‘More research needed’: plugging gaps in the evidence base on health inequalities

Introduction

Policymakers and public health researchers have demanded better evidence of the effects of interventions on public health and health inequalities and have pointed to the relative absence of rigorous outcome evaluations in this area. While most of these commentators agree that health inequalities are a problem, the means of tackling them often seems to rest on poor evidential foundations. This is no longer surprising, for the problem has been well-documented and the solution probably seems clear: if there is not enough evidence, then obviously we need more; and if the evidence is ‘weak’, well, then, it just needs to be ‘stronger’ in future. The funding environment in the UK, at least, is supportive of the production of this new, stronger evidence, with several new funding initiatives which are aiming to foster the production of new evaluations of public health interventions. However, it might be useful to consider further what ‘more evidence’ might look like (that is, what sort of evidence is needed) and how might it be ‘stronger’ than in the past and also to consider what additional challenges collecting this new evidence might bring.

What sort of evidence is needed?

It is clear that evidence on the effectiveness and cost-effectiveness of public health interventions is often missing. Sometimes this is because policies are insufficiently subjected to outcome evaluation, perhaps because it is assumed that they are mostly beneficial and any positive outcomes can be taken as read. Randomized controlled trials of policies are particularly uncommon and there are a variety of real and perceived barriers to trials. This means that the public are frequently ‘enrolled’ in real-life policy ‘experiments’ without giving their explicit consent, or indeed without any real prospect of anyone learning anything substantial about the effects of those interventions. Although the experimental nature of most policies is generally unacknowledged, this is not always the case; in a few cases in the UK in the 1990s it was acknowledged explicitly that government policies really were a form of ‘experiment’, though not in the epidemiological sense of the word. For example, the former Health Secretary Kenneth Clark famously called the NHS ‘a great experiment’, and he also referred to his health service reforms in the early 1990s in these same terms. His fellow minister, Nigel Lawson, also described his macro-economic policies as ‘The British Experiment.’ However policy interventions like these are only ‘experiments’ in the weakest sense of the word – in the words of a recent report, ‘experiment’ often just means ‘Trying It Out’.

There is no lack of non-experimental evidence, however. There is, for example, a wealth of aetiological evidence, and we also have several influential empirically based theories about the societal, social, psychosocial and other causes of inequalities. However, it often appears to be difficult to translate this information into new interventions and even when the interventions are implemented, their evaluation is often problematic. Some of the major social determinants of health and health inequalities are difficult to randomize for practical or political reasons; one’s social class for example, is not amenable to experimental manipulation and some forms of major structural intervention may be similarly problematic. The relative paucity of experimental evidence in the social sectors compared to healthcare can be seen by comparing the number of RCTs and possible RCTs in the Campbell and Cochrane databases: about 11 000 trials appear in the Campbell database, compared with over 400 000 in CENTRAL – the Cochrane Central Register of Controlled Trials. This disparity is only partly accounted for by the difference in age and resources between the Campbell and Cochrane Collaborations.

Producing more and better, evidence will however take time and while new studies are being commissioned, it is likely that the evidence base about effects of interventions is likely to remain ‘weak’ for some time to come. This suggests that while waiting for the gaps to be filled, we need to make better use of some of the evidence we do have – that is, observational epidemiological research on the social determinants. Such studies can be employed in the cause of evaluation (monitoring of changes in outcomes over time), in directly identifying points where new interventions may be directed and in suggesting which new interventions may be appropriate. However, such studies are probably under-used for these purposes and are certainly under-synthesized.

Discussions of public health evidence often tend to be quite fragmented and we usually focus on intervention studies (and their absence) to the exclusion of the rest of the evidence base. We also tend to overlook the extent to which these different types of evidence inter-relate (or should inter-relate), for example, how observational studies may be carried out as a step towards the design of new RCTs or are sometimes conducted where RCTs are not feasible. Focusing on the lack of RCTs however means that researchers tend to err on the side of scientific conservatism, where the evidence is simply never strong enough. Policymakers however are probably less interested in the evidence we don’t have, than in which direction the evidence is pointing (with suitable caveats). To quote one policymaker: ‘Rather than admitting to politicians that the studies are ‘poor’, [researchers] have to say that we have accumulated clear and consistent evidence within limits that points to certain definite impacts’.

Ranking evidence

One reason why we focus on the gaps in the evidence base is that we tend to separate it into mutually-exclusive categories, with an implicit ranking – sifting it into ‘wheat’ and ‘chaff’ in one widely-used metaphor. Sorting evidence in this fashion is useful when there is a lot of evidence to sift through; given the choice between a trial and an uncontrolled study of the effects of an intervention, we would probably prefer the trial and in turn we would prefer the uncontrolled study to the unsupported anecdote. However – to stretch the
winnowing metaphor further – when wheat is in scarce supply, no amount of sifting makes it better or any more abundant. Similarly, given a weak evidence base, ranking evidence does not tell us how to integrate different sorts of evidence, nor does it recognize the fact that different types of evidence are generally collected for different purposes. In public health we need to be more concerned with integrating evidence and using the evidence we already have, to identify plausible new evaluative studies, than with ‘ranking’ evidence.

### Problems with outcomes

The problems of the patchy evidence base are often discussed, but there are other more basic problems which generally receive less of an airing. One such problem relates to the choice of appropriate outcomes to measure. In public health we are sometimes guilty of a sort of naive ‘healthism’ in talking about our need for ‘better evidence’. For most of the main social determinants of health and health inequalities, ‘health’ is not the primary intended outcome and so, by default, evaluations which measure change in health status are uncommon. To take one example: housing is widely assumed to be a major determinant of health and health inequalities. However, the primary intended outcome of investment in new housing is not ‘health’. It is the provision of new, warm, dry, safe houses. The evidence that new houses were actually provided is easily obtained and a ‘good enough’ answer to the question requires evidence from pretty near the bottom of the hierarchy of evidence (i.e., the evidence of ones’ eyes). At most, it requires ascertaining that the houses were built and asking tenants if they are more satisfied and happier with their houses. This sort of evidence is indeed already collected routinely by most housing providers, whereas evidence on less obvious outcomes – the health impacts of their activities is generally not collected. The same issue applies to many other social interventions. It is not surprising that there is little robust evidence on the health impacts of social policies, because health improvement is often an unintended, spillover effect. Identifying meaningful primary and secondary outcomes and powering studies to detect them in evaluations of social interventions can be difficult and risks taking a narrow view of the intended purposes of the intervention in question.

So, while it is widely accepted that the main determinants of inequalities are social determinants, ‘health’ and ‘health inequalities’ may understandably not be at the top of most pressured policymakers list of priorities and developing the evidence base means making public health and health inequalities part of their business and this remains a major challenge. Health Impact Assessment (HIA) is intended to help with this task by ensuring that all major policies are assessed with respect to their impact on health and inequalities. However, the Catch-22 is that HIA depends on the provision of evidence about those effects – and we already know that this evidence is often absent.

### Ecologies of evidence?

McLaren and Hawe suggest that ‘ecological thinking’ may be helpful in public health. Such an ecological perspective takes into account the individual and environmental determinants of health and other behaviours.

Perhaps such ecological thinking can also be applied to public health evidence; it may be useful, for example, to consider the wider ‘ecology’ within which public health evidence exists and in which different sorts of evidence occupy different niches (i.e., where different study designs have different degrees of ‘fitness’ for different purposes). To pursue the analogy, in a food chain certain animals interact with, feed on and share space with other organisms. In a linear food chain polar bears eat seals, which eat shrimp, which in turn eat plankton. More complex organic relationships are described as ‘food webs’. A similar ‘ecology of evidence’ would highlight the relationships between different types of evidence and what they each contribute over time to the evidence base: for example, observational studies are used to inform the development of RCTs, which may themselves involve parallel qualitative studies, all of which are eventually subsumed within new systematic reviews, which inform the development of new observational and evaluative studies.

This does not deny the strength of experimental designs and the fact that they provide strongest evidence of effectiveness, but recognizes that they may not always be possible and that we need to use what evidence is available. In some cases, Wanless and others (including some health inequalities researchers) have suggested that the gaps in the evidence base may be partially filled by exploiting the opportunities offered by ‘natural experiments’, such as changes in employment opportunities, housing provision or cigarette pricing.

Natural experiments in the case of health inequalities can help investigate the determinants of those inequalities as well as assisting in the identification of potentially effective interventions. They can also indicate the scale and nature of impacts in situations where randomized controlled trials are not available. However, they bring their own challenges, chief among them being potential problems with internal validity, which makes drawing strong, unbiased causal inferences from such difficult studies and this has led to the occasional suggestion that such non-randomized studies should carry a ‘health warning’. This does not mean however, that they cannot provide valuable and valid information.

### Conclusion

Implicit in debates about public health evidence is the assumption that the evidence base is something which needs to be assembled, as quickly as possible, in order to tell us definitively what to do. Of course, it is never that easy. The amount of evidence we ‘need’ cannot easily be ascertained. There will never be ‘enough’ evidence in absolute terms and the evidence base will probably always be criticized as being ‘weak’. Deciding how much evidence is ‘enough’ may just be another way of deciding how much uncertainty we are prepared to live with.

A gap – in an evidence base or anywhere else – seems a simple enough problem to deal with – you just need to close it. However, not all gaps are easily closed and part of the challenge is to correctly identify those gaps and then use that information to identify where more evaluations of the effects of public policies on human health and wellbeing are most needed. The results of those studies may take time to emerge. In the interim it therefore seems particularly important to learn to make better use of whatever evidence is already available and whatever its limitations.

### Acknowledgements

I would like to thank Matt Egan and Hilary Thomson for their advice. The idea from the article came from discussions in Calgary in 2006 at the International Collaboration on Complex Interventions funded by the Canadian Institutes of Health Research.

### References

Use evidence to expose the unequal distribution of problems and the unequal distribution of solutions

Mark Petticrew argues that we need to increase the size and strength of the evidence base on policies and programmes to reduce inequalities in health. He cautions us not to be drawn unnecessarily into (ill informed or tangential) debates about methods. It would be if, while our research is proceeding, much can be done with what evidence we have. He is right on all three fronts.

Primary research is sorely needed to establish the effectiveness and economic efficiency of many interventions that could make a major impact in health inequalities. The current evidence imbalance is an embarrassment. A vet treating a puppy for an ear infection can draw on level 1 evidence on the options available, while policy makers responsible for reducing inequalities in human populations are forced to work with far less certainty.

But what to do with what we have got? More than 20 years ago Murrell argued that endless tabulations of data about the distributions of social problems are among the least likely of all research processes to alter the status quo regarding what is done to address them. Yet that is what most of us find easiest to do in population health: our report cards on our cities, regions or countries documenting the finer points of how health inequalities are getting worse.

One pathway is the type of data-for-action impetus that comes from the intense, in-it-for-the-long-haul university-community research partnerships that colleagues in some of the poorest communities have established. This requires a special type of research and we would argue, a special type of researcher.

In addition, we could simply reorient the day jobs of regular population health data analysts. Imagine how powerful it would be if, alongside maps of the areal distribution of smoking rates, obesity, premature mortality or per capita prescriptions for depression, our analysts were mandated to report on the distribution of polices and practices known to be effective in their prevention? Like smoke free public places, fair wage policies, affordable fresh food, confectionery-free schools, family friendly workplaces and so on? Imagine how much easier it would be for the public to claim their entitlement to prevention if they could see how unequally (and potentially unfairly) the policies and practices are currently distributed. ‘Geography is destiny’ is a phrase well known in some circles. If this were made more public it would be political dynamite, not for the factors the public tend to think they cannot change (who they are and where they live), but because such maps would expose the inertia of those responsible for allocating the resources that can alter those destinies.

In the field of tobacco control, Glantz and Balbach tell us that research evidence documenting the harmful effects of tobacco and the effectiveness of measures to prevent its use provided an essential backdrop. But it was never going to be enough to change policy. What was critical was the creation of a constituency for change mobilized in part by creative re-framing of the issues being addressed (from individual choice to the public good) and of the solutions being offered. This strengthened decision maker resolve and helped overcome opposition. It was not the quality of the evidence that counted so much as its weight and the way this was brought to bear on the problem.

Let’s learn from this. Population health data observatories have significant expertise and experience in mapping health behaviours, health outcomes and health inequalities [see for example the annual reports of the Association of Public Health Observatories in England (www.AHPO.org.uk)]. Inventories of relevant policies in some sectors are already being collected systematically [see Journal of School Health 2001; 71(1) Supplement, which describes the CDC School Health Policy and Programs study]. The next step is to bring these two together. In the field of health inequalities, our capacity to make a difference may be closer than we think.

Acknowledgements

P.Hawe and A.Shiel are Health Scientists funded by the Alberta Heritage Foundation for Medical Research. P.Hawe holds the Markin Chair in Health and Society. The Population Health Intervention Research Centre is a centre for research development in population health funded by the Canadian Institutes of Health Research.

References


Penelope Hawe, Alan Shiel
Population Health Intervention Research Centre, University of Calgary, 3330 Hospital Drive NW, Calgary T2N 4N1, Alberta, Canada
Correspondence: Alan Shiel
e-mail: ashiel@ucalgary.ca

doi:10.1093/eurpub/ckm095
Advance Access published on 30 August 2007