DRAFT: 2009 06 21

Helping to ensure well-informed public policy decisions:
A framework for mandatory impact evaluation

Andrew D. Oxman, Norwegian Knowledge Centre for the Health Services, Norway
Arild Bjørndal, Norwegian Knowledge Centre for the Health Services and University of Oslo, Norway
Francisco Becerra, Council on Health Research for Development, Mexico
Mark Gibson, Center for Evidence-based Policy, USA
Miguel Angel Gonzalez Block, Centre for Health Systems Research, National Institute of Public Health, Mexico
Andy Haines, Director’s Office, London School of Hygiene and Tropical Medicine, UK
Maimunah Hamid, Institute for Health Systems Research, Ministry of Health, Malaysia
Carmen Hooker Odom, Milbank Memorial Fund, USA
Haichao Lei, Department of Policy and Regulation, Ministry of Health, P.R. China
Ben Levin, Ministry of Education, Ontario, Canada
Mark W. Lipsey, Peabody Research Institute, Vanderbilt University, USA
Julia H. Littell, Bryn Mawr College, USA
Hassan Mshinda, Commission for Science and Technology, Tanzania
Pierre Ongolo-Zogo, Centre for the Development of Best Practices in Health, Yaounde Central Hospital, Cameroon
Tikki Pang, Research Policy & Cooperation Department, World Health Organization, Switzerland
Nelson Sewankambo, Office of the Principal, College of Health Sciences, Makerere University, Uganda
Francisco Songane, Partnership for Maternal, Newborn & Child Health, Switzerland
Haluk Soydan, Hamovitch Center for Science in the Human Services, University of Southern California, School of Social Work, USA
Carole Torgerson, Institute for Effective Education, University of York, UK
David Weisburd, Institute of Criminology, Hebrew University, Israel and George Mason University, USA
Judith Whitworth, John Curtin School of Medical Research, Australian National University, Australia
Suwit Wibulpolprasert, Office of Permanent Secretary, Ministry of Public Health, Thailand

Acknowledgements
We thank the following for comments on drafts of this article: Bertha Briceño, Iain Chalmers, Luis Gabriel Cuervo Amore, Octavio Gómez Dantés, Manuel M. Dayrit, Mark Petticrew, and Howard White.
Trillions of dollars are invested yearly in programmes to improve health, social welfare, education and justice (which we will refer to generally as public programmes). Yet we know little about the impacts of most of these attempts to improve peoples’ lives, and what we do know is often not used to inform decisions. We propose that governments and nongovernmental organizations (NGOs) address this failure responsibly by mandating more systematic and transparent use of research evidence to assess the likely impacts of public programmes before they are launched, and better use of well-designed impact evaluations after they are launched.

Resources for public programmes will always be limited. In low and middle-income countries (LMIC’s), where there are often particularly severe constraints on resources and many competing priorities, it is especially important to use available resources as efficiently as possible to address important problems and goals, such as the Millennium Development Goals (MDGs). Using research evidence to inform decisions, far from being a luxury, is crucial. As suggested by Hassan Mshinda, the Director General of the Commission for Science and Technology in Tanzania: “If you are poor, actually you need more evidence before you invest, rather than if you are rich.” But neither the problem nor the need for solutions is limited either to health or LMIC’s. Expenditures and the potential for waste are much greater in high-income countries, which also have limited resources and unmet needs, particularly in the face of a financial crisis. Having good evidence to inform difficult decisions can be politically attractive as illustrated, for example, by the US government’s recent decision to include $1.1 billion for comparative research (including systematic reviews and clinical trials) as part of its $787 billion economic stimulus bill.

Why would policymakers want to mandate impact evaluation?

To paraphrase Billy Beane, Newt Gingrich and John Kerry, who have argued for “a health care system that is driven by robust comparative clinical evidence” by substituting “policymakers” for “doctors”: “Evidence-based health care would not strip [policymakers] of their decision-making authority nor replace their expertise. Instead, data and evidence should complement a lifetime of experience, so that [policymakers] can deliver the best quality care at the lowest possible cost.”

We use the term ‘impact evaluation’ to refer both to assessing the likely outcomes of programmes before they are launched and to prospectively planned evaluations that are undertaken after programmes are launched to document actual outcomes relative to what would have happened without the programmes. Some policymakers may not perceive it to be in their interest to commission impact evaluations because their term of office may be short or their motivation may be ideological.

Nonetheless, making impact evaluation mandatory could have a number of advantages for a growing number of policymakers who do not share that perception, including both elected policymakers and non-elected policymakers such as civil servants. Firstly, it can reduce political risk, because it allows politicians to acknowledge that there is imperfect information to inform decisions about public programmes, and to set in motion ways to alter course if
programmes do not work as expected. There is a far greater political risk when they advocate a programme and then are stuck with it no matter what the results.

More systematic and transparent use of research to assess the likely impacts of proposed programmes could in addition better enable politicians to manage researchers acting as advocates and lobbyists misusing research evidence. It could enable them to ask critical questions about the research underlying what is being advocated, and to demonstrate that they are using good information on which to base their decisions. It also could enable them to ensure that research evaluating their initiatives is appropriate and that the outcomes being measured are realistic and agreed in advance. It puts them in the politically attractive position of continuous policy improvement and it gives them standing in the research process that they otherwise might not have.

The problem

“Because professionals sometimes do more harm than good when they intervene in the lives of other people, their policies and practices should be informed by rigorous, transparent, up-to-date evaluations.” (Iain Chalmers, Editor, The James Lind Library, 2003)

Decisions about programmes are too often made without systematically or transparently accessing and appraising relevant research evidence and without adequate evaluation of their impacts. We need to make better use of what we already know and we need to evaluate better the impacts of what we do. Reasons for our failure to do this include inadequacies with all of the following: research, access to available research, capacity to use research appropriately, and management of conflicts of interest.

Research evidence is just one input into policymaking. Other information, values, political considerations and judgements are also important. However, good intentions and plausible theories alone are an insufficient basis for decisions about public programmes that impact on the lives of others.5 6

Acknowledgements of the problem

The need to improve the use of research evidence to inform decisions about public programmes is widely recognised. For example, the 58th Session of the World Health Assembly passed a resolution acknowledging the 2004 Mexico Statement on Health Research resulting from the Ministerial Summit on Health Research and urged member states “to establish or strengthen mechanisms to transfer knowledge in support of evidence-based public health and health-care delivery systems, and evidence-based health-related policies.” The Summit requested the Director-General of WHO “to assist in the development of more effective mechanisms to bridge the divide between ways in which knowledge is generated and ways in which it is used, including the transformation of health-research findings into policy and practice.”8 In the run-up to the Summit a case was made for greater investment in health systems research to address many of the gaps in evidence on appropriate policies and strategies for improving health care.9 Currently around 0.02% of expenditure on health is devoted to such research in low and middle-income countries - far too little to provide the quality and quantity of evidence needed for informed decision making.10 The need to continue to build on what progress that has been made since the Mexico Ministerial Summit was recognised in the 2008 Bamako Statement by Ministers of Health, Ministers of Science and Technology, Ministers of Education, and other ministerial representatives from 53 countries.11
There have been numerous other calls for better use of research evidence to improve decisions about public programmes both internationally and at a national level. For example, the Blair government in the UK in its modernising government agenda stated that “government must be willing constantly to re-evaluate what it is doing so as to produce policies that really deal with problems; that are forward-looking and shaped by the evidence rather than a response to short-term pressures” and that “better use of evidence and research in policy making” was needed.12

Responses to the problem
A wide range of initiatives have been advocated and implemented to improve the use of research evidence in decisions about public programmes and to address underlying problems with research, access to available research, capacity to use research appropriately, and management of conflicts of interest. These include initiatives to:

- prioritise research and align it with countries’ needs
- commissioning research to meet the needs of policymakers for better information,
- improve the quality of research syntheses and impact evaluations,
- increase funding for research syntheses and impact evaluations,
- make research evidence more accessible to policymakers (e.g. through summaries of systematic reviews, clearing houses, and policy briefs),
- build capacity, and
- manage conflicts of interest.

A 2005 survey of organisations engaged in supporting evidence-informed policymaking identified many health technology assessment agencies and clinical practice guideline developers.13 However, the survey found few examples of organisations that support the use of research evidence for decisions about health programmes. The Evidence-Informed Policy Network (EVIPNet), one response to the Mexico Statement,14,15 is trying to address this. EVIPNet was launched by WHO in 2005. It is an attempt to strengthen the links between research and policy in LMIC’s. Teams of policymakers from ministries of health and researchers in over 20 countries have joined the network. The Region of East Africa Community Health (REACH) Policy Initiative, which was established by Kenya, Tanzania and Uganda, is another example of efforts to foster greater use of evidence in policy-making.16 However, neither EVIPNet nor REACH has sustainable funding yet.

There have been a number of initiatives to develop the capacities that are needed for evidence-informed health policymaking, but very few of these have been rigorously evaluated.Error: Reference source not found Most have been relatively limited in terms of the breadth of capacities they address, for example, focusing on the capacity for doing research. Few have focused on the capacity of policymakers and civil society to use research, or on making research evidence more accessible to policymakers.

There has been a substantial increase in the production of systematic reviews over the past two decades, including reviews that are relevant to decisions about public programmes. Two major initiatives to ensure the production, maintenance and accessibility of systematic reviews are the Cochrane Collaboration (www.cochrane.org) and the Campbell Collaboration (www.campbellcollaboration.org). The Cochrane Collaboration now has thousands of contributors from around the world and has published over 3500 reviews. However, there are still many important questions for which up-to-date reviews are not available, particularly
focusing on health issues that are relevant to people living in LMIC’s, and although many governments are providing long-term support to these efforts, most governments are not.

There also have been a number of initiatives to promote impact evaluations, including the International Initiative on Impact Evaluation (www.3ieimpact.org), efforts by the World Bank, and others.

**Examples of legislation**

Recognising its political and ethical obligation to evaluate the impact of policy decisions, the government of Mexico passed legislation requiring that impact evaluations be conducted for a variety of public programmes, explicitly recognising the value of learning what works and why as a guide for future budget decisions (Box 1).

In the US, the Education Sciences Reform Act of 2002 established within the US Department of Education, the Institute of Education Sciences (IES) (http://ies.ed.gov/). The mission of IES is to provide rigorous evidence on which to ground education practice and policy. This is accomplished through the work of its four centers.

In China, the Ministry of Health has created networks of policymakers and researchers and initiated activities to build capacity and collaboration in addressing public health challenges (http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohzcfgs/s2908/list.htm). One of the functions of these networks is to foster the generation of relevant evidence and better use of evidence in decision-making.

In Thailand, article 67 of the 2007 constitution mandates a participatory health and environmental impact assessment, before any public program that may affect health, natural resources or the environment can be started. Article 10 and 11 of the 2007 National Health Act also mandate the government to transparently provide adequate information and give the community and individuals the right to request for and participate in the assessment of the health impact of public policies.

In Colombia a law was passed in 1994 that gave the National Planning Department the mandate of evaluation, which led to the establishment of a national system for monitoring and evaluation (SINERGIA), which has been assessed by the World Bank to be one of the strongest in Latin America.  

None of these or other commitments by governments to make better use of research evidence to improve the lives of their citizens are without shortcomings. They all display different elements of what an ideal commitment might look like and there is a great deal that could be learned from both the successes and failures of efforts such as these to guide and strengthen commitments by governments and NGOs to ensure that decisions about public programmes are well informed by research evidence.

**A proposal to help ensure that decisions about public programmes are well informed**

We propose that governments and NGOs recognise their political and ethical obligations to make well-informed decisions and evaluate the impacts of their programmes in legislation. The design of this legislation should be developed based on a thorough review of international experience both with directly relevant legislation, such as the legislation in Mexico, and other relevant initiatives and legislation, for example with legislation mandating environmental
impact assessments and legislative budget processes that require fiscal impact statements to be included for new program proposals.

The rationale for developing a framework is that formal commitments (legislation) and an international framework to help design these could increase the likelihood that good intentions regarding the use of research evidence to inform policy will be translated into effective actions. The details of any legislation clearly would need to be adapted to a specific context and there are several ways in which such legislation might be introduced. For example, as part of a general appropriations act, as part of social development legislation, as in Mexico, as part of health legislation, or through regulations for government auditors.

To facilitate the development and passage of such legislation, we call on WHO to develop a framework for formal commitments by governments to improve the use of research evidence, and the World Health Assembly to endorse the adoption of such a framework. The WHO Framework Convention on Tobacco Control (FCTC) (http://www.who.int/fctc/en/) is a potential model for how such a framework might be developed. The FCTC is the first treaty negotiated under the auspices of the WHO and has been signed by 163 countries. Potential lessons to be learned from that experience include the importance of wide involvement of countries and stakeholders in developing the framework, basing it on evidence, mandating reporting to enable signatories to learn from one another’s experiences in implementing the FCTC, and assessment of progress based on summary analysis of those reports and impact evaluation.19

Key elements of a framework for ensuring well-informed decision-making about public programmes include commitments to:

- Adequate funding
- Development of capacity of both researchers to undertake evaluative research and to support policymakers’ needs for research, and of policymakers’ understanding and ability to use research appropriately
- Organisational structures or processes to support more systematic and transparent use of research evidence to inform decisions before starting new programmes
- Rigorous prospective impact evaluations of programmes, including
  - Planning evaluations in advance rather than as an afterthought
  - Ensuring clear objectives that are aligned to the programmes’ goals
  - Measurement of important outcomes as well as processes
  - Processes for prioritising which programmes are most in need of evaluation
- Effective mechanisms for managing conflicts of interests
- Involvement of the public, including civil society organisations, unions and professional organisations
- International collaboration to ensure that knowledge and learning are shared, to reduce unnecessary duplication of efforts, and to help develop capacity
- Avoidance of bureaucratic ineffective structures
- Monitoring and evaluation of the implementation and impacts of the legislation

Monitoring and evaluation of the proposed legislation is important, and in keeping with the spirit of this proposal, since there are important uncertainties about the impacts of legislating these commitments. The legislation could, for example, have undesirable impacts, such as inefficient bureaucratic processes, inappropriate inhibition or delay of promising programmes, political scapegoating, and litigation around the interpretation of evidence. Potential
downsides of the legislation that we are proposing need careful consideration and debate, as well as monitoring, to ensure that legislation is designed to maximise the chances of success and minimise the risks of undesirable consequences.

It is uncertain whether countries would want to negotiate these issues in an inter-governmental process, which could take time and resources. On the other hand, this is a global problem. Although specific details of any legislation will vary, there are lessons to be learned through collaborating on the development of a framework. It could also give teeth to resolutions such as the Mexico Statement. This is not to suggest that the framework we are proposing should be developed or imposed prescriptively by donors or international institutions. It is unlikely that would be helpful and it might be harmful. It should be developed consultatively with a strong focus on building within countries the full range of capacity needed to translate research into policies and practices that will improve people’s lives.

Commitments such as those listed above are needed to ensure politically and ethically responsible investments in programmes to improve health, social welfare, education and justice. They can help ensure that good intentions do more good than harm and that appropriate use of research evidence and impact evaluations become an expected element of decisions about public programmes rather than an optional add-on.
The government of Mexico passed legislation mandating the evaluation of social development policies in December 2003, which was signed by the president in January 2004. The legislation established a National Council for the Evaluation of Social Development Policy (CONEVAL) (www.coneval.gob.mx). A presidential decree published in August 2005 regulates the Council. CONEVAL was strengthened and given an expanded scope that now includes all government programmes in the National Development Plan for 2007 to 2012. Every ministry now has staff responsible for evaluation of its programmes with links to CONEVAL, there are regulations guiding these evaluations, and each ministry must budget for the evaluations of its programmes.

According to the legislation, CONEVAL is an independent public agency under the federal government. It is intended to have the autonomy and technical capacity to generate objective information to improve decision-making about social development. The legislation that led to the creation of CONEVAL stipulates that:

- evaluations may be undertaken by the Council or by independent agencies in institutions of higher education or scientific research, or non-profit organizations;
- programmes should be reviewed regularly to assess whether they are meeting the goals of the Social Development Policy so that they can be corrected, modified, extended, reoriented or suspended in whole or in part as needed;
- evaluations must include performance indicators for the management and quality of services, coverage and impact;
- federal agencies and programmes need to provide necessary information for evaluations;
- the Ministry of Finance and the House of Representatives through the National Auditor can recommend relevant indicators;
- evaluation results must be published in the Official Journal of the Federation and must be submitted to Congress; and
- the Council should aim to regulate and coordinate evaluation of policies and social development programs and establish guidelines and criteria for the definition, identification and measurement of poverty, ensuring the transparency, objectivity and technical rigor in this activity.

CONEVAL now has over 70 senior staff and a budget of close to $10 million US for 2009. In addition, each ministry is now mandated to commit a part of its budget to evaluation of its programmes. Up to now there has been no formal evaluation of the impacts of CONEVAL or the legislation that established it and it is perhaps too early to judge the pros and cons of this approach to mandating impact evaluations. There has been a clear increase in evaluation, but it is not known to what extent the evaluations are appropriately rigorous and meet the information needs of policymakers. There is some concern about the need for ministries to commit a proportion of their budget to evaluation without an increase in funds to pay for this and the impacts that might have on their programmes.

The legislation does not mandate the assessment of the likely impacts of programmes before they are launched and there is not a federal agency similar to CONEVAL responsible for preparing systematic reviews or policy briefs for this purpose.


5 Chalmers I. Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date evaluations. ANNALS of the American Academy of Political and Social Science 2003; 589:22-40. http://ann.sagepub.com/cgi/content/abstract/589/1/22

6 Macintyre S, Petticrew M. Good intentions and received wisdom are not enough. Journal of Epidemiology and Community Health 2000; 54:802-3. http://jech.bmj.com/cgi/content/full/54/11/802


