Theory and methods

Using natural experiments to evaluate population health interventions: new Medical Research Council guidance

Peter Craig,1 Cyrus Cooper,2 David Gunnell,3 Sally Haw,4 Kenny Lawson,5 Sally Macintyre,6 David Ogilvie,7 Mark Petticrew,8 Barney Reeves,9 Matt Sutton,10 Simon Thompson11

ABSTRACT

Natural experimental studies are often recommended as a way of understanding the health impact of policies and other large scale interventions. Although they have certain advantages over planned experiments, and may be the only option when it is impossible to manipulate exposure to the intervention, natural experimental studies are more susceptible to bias. This paper introduces new guidance from the Medical Research Council to help researchers and users, funders and publishers of research evidence make the best use of natural experimental approaches to evaluating population health interventions. The guidance emphasises that natural experiments can provide convincing evidence of impact even when effects are small or take time to appear. However, a good understanding is needed of the process determining exposure to the intervention, and careful choice and combination of methods, testing of assumptions and transparent reporting is vital. More could be learnt from natural experiments in future as experience of promising but lesser used methods accumulates.

INTRODUCTION

Natural experimental studies are often recommended as a way of understanding the impact of population-level policies on health outcomes or health inequalities.1–4 Within epidemiology there is a long tradition, stretching back to John Snow in the mid nineteenth century,5 of using major external shocks such as epidemics, famines or economic crises to study the causes of disease. A difficulty in applying similar methods to the evaluation of population health policies and interventions, such as a ‘fat tax’ or a legal minimum price per unit of alcohol, is that very often the change in exposure is much less extreme, and its effects may be subtle or take time to emerge. Although they have certain advantages over planned experiments, for example by enabling effects to be studied in whole populations,6 and may be the only option when it is impossible to manipulate exposure to the intervention, natural experimental studies are more susceptible to bias and confounding. It is therefore important to be able to distinguish situations in which natural experimental approaches are likely to be informative from those in which some form of fully experimental method such as a randomised controlled trial (RCT) is needed, and from those in which the research questions are genuinely intractable.

The Medical Research Council (MRC) has recently published guidance to help researchers and users, funders and publishers of research evidence make the best use of natural experimental approaches to evaluate population health interventions (http://www.mrc.ac.uk/natural-experiments-guidance). Following the model of the MRC complex interventions guidance,7 it was written by a multidisciplinary team with experience of evaluation using a wide range of research designs. The ideas were developed and tested in two specially convened workshops of population health researchers. Drafts were reviewed by workshop delegates and by the MRC’s Methodology Research Panel. The guidance is meant to help researchers to plan and design evaluations of public health interventions, journal editors and reviewers to assess the quality of studies that use observational data to evaluate interventions, and policy-makers and others to recognise the strengths and limitations of a natural experimental approach. In this paper we summarise the main messages of the guidance.

WHAT ARE NATURAL EXPERIMENTS?

The term ‘natural experiment’ lacks an exact definition, and many variants are found in the literature.3–10 The common thread in most definitions is that exposure to the event or intervention of interest has not been manipulated by the researcher. Outside an RCT it is rare for variation in exposure to an intervention to be random, so special care is needed in the design, reporting and interpretation of evidence from natural experimental studies, and causal inferences must be drawn with care.

WHY ARE NATURAL EXPERIMENTS IMPORTANT?

Alternatives to RCTs have been advocated by policymakers and researchers interested in evaluating population-level environmental and non-health sector interventions11 and their impact on health inequalities.5 Such interventions may be intrinsically difficult to manipulate experimentally—as in the case of national legislation to improve air quality, or major changes in transport infrastructure12—or be implemented in ways that make a planned experiment difficult or impossible, for example with short timescales or extreme variability in implementation.13 It may also be unethical to manipulate exposure in order to study effects on...
When should natural experiments be used?
The case for adopting a natural experimental approach is strongest when: there is a reasonable expectation that the intervention will have a significant health impact, but scientific uncertainty remains about the size or nature of the effects; an RCT would be impractical or unethical; and the intervention or the principles behind it have the potential for replication, even if the expected effect size is small, and inform the content of suicide prevention strategies in the UK and worldwide.

Design
A study protocol should be developed, and ideally published, whatever design is adopted. Good practice in the conduct of observational studies, such as prior specification of hypotheses, clear definitions of target populations, explicit sampling criteria, and valid and reliable measures of exposures and outcomes, should apply equally to natural experimental studies.

Analysis
The defining feature of a natural experiment is that manipulating exposure to the intervention is impossible. There are a few examples where assignment is by a ‘real life’ lottery, but selection is the rule and a range of methods is available for dealing with the resulting bias.

Where the factors that determine exposure can be measured accurately and comprehensively, matching, regression and
propensity scores can be used to reduce confounding (box 2). Bias will remain if there are unobserved or imperfectly measured factors that influence both exposure and outcomes. Given the difficulty of measuring accurately all of the characteristics associated with exposure to an intervention, methods such as difference in differences, instrumental variables and regression discontinuity designs that deal with unobserved factors are a potentially valuable advance on those that only deal with observed factors (box 1).

In practice, none of these approaches provides a comprehensive solution to the central problem of selective exposure to the intervention.20 Methods of controlling for observed factors associated with exposure are vulnerable to selection on unobservables. Methods for dealing with selection on unobservables require strong and untestable assumptions28 and their use is restricted by the often very limited availability of variables that can be used to model exposure. They are therefore best used in conjunction with additional tests for the plausibility of any causal inferences.

Combining methods that address different sources of bias and comparing the results is one such approach and there are several examples in supplemental table 1. In their evaluation of a conditional cash transfer scheme to encourage women to use health facilities to give birth, Lim et al30 combined methods for dealing with selection on both observable and non-observable characteristics. Another useful technique is to analyse outcomes that are not expected to change. Dusheiko et al31 used trends in emergency admissions as a non-equivalent dependent variable to test whether changes in elective admissions could plausibly be attributed to GP fundholding, while Ludwig and Miller28 compared mortality from causes that were likely or unlikely to respond to Headstart services.

Given the difficulty of eliminating bias, single studies are unlikely to be definitive. Replication and careful synthesis of evidence across studies will be needed to support confident inferences about effectiveness. Exact replication of a natural experiment is unlikely, but partial replication is often possible and may be more informative. Consistent findings from studies using varying designs makes it less likely that common biases are present, and consistent findings across settings or populations increase confidence in the generalisability of causal inferences. For example, a number of studies in different countries have shown that legal restrictions on smoking in public places reduce hospital admissions for heart attacks. Although the size of the effect varies widely, as might be expected given variation in smoking rates and the extent of partial restrictions prior to outright bans, the balance of evidence suggests a real effect.22

Reporting
Transparent reporting of natural experimental studies is vital. Established guidelines such as STROBE32 should be followed, with particular attention to: clearly identifying the approach as a study of a natural experiment; providing a clear description of the intervention and the assignment process; and explicitly stating the methods used to estimate impact. Procedures used to reduce bias should be discussed in a detailed and balanced way. Ideally, qualitative judgements about the risk of bias, and how well it has been dealt with, should be supplemented by a quantitative assessment.33 34 If a study has used multiple methods, variation in the estimates should be highlighted. The context within which the intervention was implemented should be described as this may affect interpretation and help users assess the generalisability of the findings. Wherever possible, the results
should be compared with those of other evaluations of similar interventions, paying attention to any associations between effect sizes and variations in evaluation methods and interventions, taking a pragmatic approach based on the design and analysis of such studies is more demanding. Priorities for the future are to build up experience of promising but lesser used methods, and to improve the infrastructure that enables research opportunities presented by natural experiments to be seized.

CONCLUSION
There are important areas of public health policy—such as suicide prevention, air pollution control, public smoking bans and alcohol taxation—where natural experimental studies have already contributed a convincing body of evidence. Such approaches are most readily applied where an intervention is implemented on a large scale, the effects are substantial and good population data on exposure and outcomes are available. But they can also be used to detect more subtle effects where there is a suitable source of variation in exposure.

Even so, it would be unwise to assume that a particular policy or intervention could be evaluated as a natural experiment without very detailed consideration of the methodological challenges. Optimism about the use of a natural experimental approach should not be a pretext for discounting the option of conducting a planned experiment, where this would be possible and more robust.

Research effort should focus on addressing important and answerable questions, taking a pragmatic approach based on combinations of research methods and the explicit recognition and careful testing of assumptions. Priorities for the future are to build up experience of promising but lesser used methods, and to improve the infrastructure that enables opportunities presented by natural experiments to be seized, including good routine data from population surveys and administrative sources, good working relationships between researchers and policy makers, and flexible forms of research funding.

REFERENCES


Using natural experiments to evaluate population health interventions: new Medical Research Council guidance

Peter Craig, Cyrus Cooper, David Gunnell, Sally Haw, Kenny Lawson, Sally Macintyre, David Ogilvie, Mark Petticrew, Barney Reeves, Matt Sutton and Simon Thompson

*J Epidemiol Community Health* 2012 66: 1182-1186 originally published online May 10, 2012
doi: 10.1136/jech-2011-200375

Updated information and services can be found at:
http://jech.bmj.com/content/66/12/1182

**Supplementary Material**

Supplementary material can be found at:
http://jech.bmj.com/content/suppl/2012/05/09/jech-2011-200375.DC1

**References**

This article cites 32 articles, 7 of which you can access for free at:
http://jech.bmj.com/content/66/12/1182#BIBL

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**

Articles on similar topics can be found in the following collections

Editor's choice (129)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/