Evidence for public health policy on inequalities: 2: Assembling the evidence jigsaw

Margaret Whitehead, Mark Petticrew, Hilary Graham, Sally J Macintyre, Clare Bambra, Matt Egan

EVIDENCE BASED PUBLIC HEALTH POLICY AND PRACTICE

Study objective: To garner research leaders’ perceptions and experiences of the types of evidence that influence policy on health inequalities, and their reflections on how the flow of such research evidence could be increased.

Design, setting, and participants: Qualitative two day residential workshop with senior research leaders, most of whom were currently involved in evaluations of the health effects of major policies. In four in depth sessions, facilitated by the authors in turn, focused questions were presented to participants to reveal their views and experiences concerning evidence synthesis for policy on inequalities. These wereanalysed thematically.

Main results: Five types of evidence for policy on health inequalities were felt to be particularly persuasive with policymakers: observational evidence showing the existing of a problem; narrative accounts of the impacts of policies from the household perspective; controlled evaluations; natural policy experiments; and historical evidence. Methods of improving the availability and use of these sources of information were put forward.

Conclusions: This paper and its companion have considered the current evidence base for policies to reduce health inequalities, and how this could be improved. There is striking congruence between the views of the researchers in this study and policy advisers in paper 1, suggesting that a common understanding may be emerging. The findings suggest significant potential for rapid progress to be made in developing both evidence based policy, and policy relevant evidence to tackle inequalities in health.

METHODS

The authors, as members of the ESRC Centre for Evidence-based Public Health Policy, organised a two day residential workshop at a conference centre in Sussex, England, in October 2002. Eight senior researchers from the UK and one from overseas were selected purposively to have the following characteristics: over 10 years’ experience of synthesising and evaluating the evidence on major policies related to health inequalities; current leaders in their field as indicated by relevant research awards for which they were principal investigators; publications and supervision of teams of researchers under these grants; expert advisors to national and international policymaking on the implications of research for policy. As comparatively few people currently work in this field and have the above characteristics, most participants were known to each other and knew one or more of the authors. Participants were assured that their remarks would not be attributed and would be kept confidential. Four focused sessions over the two days were facilitated by the authors in turn, whose role was strictly neutral in that they introduced the focused questions and chaired the sessions, but refrained from voicing their own opinions. Participants were asked to concentrate on a series of questions and the ensuing dialogue was transcribed independently by two rapporteurs (CB, ME). The participants were first asked to give illuminating examples of research evidence that had influenced public health related policy around health inequalities. Secondly, they were asked about the nature and location of the evidence and what constitutes good evidence more generally in relation to initiatives to reduce health inequalities. The rapporteurs compared notes to produce a single agreed transcript. MP, MMW, CB, and ME coded the transcript, with input from the other authors, and from this identified the main themes. These were discussed with all the...
authors to achieve agreement. We then analysed this transcript for emerging pointers toward the means of assembling policy relevant evidence for tackling health inequalities in the future.

RESULTS
What type of evidence has an impact on policy?
Research leaders were asked to give examples from their own experience of research that had had an impact on public health policy related to health inequalities. Their replies illustrated that such research has a diversity of purposes and sources, grouped into five main categories (box 1).

(1) Observational studies identifying a problem
Perhaps the simplest example of evidence leading to policy action concerned research that identified the existence of a problem, and for which the intervention to tackle the issue was then fairly obvious. One example given was the observational studies showing the association between environmental radon and lung cancer. This evidence highlighted the problem, and led to housing interventions to protect people living in high radon localities. For example, in certain high risk areas in the UK, such as Cornwall, grants were given for housing ventilation to reduce the risk of radon associated disease. Another classic example cited was the epidemiological studies showing the health damaging effects of tobacco exposure, which stimulated a raft of tobacco control policies around the world. These have reduced tobacco consumption to varying degrees, and the focus is now on determining which combination of interventions would be most effective in specified contexts.

Box 1 What type of evidence has an impact on policy to tackle health inequalities? Researchers’ views

(1) Observational studies that identified a problem
"Evidence based policy’s successes are where the intervention is quite easy once the problem has been correctly identified. Bringing up a problem with an easy solution will have more impact. It is trickier when there is an identifiable problem and the solutions are not clear or they are complex—health inequalities. What to do about this is difficult."

(2) Modest, but politically timely, household studies
"The power of qualitative evidence is high. Qualitative evidence paired with quantitative is much more persuasive."
"Timing is crucial. A small piece of research can have an impact if timed rightly"

(3) Controlled evaluations of interventions
"In 1960s America, there was a more straightforward relationship between research and evidence as policymakers seemed to be waiting for the results of certain trials before they made their policy decisions … this is now rare."

(4) Natural policy experiments
"There are natural policy experiments going on that are important … but … (there are) limitations here to the kinds of controls/comparisons that can be drawn. Hard to get credible research designs in natural experiments."

(5) Historical evidence with a long shelf life
"Policy-makers expect evidence to pop up when they want it. Lots of things in the past e.g. Rowntree’s studies still have influence on our social security system—a milestone effect. Research from a long time ago can find itself valuable."

The participants recognised that these examples involved single risks, so the evidence was easier to establish and the interventions themselves were simpler. The task was felt to be much more difficult when addressing health inequalities, which have multiple causes, rather than a single determinant (box 1).

(2) Modest, but politically timely, household studies
The persuasiveness of evidence from studies of the reality of life in different kinds of households was highlighted. Sometimes these were descriptive studies, reporting household budgets and analysing expenditure on the prerequisites for health for families living in different socioeconomic circumstances. Qualitative studies of decision making in families—exploring why one wanted to go over another—have also had powerful policy impacts. Two examples from the first term of the Labour government were used as illustration.

One study of the costs of raising children showed that in the 1990s parental expenditure on the essentials for young children was as high as for older ones, but the official benefits for younger children were much lower. The findings of this study were disseminated at the same time as the decisions on the 1998 budget were being made and appeared to have an immediate impact upon it, with an increase in benefit rates for the under 11 year olds. The research evidence was actually cited by ministers in support of the decision.

The speed of reaction, however, may indicate a prior decision, with the research evidence used as extra ammunition. This research is still having an impact on social security policy as in the new Child Tax Credit (effective from April 2003) all children over 1 year old attract the same rate of allowance.

The second example cited was a qualitative study of income distribution within households in receipt of benefits. This study showed significant gender inequalities in the control and distribution of household income. Women were more likely than men to have the burden of financial management in low income families, and to “go without” to meet their children’s needs, whereas men had greater financial control of overall income and were much more likely than women to indulge in high levels of personal spending. The study’s findings implied that the government’s 1998 plans to replace Family Credit (paid to the mother) with the Working Families Tax Credit (paid through the wage packet to the main earner) might change the perception and allocation of the money. It was a modest study (31 subjects), but its timing was crucial, as it influenced the final legislation that permitted couples to choose whether the benefit would be paid as a tax credit via the wage packet or as a cash benefit. Again, this research has had a continued impact on social security policy, as the new Child Tax Credit will be paid to the main carer of the child.

(3) Controlled evaluations of interventions
Several examples of experimental evidence influencing non-healthcare policy were given, although it was emphasised that controlled evaluations were still the exception rather than the rule. One experimental study, funded by the US Department of Labor in 1971, was started in response to a long running debate about whether extending unemployment benefit to recently released prisoners would reduce re-offending rates. Some 432 prisoners who were about to be released from prisons in Maryland were randomly assigned to one of three treatment groups (payments only, counselling only, both combined) and a control group. The results of the study showed that the likelihood of re-offending was lower in the payment intervention group than in the controls, and this finding was supported by later studies. Policy changes
directly followed from this finding in some states; for example, California extended unemployment benefit payments to released prisoners in 1978.25 The Headstart programme of pre-school education and child care in the 1960s in the USA was also subject to rigorous evaluation. Early positive impacts on child development in children experiencing the programme compared with controls brought about a boost in funding for such programmes. When, however, the gains seemed to disappear after four or five years, the funding wavered. It was not until the children were in their teenage years that “sleeper effects” of the pre-school programme reappeared, in terms of greater likelihood of staying on at school, of getting a job and earning more money, and of lower rates of teenage pregnancy and arrests.17–20 These results have subsequently provided compelling evidence for the setting up of the UK’s “SureStart” programme of early education and care introduced in Britain in the late 1990s. The same effect of rigorous research on specific issues can be observed in developing countries. In Africa, for example, research on vitamin A and mortality, and on rehydration for diarrhoeal diseases have all had an immediate impact on worldwide policy and practice after recommendations from WHO.21 22

Contrasting examples—of evidence that has prevented ill advised policies—were also presented, including the decision not to proceed with mass screening for prostate screening in the mid-1990s. Conservative ministers at the time were attracted to the idea, as it could be seen as doing something about men’s health. However, the reviews undertaken by the NHS Centre for Reviews and Dissemination in York and internationally23 concluded that there was insufficient evidence for proceeding and that in fact prostate screening may be harmful. Initially, the evidence was ignored, but when the researchers became more active in disseminating their results through press conferences the policymakers did reverse their decision.

(4) Natural policy experiments
Participants suggested that evidence from other countries or regions could inform debate if it provided concrete information on what happens when a particular policy currently under consideration in one place has already been introduced elsewhere. Evidence from surveys of public attitudes to proposed changes could add to that debate. In 1993, for instance, the New Zealand Blood Transfusion Service became part of a new hospital reorganisation that was required to be profit driven. The results of a survey on the motivations of blood donors suggested that 52% of New Zealand’s blood donors would cease donating if profit were to be made from the blood.24 This would have a massive impact on the ability of the service to match demand. As a result of the widespread publicity these findings received, the idea to make a profit from blood donations was scrapped.

(5) Historical evidence with a long shelf life
Several researchers emphasised that evidence from the past, even from many decades ago, should not be overlooked, and in fact may be particularly influential in the public health field. Much of the influence is unconscious, but none the less powerful in setting the policy agenda. Two examples were cited, one being the Rowntree Poverty Surveys of 1901 and onwards.25–27 These painted a vivid picture of life in the slums of Britain’s industrial cities and were shocking to the general public at the time. Their enduring legacy, however, has been their influence on public attitudes to poverty and the poor throughout the 20th century, which underpinned the building of the postwar welfare system.28

There was also the “sleeper effect” of evidence produced in a “cold climate”, which at the time may seem to have little or no immediate impact, but was stored and used when a more favourable political climate develops. One example related to suicide and health inequalities. In Scotland observational evidence on the association between unemployment and parasuicide produced in the 1980s29 was greeted frostily by the Scottish Office, and attempts were made to discredit it. During the late 1990s, however, the subject was revived when studies showed that completed suicide rates among young men in Scotland had increased for 30 years.30 A 1998 conference on the topic received wide media coverage, and in 2000 an opposition initiated debate on suicide was held in the Scottish Assembly. This in turn kick started the development of a suicide prevention strategy.31 32

How can the availability of these types of evidence be improved?
The participants then turned to the question of how the availability of these types of evidence could be improved (box 2).

(1) Assembling “the evidence jigsaw”
One prominent theme running through the dialogue was the diversity of sources of policy relevant evidence. The examples cited above came from classic observational studies, intervention research, natural policy experiments, and from quantitative and qualitative explorations of the real life impacts of policies. This reflects participants’ views of a “jigsaw” of evidence, by which they meant the need to synthesise evidence from diverse sources to make the causal links that would inform policy.

More often than not, the most valuable policy relevant information was not one single piece of evidence, but rather many different bits, of varying quality, creatively pieced together. For example, participants pointed out that
observational studies—such as in the example of radon and lung cancer—can identify public health problems, while other types of descriptive study (such as household studies and qualitative research) can highlight the impact on people, and experimental and quasi-experimental data (such as that derived from “natural experiments”) can guide interventions. Historical information—such as the Rowntree surveys—would shed light on the wider context. As well as new studies, participants recommended the appropriate replication of existing studies—for example, testing the same intervention in different areas, communities, and time periods.

(2) Nurturing an “evaluation culture”

One reason for the lack of relevant evidence was felt by participants to be the historical absence in the UK of an “evaluation culture”, meaning a research and policy environment within which the evaluation of the outcomes of interventions is common. This situation was contrasted with an earlier period in the USA, where in the 1960s and 1970s a strong “evaluation culture” existed and the costs of evaluation were actually built into social programmes. A 1969 Presidential Executive Order, for example, mandated that 1% of programme budgets were to be spent on evaluation, and in 1972 this amounted to $100 million. Subsequently it is possible to identify a string of studies from this period that have influenced policy. The fostering of an evaluation culture in the UK in the policy field was felt by researchers to require the removal of practical barriers, such as the bias towards short term grants, as well as the funding of longitudinal research, and improving training for researchers in evaluation and research synthesis methods. Redressing some of these problems would, it was felt, help promote the same intervention in different areas, communities, and time periods.

(3) Closer engagement between research and policy

Researchers felt that some policymakers lacked understanding of research and had inflexible ideas about research methods. They could hinder project evaluations by, for example, objecting to certain methodological principles (such as randomisation), and were unable to tolerate uncertainties in evidence—while their own certainties hindered evaluation. Researchers acknowledged, however, that they themselves needed to express uncertainty in a more accessible way, and to learn to summarise and disseminate their own research much more effectively. Researchers proposed that closer engagement could be promoted if both parties consulted at the earliest possible stage on tenders, and that more researchers and policy makers should experience periods of working in each others’ fields. Secondment of researchers to policy units for a period of time and vice versa was one suggestion.

DISCUSSION

Starting from the perspectives of researchers and policymakers, this paper and its companion have considered the current evidence base for policies to reduce health inequalities, and how this could be improved. It was striking that there was significant congruence between the views of the research and policy communities, suggesting that there may be a common (though not necessarily shared) understanding.

The most noticeable similarity relates to the types of evidence thought to have the most powerful impact on policy. Although evidence based policymaking has sometimes been seen as pertaining solely to the use of experimental evidence, this notion was rejected by both sets of participants in relation to the complex issue of inequalities. The research leaders outlined a wider raft of evaluative and other qualitative and quantitative evidence that could be used to support policy making around inequalities. This understanding of an “evidence jigsaw” accords well with the policymakers’ views in our first study (paper 1) of a “mixed economy” of evidence, in which heterogeneous types of evidence, varying with respect to methodological “rigour,” are brought to bear on different aspects of any particular policy question. The compelling nature of evidence from qualitative case studies of everyday lives was acknowledged by both researchers and policy advisers, as was the (sometimes hidden) impact of historical data. This “evidence jigsaw” itself reflects a longstanding awareness in social science and epidemiology that “appropriateness of evidence” is usually more important than “hierarchies of evidence”. The notion of “appropriate” in this study implied that evidence should be matched to the research question. As the determinants of health inequalities are themselves complex, and socially and historically determined, different types of data are needed to identify the
mechanisms and suggest areas in which to intervene. We still need to move beyond descriptions of processes and meanings, though, towards evaluations of the actual outcomes of those interventions, if progress towards reducing inequalities is to be made.

Aside from an acknowledgement by both groups that there is a need for less narrow views of what constitutes evidence, another important message to emerge was that there are specific actions that can be taken by researchers, policymakers, and funders alike that can foster the production of meaningful new research around inequalities. Chief among these actions is the promotion of an evaluation culture, entailing appropriate training for researchers; and maintaining the investment in longitudinal research, such as large scale cohort studies. Improving the inequalities evidence base would also require a commitment by policy makers to the outcome evaluation of policies, with a certain percentage of intervention budgets routinely set aside for project evaluation, with evaluation beginning at the earliest possible stage of policy conception. Closer engagement between researcher and policymakers would also be required at all stages of the research progress—from tendering onwards. This may imply more common use of cross employment, with researchers and policy makers working temporarily in each others’ fields. On the other hand, specific tasks for researchers include exploring more sophisticated means of comparing and synthesising the results of different types of research. This also requires a clear map of the specific studies that are now needed to populate the health inequalities evidence base, and more effective dissemination of research findings. This includes producing more appropriate summaries, and exploring ways to express uncertainty about those findings in an accessible way.

This study set out to explore differing perspectives on the production and use of evidence on health inequalities, but what was perhaps surprising was the extent to which the recommendations from the two groups were complementary, rather than conflicting. Researchers were for example aware of many of the barriers to the production and effective dissemination of their work to the policy community. There was also a common understanding of the need for different types of evidence to answer different types of policy question—as well as agreement on the need for better data on the effectiveness and cost-effectiveness of policies. These findings suggest that there may now be significant potential for rapid progress to be made in developing both evidence based policy, and policy relevant evidence to tackle inequalities in health. These two papers give practical pointers to both policy and research communities to help achieve these two aims.

Authors’ affiliations

M Whitehead, C Bambra, Department of Public Health, University of Liverpool, Liverpool, UK
M Petticrew, S J Macintyre, M Egan, MRC Social and Public Health Sciences Unit, University of Glasgow, Glasgow, UK
H Graham, Institute for Health Research, University of Lancaster, Lancaster, UK

Funding: this study was funded as part of the ESRC’s ‘Evidence-Network’ project grant number H141251011. Mark Petticrew is funded by the Chief Scientist Office of the Scottish Executive Department of Health.

Conflicts of interest: none declared.

REFERENCES

28 Beveridge W. Inter-departmental committee on social insurance and allied services. London: HMSO, 1942.