



## Review Article

# Making the most of natural experiments: What can studies of the withdrawal of public health interventions offer?



Peter Craig\*, Marcia Gibson, Mhairi Campbell, Frank Popham, Srinivasa Vittal Katikireddi

MRC/CSO Social & Public Health Sciences Unit, University of Glasgow, Top floor, 200 Renfield Street, Glasgow G2 3QB, UK.

## ARTICLE INFO

## Keywords:

Policy evaluation  
Research methods  
Natural experiments  
Study design

## ABSTRACT

Many interventions that may have large impacts on health and health inequalities, such as social and public health policies and health system reforms, are not amenable to evaluation using randomised controlled trials. The United Kingdom Medical Research Council's guidance on the evaluation of natural experiments draws attention to the need for ingenuity to identify interventions which can be robustly studied as they occur, and without experimental manipulation. Studies of intervention withdrawal may usefully widen the range of interventions that can be evaluated, allowing some interventions and policies, such as those that have developed piecemeal over a long period, to be evaluated for the first time. In particular, sudden removal may allow a more robust assessment of an intervention's long-term impact by minimising 'learning effects'. Interpreting changes that follow withdrawal as evidence of the impact of an intervention assumes that the effect is reversible and this assumption must be carefully justified. Otherwise, withdrawal-based studies suffer similar threats to validity as intervention studies. These threats should be addressed using recognised approaches, including appropriate choice of comparators, detailed understanding of the change processes at work, careful specification of research questions, and the use of falsification tests and other methods for strengthening causal attribution.

Evaluating intervention withdrawal provides opportunities to answer important questions about effectiveness of population health interventions, and to study the social determinants of health. Researchers, policymakers and practitioners should be alert to the opportunities provided by the withdrawal of interventions, but also aware of the pitfalls.

## 1. Introduction

Understanding the effectiveness of interventions is a key step in evidence-informed decision-making (WHO, 2007). Evidence-based medicine has emphasised the central role of randomised controlled trials (RCTs) in producing robust evidence of effectiveness (Guyatt et al., 2008), and the use of trials is strongly advocated in other fields such as poverty relief and social policy-making (Haynes et al., 2013; Tollefson, 2015). At the same time, there are concerns that many interventions are not amenable to experimental manipulation (Barrett and Carter, 2010; Deaton, 2009; Victora et al., 2004), and that an exclusive focus on RCTs will mean that interventions with substantial direct or indirect impacts on health and health inequalities – such as health system reforms, population-wide prevention measures (e.g. sugar and alcohol taxation) and non-health sector changes (e.g. welfare reforms) – will escape robust evaluation (Craig et al., 2017; House of Commons Health Committee, 2009; Katikireddi et al., 2011; Katikireddi et al., 2014).

The United Kingdom (UK) Medical Research Council (MRC) guidance on the evaluation of natural experiments (Craig et al., 2012) argues that we can robustly study interventions that are not under the direct control of researchers, but warns that good natural experiments are scarce, and that ingenuity is needed to identify the available opportunities. Although the importance of planning evaluation alongside the introduction of an intervention is increasingly appreciated by decision-makers and researchers (Cabinet Office, 2003; Trevisan, 2007), in practice this is difficult to achieve (House of Commons Health Committee, 2009). While there has been a renewed emphasis on evaluation recently, there is an historical accumulation of policies and practices supported by precedent or tradition, rather than by evidence of effectiveness. In this paper we argue that there is value in identifying and exploiting opportunities for evaluation of public health policies and interventions which arise from intervention withdrawal as well as from intervention introduction. Studies of intervention withdrawal are widely dispersed across the public health literature and there has been no previous attempt to summarise their contribution. We start by

\* Corresponding author.

E-mail addresses: [peter.craig@glasgow.ac.uk](mailto:peter.craig@glasgow.ac.uk) (P. Craig), [marcia.gibson@glasgow.ac.uk](mailto:marcia.gibson@glasgow.ac.uk) (M. Gibson), [mhairi.campbell@glasgow.ac.uk](mailto:mhairi.campbell@glasgow.ac.uk) (M. Campbell), [frank.popham@glasgow.ac.uk](mailto:frank.popham@glasgow.ac.uk) (F. Popham), [vittal.katikireddi@glasgow.ac.uk](mailto:vittal.katikireddi@glasgow.ac.uk) (S.V. Katikireddi).

<https://doi.org/10.1016/j.ypmed.2017.12.025>

Received 14 June 2017; Received in revised form 23 October 2017; Accepted 23 December 2017

Available online 27 December 2017

0091-7435/ © 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

defining ‘withdrawal’ for the purposes of this paper and describe a number of exemplar studies. In the following two sections we summarise the reasons for studying intervention withdrawal, then consider possible drawbacks of the approach and some solutions. We finish by identifying lessons for the future and discussing some implications of this methodological perspective.

## 2. Defining intervention withdrawal

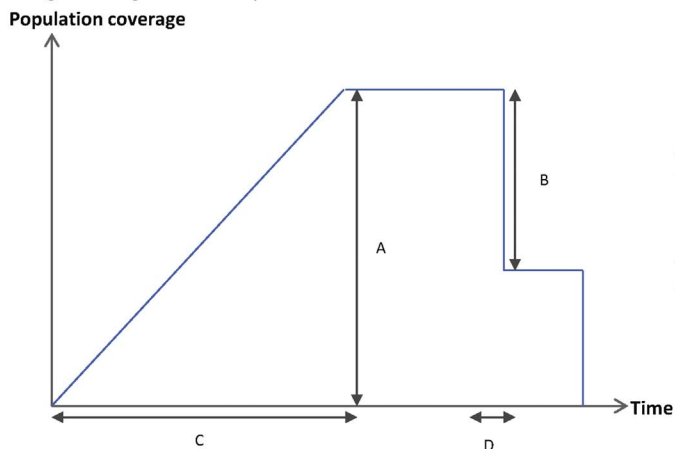
We define interventions broadly to include any kind of law, policy, programme or other action which impacts, positively or negatively, on a social, economic or health outcome. We define withdrawal as the complete or substantial reduction in provision of a longstanding intervention. Withdrawal may result from a deliberate policy change, but may also be an unintended consequence of a decision or event (such as a strike or legal judgement) motivated by other reasons. This definition of withdrawal encompasses a spectrum of processes, which may be abrupt or gradual, partial or complete. Abrupt and complete withdrawal of an intervention is most straightforward to evaluate, but gradual withdrawal or partial replacement also provides useful opportunities for evaluation. The nature of the withdrawal process has implications for the causal effect being evaluated (see Fig. 1). If an intervention that affected the whole population is partially withdrawn, this may limit the generalisability of evaluation findings (Fig. 1a). Similarly, the effectiveness of an intervention may differ over time

(Fig. 1b). For example, it is quite common for learning effects to lead to improved delivery as practitioners become more familiar with an intervention over time. Interventions may also be wholly or partly replaced with alternative interventions, rather than simply withdrawn. Just as pragmatic effectiveness studies must take account of ‘treatment as usual’ (Roland and Torgerson, 1998; Zwarenstein et al., 2008), studies of intervention withdrawal must take account of the precise nature of the comparison condition and extent of replacement.

To help understand the potential contribution of research investigating the withdrawal of interventions, we conducted a structured literature search to identify exemplar studies. We initially identified a number of topics that we were aware had been the subject of withdrawal studies, including hospital closures, alcohol tax reductions, regulatory policies, and welfare reform. We searched Web of Science, PubMed, OVID and Google Scholar using search terms developed with the assistance of an information scientist (CF).

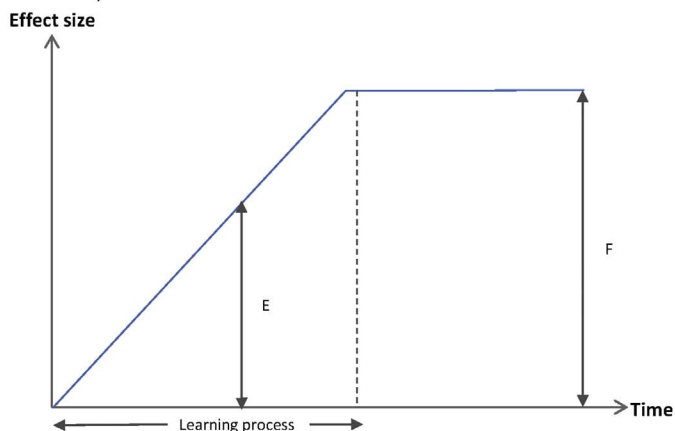
As intervention withdrawal has not been categorised in a standardised way in the literature, we were unable to use study design terminology and instead included the words “abolition”, “closure”, “cut”, “cutback”, “spending cut”, and “tax cut”. Due to the challenges identifying relevant literature and poor indexing of papers, we did not attempt to conduct a comprehensive search to identify all existing literature on withdrawal studies. Instead, we elected to focus on a selection of exemplar studies, chosen to illustrate the diverse range of topics studied and the various analytical approaches employed to

a) Timing and magnitude of implementation



Scenario 1: Evaluating the intervention’s introduction provides a causal estimate that is more generalisable due to larger population coverage than studying partial withdrawal (A vs B). However, evaluating withdrawal provides a causal estimate that may be less prone to confounding than gradual introduction, since there is less chance of a large change in confounders over a shorter time period (D vs C).

b) Consistency of effect size



Scenario 2: Evaluating the intervention’s introduction estimates the causal effect before the intervention is optimised, whereas studying withdrawal allows the optimised causal effect to be estimated (E vs F).

Fig. 1. An illustration of how causal effects may differ between evaluations of intervention introduction and withdrawal.

Scenario a: Evaluating the intervention’s introduction provides a causal estimate that is more generalisable due to larger population coverage than studying partial withdrawal (A vs B). However, evaluating withdrawal provides a causal estimate that may be less prone to confounding than gradual introduction, since there is less chance of a large change in confounders over a shorter time period (D vs C).

Scenario b: Evaluating the intervention’s introduction estimates the causal effect before the intervention is optimised, whereas studying withdrawal allows the optimised causal effect to be estimated (E vs F).

**Table 1**  
Illustrative examples of intervention withdrawal studies.

Withdrawal event	Outcome of interest	Analysis strategy	Reference
<b>Example 1:</b> Repeal of Missouri's 'permit to purchase' handgun law	Firearm related murders	Fixed effects regression of state level pre vs post repeal differences in murder rates	Webster et al. (2014)
<b>Example 2:</b> Abolition of controls on foreign direct investment in Vietnam	Consumption of sugar-sweetened drinks	Difference in differences, comparing Vietnam and the Philippines	Schram et al. (2015)
<b>Example 3:</b> 35% reduction in number of traffic police officers in Oregon	Fatal and non-fatal road traffic accidents	Difference in differences, comparing Oregon with Washington and Idaho; synthetic controls comparing Oregon with a weighted composite of US states	DeAngelo and Hansen (2014)
<b>Example 4:</b> Abolition of General Practice (GP) fundholding in England	Elective hospital admissions	Difference in differences comparing fundholding and non-fundholding GPs	Dusheiko et al. (2006)
<b>Example 5:</b> 33% reduction in alcohol taxes in Finland	Alcohol related sudden deaths	Interrupted time series	Koski et al. (2007)
<b>Example 6:</b> Reduction of pension credits in the UK	Mortality rates in older people	Multiple linear regression	Loopstra et al. (2016)

identify the effects of intervention withdrawal. The exemplar studies are listed in [Table 1](#), and summarised in greater detail below.

**Example 1.** There is wide variation between United States (US) states in laws regulating the purchase and use of firearms, and a close relationship between state-level murder rates and the strictness of their firearm laws (Siegel et al., 2013). However, the observed variation may reflect differences between US states in socio-economic, political and cultural factors that affect both firearm laws and murder rates. Few studies have used changes in firearm laws to identify the effect of specific legal provisions. The withdrawal in 2007 of a key element of Missouri's firearm laws dating from 1921, the permit-to-purchase law which required background checks on purchasers of handguns, provided an opportunity to test the effect on public safety of legal restrictions on purchase of firearms. Webster et al. fitted state-level fixed effects regression models to identify the effect of repeal of the permit-to-purchase law, taking account of changes in other state laws, policies and characteristics that might affect gun crime (Webster et al., 2014). They demonstrated a sharp increase in the murder rate in Missouri following repeal that was specific to firearm related murders, and was not mirrored by similar increases in other US states.

**Example 2.** Accession to international trade treaties by low and middle income countries (LMICs) can result in the rapid removal or lowering of tariff and non-tariff barriers to trade, and an increase in both food imports and inward investment in food processing and manufacturing (Thow and Hawkes, 2009). Schram et al. assessed the potential health implications of removing trade barriers by comparing the growth in consumption of sugar-sweetened drinks in Vietnam before and after its accession to the World Trade Organisation (WTO) in 2007 (Schram et al., 2015). To distinguish the effect of removing trade barriers from secular trends in consumption, they fitted difference in differences models comparing changes in consumption in Vietnam and in the neighbouring Philippines, which had joined the WTO in 1995. Vietnam's rate of growth in consumption increased sharply post-accession and the removal of controls on foreign direct investment. Increased consumption was largely attributable to rising sales of soft drinks by foreign owned rather than Vietnamese companies, and was more rapid in Vietnam over this period than in the Philippines.

**Example 3.** Failure to agree a state budget in 2003 led to the layoff of more than one third of Oregon's traffic police force. De Angelo and Hansen used the layoff to estimate the effectiveness of traffic policing in reducing road traffic injuries and fatalities (DeAngelo and Hansen, 2014). They compared injury and fatality rates in Oregon with those of two neighbouring states before and after the layoff, using difference in differences models. The results indicated that, after allowing for trends in other factors associated with road traffic accidents, such as the weather and the number of young drivers, the reduction in policing led to a 12–14% increase in fatalities. De Angelo and Hansen found similar

effects using a synthetic control approach (Abadie et al., 2010) in which trends in Oregon were compared with those in a weighted composite of other US states.

**Example 4.** No evaluation was planned when key aspects of the commissioning of healthcare were transferred to general practitioners (GPs) by the UK government in 1991–2. Under the new system (widely referred to as 'GP fundholding') GPs could choose to hold a budget for meeting the cost of some elective surgical procedures. The withdrawal of GP fundholding in 1998–9 allowed its impact on elective admissions to be investigated as a marker of rationing in the English National Health Service (NHS) (Dusheiko et al., 2006). Dusheiko et al. used difference in differences models to compare changes in elective admission rates among fundholding and non-fundholding GPs. They found fundholders had lower rates of elective admissions while the scheme was in operation, and increased their rates of admission more than non-fundholders in the two years following abolition. There was no difference in rates of emergency admission, or in the changes in the rates before and after abolition, strengthening the inference that fundholding influenced elective admissions.

**Example 5.** Alcohol excise duties were sharply reduced in Finland in anticipation of Estonia's accession to the European Union (EU) in May 2003. Finland traditionally had high rates of alcohol taxation, but alcohol was markedly cheaper in Estonia, raising fears that the abolition of restrictions would lead to an increase in imports. Using data from medico-legal examinations to identify cause of death, and fitting interrupted time series models to weekly series of deaths from 1990 to 2004, Koski et al. found a 17% increase in sudden, unexpected deaths attributable to alcohol following the tax cut in March 2003 (Koski et al., 2007). There was no change associated with the earlier removal of restrictions on imports, possibly because travellers' allowances had been gradually raised over several years. Subsequent research confirmed the impact on alcohol related hospitalisations, and on deaths from other alcohol related causes (Herttua et al., 2008, 2011).

**Example 6.** Many countries have introduced austerity policies in recent years, scaling back long established policy interventions, including many designed to protect the most vulnerable (Stuckler and Basu, 2013). Loopstra et al. (Loopstra et al., 2016) investigated the impact of substantial reductions to pension credits for low income older people in England using data on death rates and pension credit expenditure at local authority level, and fitting first difference regression models to estimate the effect of changes in pension credit spending on changes in old age mortality. The study found that the loss of credits explained most of the 5% rise in deaths among people aged 85 + between 2012 and 2013. It complements previous comparative research (Lundberg et al., 2008) underlining the effectiveness of social security expenditure in protecting the health of older people.

### 3. Reasons to study intervention withdrawals

The primary reason for studying withdrawal is to widen the range of opportunities for useful natural experimental studies of interventions that are not readily amenable to experimental manipulation, such as large scale policy changes, national legislation or population-wide prevention programmes. Although randomised trials of intervention withdrawal do exist ([Medical Research Council Antiepileptic Drug Withdrawal Study et al., 1991](#)), we know of no examples where they have been used as a proxy to estimate the effect of introducing an intervention. Interventions are most commonly evaluated when they are introduced, because that is when questions of effectiveness tend to be most salient. But frequently such opportunities are missed, as was the case with GP fundholding ([Example 4](#)). Evaluation planning may begin too late, or those responsible for the intervention may be reluctant, for political or other reasons, to expose it to a rigorous test of effectiveness. Also, many services or policies have developed gradually over a long period and may therefore never have been evaluated robustly, if at all. For example, it is often assumed that street lighting improves road safety and reduces crime. Pressure on local government budgets and concerns for the environmental impact of street lighting have led to varying degrees of withdrawal in the UK, but a recent study using fixed effect panel models found little impact on either night time traffic accidents or crime ([Steinbach et al., 2015](#)).

Another possible advantage of studying withdrawal as opposed to the introduction of an intervention is that the process may be more abrupt and at times, unexpected. For example, legislation to restrict the availability of firearms may impact on health outcomes more gradually than the sudden withdrawal of restrictions and therefore changes such as the repeal of Missouri's permit controls ([Example 1](#)) have been highlighted as providing opportunities for more robust evaluations ([Santaella-Tenorio et al., 2016](#)). Studies of the introduction of an intervention, particularly when that process is slow, may be subject to behavioural changes in advance of the intervention (anticipation effects), or require practitioners to learn new techniques that take time to master. For example, it takes time to recruit and train traffic police, so that the effect of increasing numbers of police may be more prone to confounding by trends in other factors influencing accident rates, than the sharp reduction studied by De Angelo and Hansen ([Example 3](#)) ([DeAngelo and Hansen, 2014](#)). Studies of intervention withdrawal may also usefully supplement conventional implementation studies. Hospitalisation for acute myocardial infarction (AMI) fell when a smoke-free law was implemented in Helena, Montana, in 2002, then rose when the law was suspended six months later ([Sargent et al., 2004](#)). The second change makes it less plausible that the first reflected secular trends rather than an effect of the smoke-free law.

So far we have focused on the use of withdrawals as a device for identifying the effectiveness of the intervention being withdrawn, but this is not their only purpose. Withdrawals also offer scope for studying the social determinants of health. For example, unemployment is strongly associated with poor health, but the association may reflect health-related selection into unemployment or common causes of poor health and unemployment, rather than a causal effect of unemployment on health. Factory closures, where the whole work force is made redundant, provide a clearer test of the processes involved, and a number of studies have used such events to distinguish the health impacts of unemployment from the effects of selection and confounding by other dimensions of socio-economic position ([Eliason and Storrie, 2009](#); [Keefe et al., 2002](#); [Martikainen et al., 2007](#)). Finally, to the extent that they reflect the operation of wider policies, such as fiscal austerity or trade liberalization, studies of intervention withdrawal may contribute to understanding the overall impacts of broad policy approaches.

### 4. Methodological challenges

One possible drawback with using intervention withdrawal to study

effectiveness is that the intervention may no longer be of interest to decision-makers. There is a substantial body of research on the effects of psychiatric hospital closures which could be regarded as largely of historical interest, because of a permanent shift towards community care ([Kunitoh, 2013](#)). On the other hand, this literature could be seen as highly relevant to current debates about the appropriate balance between hospital and community-based mental health care ([Winkler et al., 2016](#)). The examples we have given involve policies and programmes, such as gun control and alcohol taxation, which are still of immediate and widespread public health interest. Variation between jurisdictions means that interventions withdrawn in one area may still be in place elsewhere, as in the case of permit to purchase laws. Evidence from withdrawal studies may therefore inform decision-making elsewhere. Withdrawal may sometimes be forced on policy-makers by external events, as in [Examples 3 and 5](#), rather than reflecting an underlying policy change. Although the case for relevance needs to be considered carefully, there is no general reason to assume that interest in an intervention ceases following withdrawal.

A more serious potential drawback is that changes associated with an intervention's withdrawal may not provide useful information about its effectiveness. For the withdrawal to be informative about the effectiveness of an intervention, the effects must be reversible. Effects may be only partly reversible where interventions have been in place for long periods, during which cultural or systemic changes may have occurred that lessen the effect of withdrawing the intervention. The replacement in October 2015 of China's one child policy with a two child policy was expected to have little observable impact, because economic and social changes over the 35 year duration of the policy have led to a longstanding reduction in fertility ([The Economist, 2015](#)). In practice, there was a marked increase in the birth rate in the year following the change, and specifically in the numbers of second births, though it is not expected to reverse the rise in dependency ratio ([The Economist, 2017](#)). The extent to which the effect of policies is culturally mediated varies widely, and such changes may not matter in some cases. In Dusheiko et al.'s study of GP fundholding ([Example 4](#)), the authors carefully address the issue of reversibility and conclude that although permanent changes in GPs' behaviour as a result of fundholding cannot be ruled out, such changes would mean that inferences based on the observed impact of withdrawal would be conservative ([Dusheiko et al., 2006](#)).

Estimates of the effect of an intervention inferred from changes associated with its withdrawal may be biased by anticipation effects if the withdrawal is widely known of in advance. Withdrawal-based effect estimates may also be confounded by transitional effects, introduction or withdrawal of other interventions at the same time, or by compensating behaviors. The paradoxical finding that mortality remains stable or falls during doctors' strikes reflects the fact that emergency care is maintained, while elective procedures that carry a short term risk are postponed, rather than evidence that healthcare is pointless or harmful ([Cunningham et al., 2008](#)). In circumstances of fiscal austerity, a wholesale reduction in public spending may make it difficult to identify the effect of a specific withdrawal. A key strength of DeAngelo and Hansen's study of traffic policing ([Example 3](#)) is that the layoffs came about as a result of failure to agree a state budget, rather than from a broader economic recession.

The above drawbacks should be avoidable, at least in some cases, by using standard approaches to strengthening causal inference in natural experimental studies. All such studies make causal inferences by comparing outcomes in a population exposed to an intervention and an unexposed comparator population, with the choice of method dependent on the nature of the event or process that determines the difference or change in exposure. Withdrawal events lend themselves to approaches that exploit changes over time, such as fixed effects panel models ([Example 1](#)), difference in differences ([Examples 2–4](#)), interrupted time series ([Example 5](#)), or synthetic control methods ([Example 3](#)).

**Table 2**

Is studying intervention withdrawal appropriate? Eight key questions.

1	Is there a clearly defined intervention that has been removed or substantially reduced?
2	Was the intervention in place for long enough to abolish learning effects, etc.?
3	Has the intervention brought about cultural or other changes that are likely to influence health outcomes and to persist following withdrawal?
4	Has the intervention been replaced? If so, what with?
5	Are there health outcomes of interest that are likely to change immediately or within a known lag period?
6	Are there likely to be anticipatory effects prior to withdrawal of the intervention?
7	Were other policies withdrawn or introduced that might confound the effects of the withdrawal of interest?
8	Are there likely to be short term effects associated with disruption or other features of the withdrawal process?

As with all natural experimental studies, causal inference is most straightforward when effects are large and rapid, and additional design elements are needed to strengthen inference and rule out alternative explanations when effects are more subtle. These additional elements include a detailed understanding of the change processes at work (including a clear rationale for the choice of outcome and length of time expected before an effect is seen), consideration of contextual factors that may moderate the impact of withdrawal, careful choice of comparator(s) and the use of placebo or falsification tests (Craig et al., 2017). Use of a combination of methods that rely on different assumptions can further strengthen causal inference. In their study of the impact on lone mothers of the withdrawal of Assistance for Families with Dependent Children (AFDC), Basu et al. used both difference in differences and synthetic control approaches (see also Example 3) (Basu et al., 2016).

Anticipation and transitional effects are important considerations in studies that exploit change in exposure over time to identify the effect of an intervention but are not specific to studying withdrawal, and can be addressed in the same way as they are in traditional intervention studies, for example by testing for effects associated with a range of implementation dates, or over a range of observation periods. Table 2 summarises considerations that need to be taken into account in identifying instances of intervention withdrawal that may be useful to study.

## 5. Conclusions

Studies of intervention withdrawal usefully widen the range of opportunities to evaluate interventions, and in particular allow some interventions and policies that have developed piecemeal without any systematic assessment of their effectiveness to be evaluated for the first time. They can also add to the evidence base for interventions, such as water fluoridation to prevent dental caries, which have been evaluated in other ways, but where the evidence remains inconclusive (Iheozor-Ejiofor et al., 2015; McLaren and Singhal, 2016). Interpreting changes that follow withdrawal as evidence of the impact of an intervention assumes that the effect is reversible, and this assumption must be carefully justified. Otherwise, withdrawal-based studies suffer similar threats to validity as conventional intervention studies, and these should be addressed using similar approaches, including appropriate choice of interventions, identification of an appropriate comparator, detailed understanding of the change processes at work, careful specification of questions, and the use of falsification tests and other methods for strengthening causal attribution (Craig et al., 2017).

In this paper, we primarily focus on using intervention withdrawal to better understand the effectiveness of the intervention. In addition, intervention withdrawal may provide useful opportunities for studying social determinants of health, if the withdrawal event breaks the link between exposure and confounders, or the process of withdrawal may be of interest in its own right (for example, because the disruption

involved has adverse impacts (Fulop et al., 2002; Greenhalgh et al., 2011)). The studies we have discussed, and other examples in the public health and related literatures, clearly demonstrate the value of studying withdrawals, particularly of interventions that are likely to have been under-researched at the time of their implementation. Wider appreciation of the value of studying intervention withdrawal should lead to better use of the available opportunities.

## Abbreviations

AFDC	Assistance for Families with Dependent Children
AMI	acute myocardial infarction
EU	European Union
GP	General Practice
MRC	Medical Research Council
NHS	National Health Service
RCT	randomised controlled trial
UK	United Kingdom
US	United States
WTO	World Trade Organisation

## Acknowledgments

We thank Candida Fenton for assistance with literature searching, John Gilchrist for assistance with the production of the Figure, and Natalie Owens for help preparing the manuscript for submission.

## References

- Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *J. Am. Stat. Assoc.* 105, 493–505.
- Barrett, C.B., Carter, M.R., 2010. The power and pitfalls of experiments in development economics: some non-random reflections. *Appl. Econ. Perspect. Policy* 32, 515–548.
- Basu, S., Rehkopf, D.H., Siddiqi, A., Glymour, M.M., Kawachi, I., 2016. Health behaviors, mental health, and health care utilization among single mothers after welfare reforms in the 1990s. *Am. J. Epidemiol.* 183, 531–538.
- Cabinet Office, 2003. *The Magenta Book: Guidance Notes for Policy Evaluation and Analysis*. Cabinet Office, London.
- Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., Ogilvie, D., Petticrew, M., Reeves, B., et al., 2012. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *JECH* 66, 1182–1186.
- Craig, P., Katikireddi, S.V., Leyland, A., Popham, F., 2017. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. *Annu. Rev. Public Health* 38, 39–56.
- Cunningham, S.A., Mitchell, K., Venkat Narayan, K.M., Yusuf, S., 2008. Doctors' strikes and mortality: a review. *Soc. Sci. Med.* 67, 1784–1788.
- DeAngelo, G., Hansen, B., 2014. Life and death in the fast lane: police enforcement and traffic fatalities. *Am. Econ. J. Econ. Pol.* 6, 231–257.
- Deaton, A.S., 2009. *Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development*. National Bureau of Economic Research.
- Dusheiko, M., Gravelle, H., Jacobs, R., Smith, P., 2006. The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment. *J. Health Econ.* 25, 449–478.
- Eliason, M., Storrie, D., 2009. Job loss is bad for your health – Swedish evidence on cause-specific hospitalization following involuntary job loss. *Soc. Sci. Med.* 68, 1396–1406.
- Fulop, N., Protopsaltis, G., Hutchings, A., King, A., Allen, P., Normand, C., Walters, R., 2002. Process and impact of mergers of NHS trusts: multicentre case study and management cost analysis. *BMJ* 325, 246.
- Greenhalgh, T., Russell, J., Ashcroft, R.E., Parsons, W., 2011. Why national eHealth programs need dead philosophers: Wittgensteinian reflections on policymakers' reluctance to learn from history. *Milbank Q.* 89, 533–563.
- Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H.J., 2008. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336, 924–926.
- Haynes, L., Service, O., Goldacre, B., Torgerson, D., 2013. *Test, Learn, Adapt: Developing Public Policy With Randomised Controlled Trials*. Cabinet Office: Behavioural Insights Team, London.
- Herttua, K., Mäkelä, P., Martikainen, P., 2008. Changes in alcohol-related mortality and its socioeconomic differences after a large reduction in alcohol prices: a natural experiment based on register data. *Am. J. Epidemiol.* 168, 1110–1118.
- Herttua, K., Mäkelä, P., Martikainen, P., 2011. The effects of a large reduction in alcohol prices on hospitalizations related to alcohol: a population-based natural experiment. *Addiction* 106, 759–767.
- House of Commons Health Committee, 2009. *Health Inequalities: Report of Session*

- 2008–09. vol. 1 House of Commons, London.
- Iheozor-Ejiofor, Z., Worthington, H.V., Walsh, T., O'Malley, L., Clarkson, J.E., Macey, R., Alam, R., Tugwell, P., Welch, V., et al., 2015. Water fluoridation for the prevention of dental caries. In: *The Cochrane Library*: Art. No.: CD010856.
- Katikireddi, S.V., Higgins, M., Bond, L., Bonell, C., Macintyre, S., 2011. How evidence based is English public health policy? *BMJ* 343, d7310.
- Katikireddi, S.V., McKee, M., Craig, P., Stuckler, D., 2014. The NHS reforms in England: four challenges to evaluating success and failure. *J. R. Soc. Med.* 107, 387–392.
- Keefe, V., Reid, P., Ormsby, C., Robson, B., Purdie, G., Baxter, J., Incorporated, N.K.I., 2002. Serious health events following involuntary job loss in New Zealand meat processing workers. *Int. J. Epidemiol.* 31, 1155–1161.
- Koski, A., Sirén, R., Vuori, E., Poikolainen, K., 2007. Alcohol tax cuts and increase in alcohol-positive sudden deaths—a time-series intervention analysis. *Addiction* 102, 362–368.
- Kunitoh, N., 2013. From hospital to the community: the influence of deinstitutionalization on discharged long-stay psychiatric patients. *Psychiatry Clin. Neurosci.* 67, 384–396.
- Loopstra, R., McKee, M., Katikireddi, S.V., Taylor-Robinson, D., Barr, B., Stuckler, D., 2016. Austerity and old-age mortality in England: a longitudinal cross-local area analysis, 2007–2013. *J. R. Soc. Med.* 109, 109.
- Lundberg, O., Yngwe, M.Å., Stjärne, M.K., Elstad, J.I., Ferrarini, T., Kangas, O., Norström, T., Palme, J., Fritzell, J., 2008. The role of welfare state principles and generosity in social policy programmes for public health: an international comparative study. *Lancet* 372, 1633–1640.
- Martikainen, P., Mäki, N., Jääntti, M., 2007. The effects of unemployment on mortality following workplace downsizing and workplace closure: a register-based follow-up study of Finnish men and women during economic boom and recession. *Am. J. Epidemiol.* 165.
- McLaren, L., Singhal, S., 2016. Does cessation of community water fluoridation lead to an increase in tooth decay? A systematic review of published studies. *J. Epidemiol. Community Health* 70, 934–940.
- Medical Research Council Antiepileptic Drug Withdrawal Study, G, Bessant, P., Chadwick, D., Eaton, B., Taylor, J., Holland, A., Joannou, J., Johnson, A.L., Oldfield, L., et al., 1991. Randomised study of antiepileptic drug withdrawal in patients in remission. *Lancet* 337, 1175–1180.
- Roland, M., Torgerson, D.J., 1998. What are pragmatic trials? *Br. Med. J.* 316, 285–286.
- Santaella-Tenorio, J., Cerdá, M., Villaveces, A., Galea, S., 2016. What do we know about the association between firearm legislation and firearm-related injuries? *Epidemiol. Rev.* 38, 140–157.
- Sargent, R.P., Shepard, R.M., Glantz, S.A., 2004. Reduced incidence of admissions for myocardial infarction associated with public smoking ban: before and after study. *BMJ* 328, 977–980.
- Schram, A., Labonte, R., Baker, P., Friel, S., Reeves, A., Stuckler, D., 2015. The role of trade and investment liberalization in the sugar-sweetened carbonated beverages market: a natural experiment contrasting Vietnam and the Philippines. *Glob. Health* 11, 1–13.
- Siegel, M., Ross, C.S., King, C., 2013. The relationship between gun ownership and firearm homicide rates in the United States, 1981–2010. *Am. J. Public Health* 103, 2098–2105.
- Steinbach, R., Perkins, C., Tompson, L., Johnson, S., Armstrong, B., Green, J., Grundy, C., Wilkinson, P., Edwards, P., 2015. The effect of reduced street lighting on road casualties and crime in England and Wales: controlled interrupted time series analysis. *J. Epidemiol. Community Health* 69, 1118–1124.
- Stuckler, D., Basu, S., 2013. *The Body Economic: Why Austerity Kills*. Basic Books.
- The Economist, 2015. China Has Relaxed Its One-Child Policy. Yet Parents Are Not Rushing to Have a Second.
- The Economist, 2017. China's Spike in Births Will Not Alleviate Its Ageing Problem.
- Thow, A.M., Hawkes, C., 2009. The implications of trade liberalization for diet and health: a case study from Central America. *Glob. Health* 5, 1–15.
- Tollefson, J., 2015. Revolt of the randomistas. *Nature* 254, 153.
- Trevisan, M.S., 2007. Evaluability assessment from 1986 to 2006. *Am. J. Eval.* 28, 290–303.
- Victora, C.G., Habicht, J.-P., Bryce, J., 2004. Evidence-based public health: moving beyond randomized trials. *Am. J. Public Health* 94, 400–405.
- Webster, D., Crifasi, C.K., Vernick, J.S., 2014. Effects of the repeal of Missouri's handgun purchaser licensing law on homicides. *J. Urban Health* 91, 293–302.
- WHO, 2007. *Sound Choices: Enhancing Capacity for Evidence-Informed Health Policy*. World Health Organization.
- Winkler, P., Barrett, B., McCrone, P., Csémy, L., Janousková, M., Höschl, C., 2016. Deinstitutionalised patients, homelessness and imprisonment: systematic review. *Br. J. Psychiatry* 208, 421–428.
- Zwarenstein, M., Treweek, S., Gagnier, J.J., Altman, D.G., Tunis, S., Haynes, B., Oxman, A.D., Moher, D., 2008. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 337.