything.” (Policy advisor, UK)

- This suggests that the well-rehearsed ethical, methodological and other barriers are only part of the story

- “Softer,” descriptive evaluations are also more expedient/more acceptable

1. Bad, controlling, reductionist triallists, obsessed with measuring things, probably while wearing a monocle and stroking a white cat and speaking in a European accent, who want to control the world,

*compared with*

2. Complex, holistic thinkers who have sole responsibility for defending the free world from the above by using their understanding of complex causal pathways, processes and mechanisms, but without constraining their beauty by measuring or controlling them.

**RCT-lover. And cat.**

- In this black and white world, there is nothing in-between:

- No triallists who do also qualitative research;
- No RCTs of social interventions with qualitative process evaluations;
- No qualitative researchers involved in trials

# So. What counts as evidence?

- It is dependent on where you stand as a user, and as a researcher
- Improving the public health evidence base is not simply a matter of advocating that we do more RCTs. We also need more outcome evaluation (with appropriate evaluations of processes and contexts)
- Evaluations of complex interventions are themselves complex, requiring knowledge of theory, causal pathways, weighing of multiple outcomes, dealing with competing stakeholder perspectives… & an RCT *alone* is insufficient to deal with this.
- This also requires understanding the use, value, and production of evidence within different sectors
- And building capacity and understanding in evaluation methods, and seeking to overcome simplistic notions about evaluation, and evaluation methods, and their complementarity
- And that's all.