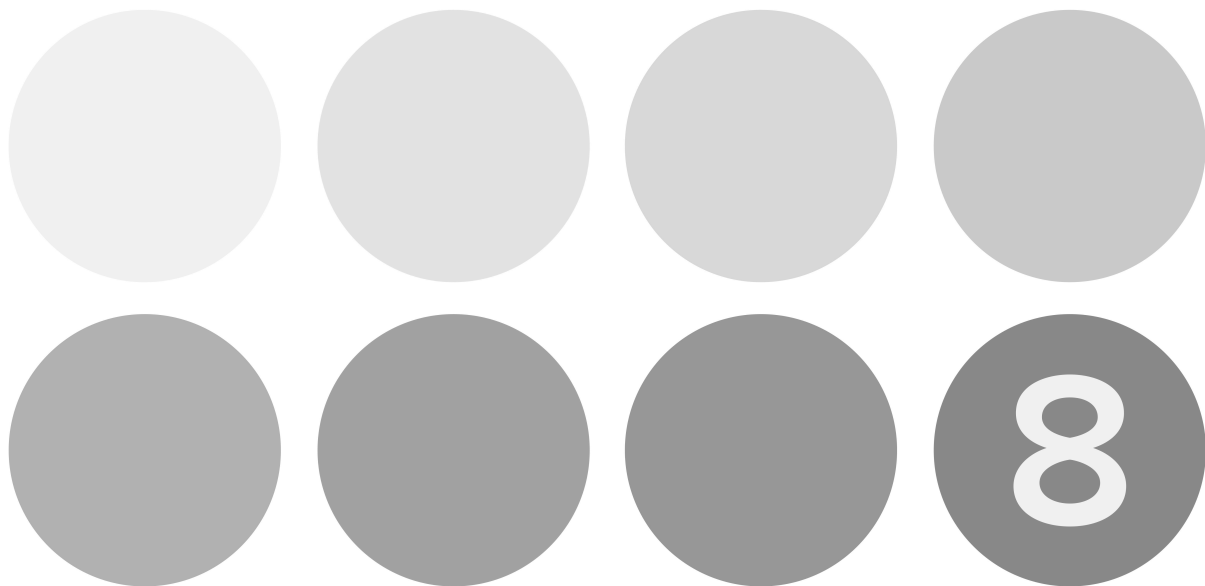


A Guide to Health Reform

Eight Practical Steps

Step Eight: Evaluating Impacts



**HARVARD
T.H. CHAN**

SCHOOL OF PUBLIC HEALTH

Department of Global Health
and Population

By Michael R. Reich,
Paola Abril Campos,
Anuska Kalita,
Anya Levy Guyer,
and Winnie Yip

Step Eight: Evaluating Impacts and Creating Sustainability for Your Health Reform

As we have emphasized throughout this Guide, the purpose of engaging in health reform is to improve the performance of the health system, especially on the performance problems selected as targets for action. Step Eight focuses on evaluating how the interventions selected influence health system outcomes. Do your interventions actually improve health system performance? While we have placed evaluation at the end of the Guide, monitoring and evaluation should in fact be considerations throughout each step. Similarly, *GHRR* presents evaluation as the important last stage in the policy cycle (Figure 2-1 above). The book does not include a separate chapter on evaluation, but Chapter 2 of *GHRR* does include a summary discussion of evaluation.

As mentioned in Step Five, a best practice is to try proposed interventions in pilot tests (or field trials) to demonstrate that they will have the intended impact. When this occurs, it is typically done after assessing the feasibility of implementing the interventions in a real-world context (Step Seven). However, implementation of health reform is sometimes pushed through without preliminary trials. Even in these cases regular monitoring is critical to implementation, and comprehensive evaluations should be conducted after implementation of the health reform plan.

Key actions in Step Eight:

<input checked="" type="checkbox"/>	Decide on your evaluation strategy early in the reform process, before starting implementation, especially whether you will use before-and-after comparison, or control groups, or region-by-region implementation at different times
<input checked="" type="checkbox"/>	Decide on how you will collect data needed for evaluation, including who will collect the data, how much it will cost, measures to assure reliability, and how to avoid collecting too little or too much data
<input checked="" type="checkbox"/>	Decide whether to use an external organization or an internal agency to perform the evaluation (after assessing the advantages and disadvantages of both approaches)
<input checked="" type="checkbox"/>	Answer five questions for planning the evaluation: <ol style="list-style-type: none"> 1. Why conduct an evaluation? 2. Who does the evaluation? 3. What do you evaluate? 4. Who are the main audiences and how to you communicate the evaluation? 5. Who will evaluate the evaluators?

Decide on your evaluation strategy

It is worthwhile to distinguish between *monitoring* and *evaluation* as concepts.

- *Monitoring* refers to data collection and analysis done while you are implementing your reform measures so that you can take corrective action during implementation to improve both the process and effects of your efforts. The intention of monitoring is to create relatively rapid and simple feedback loops on how you are doing, to provide information that assists in making immediate changes and in achieving specific targets. As noted in Step Seven, monitoring is typically conducted by the Implementation Team for its own use.

- *Evaluation*, on the other hand, occurs after significant implementation has been completed (either at the end of a project or at regular points in implementation, such as after one year or every couple of years) in order to determine whether the reform is achieving its intended goals. Some evaluations seek to determine whether the reform interventions are actually the drivers of changes in performance; others may be designed to capture data on unintended consequences of reform efforts.

Evaluation should be conducted by an entity with some (arm’s-length) objectivity. It could be performed by (or contracted to) an external agent, by an audit group within the government, or by a group situated at some distance from the implementers. Assigning evaluation to an internal group involved with implementation can create conflicts of interest that may undermine the legitimacy, rigor, and conclusions of the analysis and conclusions. The rest of Step Eight focuses on evaluation.

The Health Reform Team should begin considering the evaluation strategy at the start of the reform process, and continue thinking about evaluation throughout.

There are different types of evaluations to consider. For example, should you plan to do a before-and-after comparison evaluation? If so, the Health Reform Team needs to collect baseline data *before* starting implementation and set up administrative systems to collect (and preserve) relevant information during implementation. Another option is to include control groups (where the reform is not implemented, or more often is implemented after a certain period) in order to enable the evaluation to draw causal inferences between the interventions and the outcomes. If the latter option is preferred, then control groups need to be carefully selected and data collection begun at an appropriate time before implementation starts. If a reform is implemented region by region over time, this can create a kind of natural experiment to assess impacts (King et al., 2007). Broad strategic questions about evaluation designs need to be discussed, debated and decided on by the Health Reform Team early in the health reform process.

Decide on data collection for evaluation

GHRR (Chapter 2) presents several lessons and cautions about data collection for evaluation: First, data collection is not free, and better data typically cost more to collect. Second, the costs of data collection typically fall on the people doing the reporting. If the costs to them of gathering good data are too high, they will usually provide poor data. Third, it is possible to collect too much data, creating “data cemeteries” in which piles of “dead” data accumulate but are not analyzed.

For certain types of evaluations (such as those requiring household or facility surveys, for example) and as discussed in relation to the performance assessment in Step Three, the Health Reform Team may decide to hire an external organization to collect data and do the analysis. The extent of data collection required for evaluation depends on many factors, including the evaluation’s objectives, budget, and timetable, as we discuss next.

Plan the evaluation of your health reform

As the Health Reform Team plans to evaluate the health reform, the following five questions can be used to guide strategic and operational decisions. (They may seem very similar to questions asked in the previous steps—however, the answers will vary in each step.)

1. Why conduct an evaluation? Think about the technical, ethical and political reasons for evaluation. Evaluation serves important technical objectives by seeking to determine whether the interventions selected are having the intended consequences on key measures of health system performance (both intermediate and final outcomes). The lessons from evaluation can provide lessons about how to improve the reform's impacts on health system performance and can promote learning about the reform by many audiences.

Evaluation also can help assess core ethical concerns, including the transparency and accounting of using public resources as well as achieving important distributional and equity goals (for example, in improving maternal mortality of disadvantaged ethnic groups or of specific geographic regions). Are the intended beneficiaries actually receiving the services targeted at them, with the expected impacts on outcomes?

Evaluation can also serve political purposes, for instance, in seeking to create evidence for the effectiveness of the health reform in achieving an administration's promised improvements in the health system, with the potential for rebutting possible future efforts to roll back or even eliminate the policy changes introduced. In this sense, evaluation, by showing that the reform is having its intended impacts, can serve as a kind of insurance policy against potential political opposition.

2. Who does the evaluation? As noted above, an evaluation is usually not conducted by the organization responsible for implementation, but a decision still must be made about whether to use an evaluation group within government or outside government. This decision, along with the allocation of financial resources and budget, can affect the quality, timetable, and accountability of the evaluation. Using an external organization may require a competitive bidding process; it will usually require a contract and negotiation with the evaluation group, to assure that the evaluators have the necessary technical skills and capacity to complete the evaluation in the required time (which can be influenced by political factors, such as the end of an administration and upcoming elections). The selection of the evaluation group also has important implications for the legitimacy and influence of the evaluation and final report.
3. What do you evaluate? An evaluation can focus on specific outcomes as measurable objectives as well as various processes required for implementation. The decision about what you evaluate will determine the kinds of quantitative and qualitative methods that are used in assessing health reform interventions. The design of the health reform may include specific objectives as targets. The evaluation group may also decide to assess performance achievements according to benchmarks, using other similar entities (other states within a

country, or nearby countries for a national level evaluation). The evaluation can include an assessment of health system processes, using qualitative methods, as well as specific numeric targets, using quantitative methods. The decisions about what to evaluate should be clearly defined in the contract with the evaluation group, and will be shaped by the overall evaluation strategy (as discussed above, for example, a before-and-after evaluation versus an evaluation with control groups). The contract can include specific benchmarks on what and when the evaluation group decides to evaluate, to allow for sufficient discussion and negotiation between the evaluation group and the Health Reform Team.

4. Who are the main audiences for the evaluation and how do you communicate with them? Deciding on the audiences for your evaluation is critical to shaping your communication plan for the evaluation and its conclusions. The audiences could include the top political leaders for the country, national legislators, administrators for different parts of the health system, labor unions for the health sector, and the general public—as well as international experts, multilateral organizations, and donor agencies (depending on the country and the reform). Each audience could require different messages and communications strategies, ranging from social media to top-ranking scientific journals. For a major health reform evaluation, it may be appropriate and necessary to hire a professional communication strategy company to plan the messages and their effective delivery. This decision is best made early in the evaluation planning process. As noted above, evaluation has technical objectives as well as ethical and political objectives. Evaluation is more than a technical exercise, and the Health Reform Team needs to manage the process and perceptions of evaluation carefully to help assure success of the overall reform process.
5. Who will evaluate the evaluation? This question concerns the legitimacy and credibility of the evaluation. Three possible mechanisms to support positive perceptions of the evaluation are: to assign the evaluation to an institution with a strong international reputation for rigorous studies; to seek publication of the evaluation in a high-reputation international peer-reviewed scientific journal (such as *Lancet*, *Nature*, or *Science*); and to create a blue-ribbon advisory committee to oversee the evaluation process. These mechanisms, however, may not be sufficient, in a polarized political environment, to deter public and opposition criticism of an evaluation and its underlying reform.

Examples of evaluations

There are many textbooks, guides, articles, courses, and other publications on how to do evaluations of health policies. We will not attempt here to review the array of materials on policy or program evaluation, due to limited space and objectives of this Guide. [Appendix 8-1](#) highlights and provides references for a selection of health reform evaluations to illustrate what can be done, depending on your objectives, capacity, budget, and reform, as noted above. These examples can give your Health Reform Team some possible “models” of different kinds of evaluations.

Promote the sustainability and resilience of your reform

A positive evaluation of a health reform and its impacts, by itself, does not guarantee policy sustainability and political resilience. The continuity of your health reform into the future depends on multiple factors, including financial resources, political competition, and public support for the policy changes that have been implemented. There is a tendency among policy makers to assume that path dependency will occur, that positive feedback loops from key beneficiaries will sustain the reform against opposition (Pierson, 2000). But things do not always happen that way. The near demise of President Barack Obama’s health reform in the United States, which survived an attempted rollback by one single vote in the US Senate (from Republican John McCain), shows how precarious that expected “path dependency” can be (Scott and Kliff, 2017). And the elimination of *Seguro Popular* in Mexico, by the government of President Andrés Manuel López Obrador, demonstrates that “path dependency” can be reversed even after 16 years of implementation, if a political opponent gains power and control of government (Reich, 2022).

The key question for the Health Reform Team is: How can you build in political resilience against the low probability event that your opposition will come to power and seek to weaken or reverse or eliminate your reform? How can health reformers build legislative and judicial safeguards so that their changes will be sustained if the political enemy comes to power? How can health reformers create political support among key stakeholders, and popular understanding and support among the general public and beneficiaries?

In a real sense, the success of a health reform is measured not only by its policy impacts (its consequences for key health system performance indicators), but also by its political resilience (its ability to survive and continue when the political opposition comes to power).

References

- Bergkvist S, Wagstaff A, Katyal A, Singh P, Samarth AM, Rao MI. 2014. What a difference a state makes: health reform in Andhra Pradesh. Policy Research Working Paper 6883. Washington, DC: The World Bank Development Research Group. <https://doi.org/10.1596/1813-9450-6883>
- King G, Gakidou E, Imai K, et al. 2009. Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. *Lancet* 373: P1447-1454. [https://doi.org/10.1016/S0140-6736\(09\)60239-7](https://doi.org/10.1016/S0140-6736(09)60239-7)
- Nigenda G, Wirtz VJ, González-Robledo LM, Reich MR. 2015. Evaluating the impact of Mexico’s health reform: The case of *Seguro Popular*. *Health Systems & Reform* 3:217-228. <https://doi.org/10.1080/23288604.2015.1031336>
- Pierson P. 2000. Increasing returns, path dependence, and the study of politics. *American Political Science Review* 94(2): 251-267. <https://doi.org/10.2307/2586011>

Reich MR. 2020. Restructuring health reform, Mexican style. *Health Systems & Reform* 6(1): e1763114. <https://doi.org/10.1080/23288604.2020.1763114>

Roberts MJ, Hsiao W, Berman P, Reich M. 2004. *Getting Health Reform Right: A Guide to Improving Performance and Equity*. New York: Oxford University Press.

Scott D, Kliff S. 2017. Why Obamacare repeal failed—and why it could still come back. *Vox.com*; 31 July 2017. <https://www.vox.com/policy-and-politics/2017/7/31/16055960/why-obamacare-repeal-failed> (Accessed 12 June 2023)

Appendix 8-1: Examples of different models used to evaluate health reforms

King G, et al. 2009. *Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. Lancet* 373: P1447-1454. This high-quality technical evaluation was conducted 10 months after the start of implementation, to assess impacts on both intermediate and final performance objectives. The study found a reduction in catastrophic health expenditures, but not in medication spending, health outcomes, or health service utilization. This assessment contributed to the health reform's continuity across two subsequent administrations in Mexico, but not when a new opposition government came to power in 2018. [https://doi.org/10.1016/S0140-6736\(09\)60239-7](https://doi.org/10.1016/S0140-6736(09)60239-7)

Nigenda G, et al. 2015. *Evaluating the impact of Mexico's health reform: the case of Seguro Popular. Health Systems & Reform* 3:217-228. This article combines information from four external evaluations of Mexico's health reform (conducted by a team at the National Institute of Public Health) to do a process evaluation on the use of financial resources in purchasing health services. The article focuses on implementation processes related to the federal government's financial transfer mechanisms to the states, the purchase of medicines, and the contracting of health workers. The assessment found a number of significant challenges and identified various government responses to problems that have sought to improve performance, "with mixed results." <https://doi.org/10.1080/23288604.2015.1031336>

Blanchet NJ, et al. 2012. The effect of Ghana's National Health Insurance Scheme on health care utilization. *Ghana Medical Journal* 46(2): 76-84. This study examines the impact of Ghana's health reform, implemented in 2003, on utilization of health services, by using data from the Women's Health Study of Accra on medicines and health services. The assessment found that women with health insurance are "significantly more likely" to obtain prescriptions, visit clinics, and seek health care from formal providers when sick. In short, they have been access—an important intermediate objective for health reform. This academic study, however, was not conducted with the goal of a government-sponsored evaluation of the policy. PMID: [22942455](https://pubmed.ncbi.nlm.nih.gov/22942455/)

Bergkvist S, et al. 2014. *What a difference a state makes: Health reform in Andhra Pradesh. Washington, DC: The World Bank Development Research Group. Policy Research Working Paper 6883.* This World Bank study assesses Andhra Pradesh's Aarogyasri health insurance scheme, which provides coverage for around 900 high-cost procedures delivered in secondary and tertiary hospitals. The study found that patients still paid "quite large" out-of-pocket expenses during hospitalization, even for services covered by the scheme. A comparison with a neighboring state (Maharashtra), however, found that Andhra Pradesh showed better access (higher rates of inpatient and surgery admissions) and better financial risk protection (lower growth rates of certain costs). The report indicated a number of areas for improvement in the health reform, although this was not an official government evaluation. <https://doi.org/10.1596/1813-9450-6883>