Same words, different meanings: How epidemiological terminology struggles with population health intervention research

Mark Petticrew,[1] Sarah Viehbeck,[2] Steven Cummins,[1] and Thierry Lang [3]

1. Department of Social and Environmental Health Research, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London WC1H 9SH.

2. Sarah M. Viehbeck, School of Public Health and Health Systems, University of Waterloo, Waterloo, Canada.

3. Département d'Epidémiologie et Santé Publique, UMR1027 Inserm-Université Paul Sabatier, Toulouse, France.

This is the English version of the following paper, which was originally published in French:

Citation: Petticrew M, Viehbeck S, Cummins S, Lang T. À mêmes mots, sens différents – les difficultés de la terminologie épidémiologique avec la recherche en interventions en santé des populations. *Revue d'Épidémiologie et de Santé Publique*, online first, 31st March 2016.

URL: http://www.sciencedirect.com/science/article/pii/S0398762016001462

Abstract

Public health research differs from clinical epidemiological research in that its focus is primarily on the population level social and structural determinants of individual health and the interventions that might ameliorate them, rather than having a primary focus on individual-level risks. It is typically concerned with the proximal and distal causes of health problems, and their location within complex systems, more than with single exposures. Thus, epidemiological terms and concepts may have very different implications when used in the context of population health. This paper considers some key differences in relation to terms like ‘population’, ‘baseline’, ‘control group’ ‘outcome’ and ‘adverse effects’. Even the concept of an ‘intervention’ often needs careful handling. The paper concludes that there is a need for an expanded, and more realistic use of these terms in the population health intervention research context.

Introduction

Population health intervention research (PHIR) is emerging as a field aimed at generating evidence about the population-level impacts of policies, programs, and actions and policies from within or outside of the public health sector. The interventions in question include, for example, income, education, and employment opportunities and their impacts on health.(1) Though the evaluation of such interventions has a long history, there has been an upsurge of interest in the past ten years driven by a greater understanding of the need to intervene on the "upstream" determinants of health.(2)

Much of the debate about the sort of evidence that is needed for this enterprise - in particular discussions about "better evidence" and the need to identify "what works" – frequently draws on clinical epidemiological language and on concepts and approaches developed in evidence-based medicine. However there are obvious differences between clinical and population health interventions, and these differences pose problems for the translation of clinical epidemiological language and concepts between these fields. Firstly, and most obviously, population health can be distinguished from other areas of epidemiology and public health in that its focus is primarily on the population level social and structural determinants of individual health and on the interventions that might ameliorate them, rather than having a primary focus on individual-level risks. Population health is concerned with the proximal and distal causes of health problems, and their location within complex systems, more than with single exposures.

For these and other reasons methodological terms and concepts used in many areas of clinical epidemiology may have different implications when used, without adaptation, in the context of population health. One possibility is that we may appraise evidence created within a population health paradigm quite differently – and more harshly - than evidence created to inform clinical decision-making. Indeed it is sometimes said informally that the public health research are overly critical of their own and other's research, placing emphasis on different factors during funding peer review than, say, basic scientists. For example, an analysis of over 300 Canadian Institutes of Health Research grants has shown that, compared to biomedical/clinical panels, health services and public health panels score grant applications lower.(3) Biomedical committees place greater importance on the track record of the author, the logical derivation of the research ideas, appropriate laboratory techniques, and provisions for graduate student support. In contrast, health committees place greater importance on research design, methods and statistics, literature reviews, and budgets. In addition the analysis found less consensus among the ratings of health committees, than among medical committees.(3) An evaluation of the CIHR Open Operating Grant program in 2012 reached similar conclusions.(4)

One response might be to move the goalposts and assess the methodological quality of PHIR more generously. A more credible position might be to explore how these methodological concepts differ between disciplines, why this matters, and then to calibrate judgments about strength of evidence accordingly. Examining the methodological language underpinning such judgments is an important starting point; there is power enshrined in language, and methodological terms like "control" and "baseline" are not just descriptions of groups or points in time, but also carry implications about methodological rigour.

The purpose of this paper is therefore to interrogate some of the more salient methodological terms for which there appears to be a difference in implications between clinical epidemiology and PHIR, and to use this frame to argue for an expanded, and more realistic use of these terms in the population health intervention context. As a broad organizing principle, we have adopted a framework widely used in evidence-based medicine to help think through the framing of research questions: the PICO framework

Population, Intervention (or exposure), Control/Comparison, and Outcomes (PICO).(5) We chose this as a starting point as the PICO items may often be particularly difficult to define and apply when we move beyond clinical/healthcare interventions. The discussion however extends beyond PICO in order to bring in other key methodological terms and concepts that may be encountered in the development or evaluation of population health interventions. To select the terms for discussion, all the items in the Cochrane Handbook's "Glossary of Terms" were reviewed independently by the authors for possible inclusion.(6) Terms were selected on the basis of likely differences in interpretation for population health intervention research, compared with clinical epidemiology contexts. A final set of terms for discussion was selected based on convergence across at least two of three reviewers. We hope that this discussion may be of relevance to several audiences, in particular those who need to critically appraise population health research before it is funded, or implemented, as well as researchers and systematic reviewers more attuned to appraising clinical research than population health research. It may also be of relevance to journal editors who often have to assess the comparative merits of public health research alongside clinical studies.

Populations and baselines

What is a "population"? In clinical trials, the "population" refers to the demographic, clinical and other variables collected for each participant at the beginning of the study, before the intervention is administered. The Cochrane Handbook defines this broadly, to include relevant demographic factors as well as geography and diseases.(6) A clearly defined population is important not just for purposes for description, but because this is often used as a marker for study "quality". A population is the base from which a study sample is drawn – it needs to be identified and quantified at the start and end of a quantitative longitudinal study or trial, along with the ability to account for any attrition during the research. An ill-defined or non-representative population may also have implications for the generalizability of the findings. As a result, defining a population and its representativeness often has implications for whether the study scores "high" or "low" in terms of study quality on critical appraisal tools.(7) However in population health, the nature and definition of populations may be contested. Crucially, populations are not just aggregates of individual characteristics, but as Krieger (2012) describes, social relations may also shape population health distributions, as shown in the analysis of the spread of obesity through social networks in the Framingham Offspring study.(8) (9)

Describing aggregates of sample characteristics at baseline is nonetheless important because it allows the "before" population to be compared to the "after" when evaluating change in response to an intervention. For complex public health interventions, we are often less interested in the straightforward change from "before" to "after" a given intervention exposure (e.g., introduction of a policy or program at a population-level), and more interested in the extent to which the intervention interacts with different populations within and across different contexts. Is it more effective in some populations than others? And are there features of the context within which the intervention is introduced which mediate or moderate its effectiveness? If so, then it becomes important to collect adequate information not just about baseline characteristics of populations in isolation, but also about the initial state of the contexts (structures, environments, and systems) into which they are introduced.

However identifying when the population was exposed to the intervention may often be easier said than done. For example, individuals in populations often live in one context but work in another. As a result, they are operating within fluid boundaries. There may have been a long prior period during which the population was exposed to different, or weaker "doses" of the intervention, with no clear "start" or "end" to the intervention phase. In the case of introducing new policies, for example, the intervention may not

start simply when the policy is launched – instead, it is possible that the population's behavior (and knowledge, and attitudes) may have been in a process of change long before that. The intervention may even be said to begin during the policy agenda-setting phase, and broader awareness about the policy is being raised, or when the policy is first announced publically, or when the system within which the policy is delivered begins to adjust in anticipation of the formal start date.

As a result, thinking of interventions as "events in systems" as Hawe, Shiell et al. (2008) describe is a more accurate description of the long, slow social change process which social interventions often involve.(10) One example is the introduction of "Smokefree" legislation in many countries, which places restrictions on smoking in public places.(11) Should the "baseline" for evaluating such legislation be considered to be the day before the policy was introduced? (or the same day the year before to control for seasonal effects). Or does the baseline begin before this? Many interventions (like policies) have the potential to affect behavior in advance of implementation, as individuals or communities prepare for change. In the case of the Smokefree legislation in Scotland for example smoking quit rates increased in the period before the introduction of the legislation.(12) Policies themselves may also diffuse over time, as policies in one jurisdiction influence policies in another. For example, Nykiforuk et al. (2008) have described how smoke-free spaces bylaws diffused over a 30-year period across the provinces of Alberta and Ontario, Canada.(13)

The baseline in evaluations of population-level interventions can therefore extend back in time. The success and acceptability of the Smokefree legislation may have depended on the many decades during which the public were sensitized to the need for legislation. During this period the risks of smoking became widely known and the public acceptability of smoking itself began to wane, something the tobacco industry itself realized.(14) This undoubtedly affected the population readiness for the intervention. Dating the start of the intervention in Scotland for example to 6am on March 26, 2006, the date legislation was introduced, may therefore represent the end of a long process, rather than the start of implementation. As the public acceptability of legislative interventions continues to evolve after they have been introduced, it may not even be the end of the intervention.(15)

One possible approach to this problem is to consider the use of multiple baselines, so that baseline measurement of both the population and its context is not a "one-off" event. Given the shifting nature of baselines and populations, the context too may extend back in time and understanding this may require collecting historical data. This, and consideration of the "policy diffusion" described above may help the evaluator understand the extent to which a population and context is or was "intervention ready" and to gather a more realistic sense of pre-intervention conditions. Thus, context is not something to be controlled for, but integral to understanding how and why an intervention works. This is particularly true of policies (*"Many policy choices cannot be broken down into discrete interventions evaluated in a context-free manner."* (Kouri, 2009). (16)

Interventions, and intervention and control groups

The above discussion - and the difficulty in determining when an intervention begins -suggests that the word "intervention" can itself be misleading. There is a clear tension between the fluid and dynamic nature of environments and systems, and the implication carried by the term "baseline" of a steady, measurable prior state. When intervening at the population level we are not working with discrete, de-contextualized events with clear beginnings, middles, and ends. Rather, evaluations of complex population health interventions must attempt to capture relevant aspects of the states of systems over time, and try to understand what the impact of the introduction of the intervention is on that system.(17)

The Cochrane Handbook for Systematic Reviews of Interventions describes an intervention as "*the process of intervening on people, groups, entities or objects in an experimental study*".(18) In controlled trials, the word is sometimes used to describe the regimens in all comparison groups, including placebo and no-treatment arms. In this case "Intervene" and "intervention" carry the implication that some intentional, planned action takes place. This may or may not be the case in population health. Evaluations of natural experiments may involve evaluations of policies in different jurisdictions which are implemented outside of the control of a researcher and are not often deliberately implemented in a way that facilitates controlled evaluation. (19) Instead the evaluation uses natural variations in exposure as a way of assessing impact. This approach is closer to observational epidemiology than clinical trial methodology.

Population health interventions could also legitimately include multiple rather than single interventions. While this may be problematic in terms of confounding, it may also provide the opportunity to study interactions between intervention elements. The key point however is that interventions are often not developed by researchers, and evaluation may often be post-hoc and pragmatic.

If "intervention" and "population" are flexible concepts, then by extension defining intervention and control groups must be correspondingly so. The Cochrane definition of "Intervention group" refers to a group of participants in a study receiving a particular health care intervention. "Parallel group trials include at least two intervention groups. The control arm in a trial "acts as a comparator for one or more experimental interventions". (18) This is an ideal which is not always achieved, or even appropriate in population health. Rather than a binary definition of exposure (some receiving the intervention, some not) there are more often populations which are exposed to varying amounts and aspects of the intervention at different times. This may occur because the intervention is modified by implementers according to their judgment of different populations' needs, and may be modified by populations themselves. The intervention itself may also evolve over time and may legitimately vary across settings. Hawe et al. (2004) describe how, for such interventions, the issue is not the standardization of the intervention, but the standardization of the processes and functions - that is, the mechanisms by which the change takes place – the intervention itself however may vary. (20) In Hawe's example, if the intervention is, say, "workshops for general practitioners", the form of these can adapt to the local context,

The purpose of "controls" is to provide a clear counterfactual, with the ideal that the control group should be similar in all important characteristics (other than receipt of the intervention) which are likely to affect the relevant outcomes. Given the variations in exposure described above, this requirement becomes progressively more difficult to meet as one moves from individual-level to population-level interventions within particular jurisdictions or which cross jurisdictions. In particular, the idea of a "control" population or area becomes more difficult to operationalize when one considers evaluations of the effects of area-based interventions, for example, evaluations of changes in cities, countries, or other environments; the concept of "control cities" stretches the concept of "control" - it is important to identify suitable "comparison" areas but close matching on prognostic variables may be difficult and the comparison areas are unlikely to be highly similar. Furthermore, for many policy interventions, there will be no untouched, "no-intervention" control group – particularly given the fluidity of populations described above.

One implication is that in the absence of a true control group, there could be other forms of comparison groups or areas, with different experiences of and exposure to the intervention or multiple time series approaches for a within-area "control" group (pre- vs. post-exposure). The evaluation approach to this challenge may therefore quite legitimately be closer to observational etiological epidemiology than to traditional evaluation research, involving descriptive observational studies and mathematical models or simulations. The case for non-experimental research to evaluate social and public health interventions has often been made (21) (19) but the basic principle is that the researcher seeks to draw scientifically-

based inferences about the relationship between causes and outcomes. Non-experimental studies are appropriate for doing this, with the choice of design depending on feasibility, effect size, cost, whether the intervention has already happened or not, and the possibility for the researcher to influence the intervention implementation.

Primary outcomes, intermediate outcomes and adverse effects

Epidemiological approaches to evaluating interventions require that the primary outcome should be defined *a priori*. The Cochrane definition refers to “the outcome of greatest importance” but also recommends that Cochrane Reviews should seek to include all outcomes that are likely to be important to users of the review, bearing in mind the risks (e.g., false positives). This is a flexible approach which acknowledges that there is no perfect solution. Funders and peer reviewers however are often less flexible and studies which do not nail their colors to the mast of a single, primary outcome may be not be well received, or even funded. As population health interventions often occur outside the health sector it may make little sense to users to make a health outcome the “primary” one. For population health research, particularly for interventions outside the health sector, categorizing outcomes as "primary" or "secondary" is often difficult, given that different stakeholders or sectors will have different perspectives on the value of the different outcomes.

Intermediate and surrogate outcomes may therefore be particularly important in population health intervention research. These are defined as outcome measures that are not of direct practical importance but are believed to reflect outcomes that are important nonetheless. Surrogate endpoints are often physiological or biochemical markers that can be relatively quickly and easily measured, and that are taken as being predictive of important clinical outcomes. They are often used when observation of clinical outcomes requires long follow-up, though can be misleading(22) However given the lag times between some interventions and their health impacts in population health research, careful use of intermediate outcomes may be helpful to users and help elucidate the causal pathways - for example, by providing evidence of effects at various points on the pathway between intervention and outcome. For this reason there is increasing interest in modeling approaches such as Directed Acyclic Graphs (DAGS), Agent Based Modeling, and Structural Equation Modeling (SEM).(23, 24)

The possibility of unintended adverse outcomes is a particular concern in population health interventions. Key among these adverse outcomes is the possibility that an intervention may increase health inequities in the population. Lorenc and Oliver contrast public health with clinical medicine, where there is a substantial literature on adverse events and patient safety, including how to monitor them through the research process; in public health research however unintended harms may go uncaptured or be overlooked, particularly in cases of policies or programs where researchers are not controlling the intervention. In the case of inequities, they may be difficult to identify; for example improved health outcomes at a population level may mask increasing inequities between subgroups in the same population – even where every individual experiences in that population some degree of improvement.(25) There is some evidence that downstream, individually-focussed interventions, as opposed to upstream population-level interventions may be more likely to increase inequities.(26)

Policies, programmes and projects

Aside from simply examining 'interventions', as described above, population health is also frequently concerned with the effects of policies, programmes and projects. However the distinction between these three terms is often vague, and the terms are often used interchangeably. In its Glossary of Health Impact Assessment, or HIA, (27), the World Health Organization makes a distinction between them:

"Policy: A policy can be defined as an agreement or consensus on a range of issues, goals and objectives which need to be addressed.. For example, "Saving Lives: Our Healthier Nation" can be seen as a national health policy aimed at improving the health of the population of England, reducing health inequalities and setting objectives and targets which can be used to monitor progress towards the policy's overall goal or aims.
Programme: The term programme usually refers to a group of activities which are designed to be implemented in order to reach policy objectives.. For example, many Single Regeneration Budget programmes and New Deal for Communities initiatives have a range of themes within their programmes – often including health, community safety (crime), education, employment and housing – and within these themes are a number of specific projects which, together, make up the overall programme.
Project: A project is usually a discrete piece of work addressing a single population group or health determinant, usually with a pre-set time limit. For example, "Private Rented Dwellings" was a three year project in Southport, Merseyside which provided money to private landlords in order to bring their rented properties up to housing fitness standards….A project can therefore be seen as "a component of a programme".

Some authors further distinguish these terms according to their application at strategic and tactical levels: *"Policy represents the way in which government or an organisation seeks to achieve the objectives it has set. HIA at this level can be strategic, enabling health concerns to be incorporated early on and a "global" view to be taken. […] HIA at the programme or project level allows health impacts to be assessed that are specific to a particular locality or community. It is more tactical, with aims relating to proposal modification and implementation*(28).

However authors from the fields of Health Impact Assessment, policy analysis and health promotion (e.g. Potvin and Mc Queen), have all emphasised that the distinction between these levels remains vague: *"Program, intervention, project, initiative, are often used interchangeably" (2009)."* (29)

All the terms listed above are frequently deployed to support the wider endeavour of making public health more evidence-based. However if even basic concepts are either ill-defined or are subject to widely different interpretations, this in turn presents greater problems for population health intervention research, summarised by one researcher as "Can you want evidence based policymaking if you don't really know what it is?" (30).

.

Discussion

We have used the PICO framework to structure a discussion around some of the main epistemological and methodological differences between more "traditional" epidemiology and population health research. The list of terms is obviously not exhaustive (other fruitful areas for discussion could include "context", "confounding" and "blinding").Those we have chosen to discuss are not meant to be taken as the most "important", but rather those which we believe appear to be fundamental concepts (baseline, control) yet

which may significantly differ in their application between clinical epidemiology and population health intervention research. We also highlight the tension between epidemiologically-based evaluations of highly-controlled interventions, and the reality of assessing change in complex systems where the researcher does not control the implementation of the intervention - which is generally the case with policies. Current population health intervention research is often engaged in applying methods from one paradigm (EBM) to another (systems thinking). However, there is no simple rapprochement between the two, and we are often faced with a simple choice; conducting formal, epidemiologically-based evaluations, while accepting that it is often a biased and potentially misleading representation of the complex social world; or to adopt a systems-level approach to evaluation which takes account of the fluid, dynamic and interlinked nature of populations, interventions and contexts. While the latter is much discussed, the practical application of systems thinking still requires development. In the interim population health researchers continue to force the "square peg" of complex interventions into the "round hole" of clinical epidemiology.

There are several small but valuable steps which could be taken now. The first is to develop new critical appraisal tools for population health intervention studies which are focused on evaluation of change in systems over time, and which negotiate the vagaries of multiple and changing baselines, the problem of identifying controls, and so on. This would have practical value for systematic reviews, for which current practice often involves adapting existing tools originally intended for appraising clinical or behavioural studies. It would mean that studies in this field would not be downgraded on the basis of their inability to meet inappropriate criteria; and would upgraded if they appear to be consistent with what would be considered best practice within the research area. Reporting guidelines for journals publishing such studies could also be revisited to ensure that they are fit for purpose – for example, to capture relevant aspects of populations, contexts and interventions, and outcomes that are both intended and unintended. Improving understanding among researchers, funders, and users that these different underlying epistemologies do not mean different (lower) standards is also essential. For researchers, discussion of "gold standards" in evidence needs to be shelved. It is of little utility or relevance to population health intervention research, and the concept is outdated even in EBM. (31)

The most productive change now is likely to be a change of perspective, rather than a change of tools. Adopting Hawe and Shiell's perspective of interventions as "change in systems", as opposed to discrete on/off events, will be transformative for PHIR.(10) (17) (20) Teasing out the implications of this perspective for evaluation, appraisal, synthesis, study reporting and decision-making, while learning from past examples, is an imperative.

Acknowledgements and funding: Sarah Viehbeck was funded through an Emerging Researcher Award through the Population Health Improvement Research Network and through Visiting Scholar support provided by the London School of Hygiene and Tropical Medicine. Mark Petticrew received support for this manuscript through funding from the UK National Institute of Health Research, and the CIHR International Collaboration on Complex Interventions (ICCI). Steven Cummins is supported by a United Kingdom National Institute for Health Research (UK-NIHR) Senior Research Fellowship.The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the UK-NIHR or the UK Department of Health. A preliminary draft of this manuscript was shared at a writing workshop supporting by the Canadian Institutes of Health Research-Institute of Population and Public Health and the Center for Community Health Development. The authors acknowledge Erica Di Ruggiero, Nancy Edwards, Penny Hawe, and Ken McLeroy for contributing to discussions which informed the ideas in the manuscript.

References

1. Hawe P, Samis S, Di Ruggiero E, Shoveller J. Population health intervention research initiative for Canada: Progress and prospects. New South Wales Public Health Bulletin. 2011;22(22(1-2)):27-32.
2. Bambra C, Gibson M, Sowden A, Wright K, Whitehead M, Petticrew M. Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews. Journal of Epidemiology and Community Health. 2010;64(4):284-91.
3. Thorngate W, Faregh N, Young M. Mining the archives: Analyses of CIHR research grant adjudications. Available at: http://http-server.carleton.ca/~warrent/reports/mining_the_archives.pdf. Accessed 11th March 2014. 2002.
4. Corporate affairs plan. November 25th, 1987. 2501254715/4723. Available at: http://legacy.library.ucsf.edu/tid/itj32e00. Accessed 23rd July 2014.
5. Memo on preparation and handling of halal foods, 12th November 1998. Available at: http://legacy.library.ucsf.edu/tid/uau30d00. Accessed 23rd July 2014.
6. Description of Martin Haley's activities for RJ Reynolds. Available at: http://www.sciencecorruption.com/ATN178/00514.html. Accessed 23rd July 2014.
7. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Tugwell P ea. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Clinical Epidemiology Unit, University of Ottawa, 1999., .
8. Krieger N. Who and What Is a "Population"? Historical debates, current controversies, and implications for understanding "Population Health" and rectifying health inequities. Milbank Quarterly. 2012;90(4):634–81.
9. Christakis N, Fowler J. The Spread of Obesity in a Large Social Network over 32 Years. N Engl J Med. 2007;357:370-9.
10. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. BMJ. 2008 336(7656):1281-3.
11. Motlagh B, O'Donnell M, Yusuf S. Prevalence of cardiovascular risk factors in the Middle East: a systematic review. Eur J Cardiovasc Prev Rehabil. 2009;16(3):268-80.
12. Fowkes F, Stewart M, Fowkes F, Amos A, Price J. Scottish smoke-free legislation and trends in smoking cessation. Addiction. 2008;103(11):1888–95,.
13. Nykiforuk C, Eyles C, Campbell H. Smoke-free spaces over time: a policy diffusion study of bylaw development in Alberta and Ontario, Canada. Health Soc Care Community. 2008;16(1):64-74.
14. Procedure to obtain halal certification. Bates No: 524433842. Available at: http://legacy.library.ucsf.edu/tid/exu03c00. Accessed 20th December 2013.
15. Diepeveen S, Ling T, Suhrcke M, Roland M, Marteau T. Public acceptability of government intervention to change health-related behaviours: a systematic review and narrative synthesis. BMC Public Health. 2013;13:756.
16. Kouri D. Knowledge exchange strategies for interventions and policy in public health. Evidence & Policy. 2009;5(1):71-83.
17. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. Am J Community Psychol. 2009;43(3-4):267-76.
18. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 20112011.

19. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new MRC guidance. J Epidemiol Community Health 2012;66:1182-6.
20. Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? British Medical Journal. 2004;328:1561-3.
21. Ogilvie D, Mitchell R, Mutrie N, Petticrew M, Platt S. Evaluating health effects of transport interventions: methodologic case study. American Journal of Preventive Medicine. 2006;31:118-26.
22. Gøetsche P, Liberati A, Torri V, Rosetti L. Beware of surrogate outcome measures. International Journal of Health Technology Assessment. 1996;12:238-46.
23. Tu Y-K. Commentary: Is structural equation modelling a step forward for epidemiologist? Int J Epidemiol. 2009;38(2):549-51.
24. Chalabi Z, Lorenc T. Using Agent-Based Models to Inform Evaluation of Complex Interventions: Examples from the Built Environment. Prev Med. 2013;57(5):434-5.
25. Lorenc T, Oliver K. Adverse effects of public health interventions: a conceptual framework. J Epidemiol Community Health. 2013;CHECK(CHECK):CHECK.
26. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. J Epidemiol Community Health. 2013;67(2):190-3.
27. Glossary of Health Impact Assessment: World Health Organisation. Available at: http://www.who.int/hia/about/glos/en/.
28. Mindell J, Ison E, Joffe M. A glossary for health impact assessment. J Epidemiol Community Health. 2003;57:647-51
29. Potvin L, McQueen D, editors. Health Promotion Evaluation Practices in the Americas: Values and Research. New York: Springer; 2008.
30. Paul Cairney: Politics & Public Policy. Available at: https://paulcairney.wordpress.com/tag/defining-policy/. Accessed 24th Nov 2015.
31. Sackett D, Wennberg J. Choosing the best research design for each question. British Medical Journal. 1997;315:1636.