A Results Framework Serves Both Program Design and Delivery Science¹,²

David R. Marsh,³* Juan Carlos Alegre,⁴ and Karen Z. Waltensperger⁵

¹Senior Child Survival Advisor, Save the Children, Westport, CT 06880; ²Monitoring and Evaluation, Save the Children, Washington, DC 20036; and ³Africa Regional Health Advisor, Save the Children, Pretoria 0127, South Africa

Abstract

Health programmers and researchers must collaborate despite different mandates and technical languages. A results framework is a simple model that both disciplines can use to understand complexity, clarify assumptions and hypotheses, design programs, and ask questions to inform action research. Typically, a health program’s results framework has 3 tiers and 6 boxes: a base of 4 health service intermediate results (access, quality, demand, and environment), which lead to a midlevel strategic objective (use of life-saving intervention), which leads to the goal (improved health). A situation analysis directly informs intervention selection; more difficult is selecting strategies to deliver the interventions, especially in settings of health system weakness. We propose menus and submenus of strategies to achieve each intermediate result, illustrate the use of the results framework in a program design and in clarifying research questions, and begin to propose a research agenda for “delivery scientists” responsible for recommending optimal investments to maximize use of interventions by those who need them most. J. Nutr. 138: 630–633, 2008.

The results framework

Programmers and researchers have different agendas, serve different stakeholders, and use different terminologies. Yet, to answer international health’s most vexing challenges, they must and do collaborate. Results frameworks are simple models that both disciplines can understand and use. The United States Agency for International Development adopted the results framework to strengthen the planning and evaluation of its projects (1,2). Despite a decade of growing use among the Agency and its partners, relevant literature is scant. Toffolon-Weiss, Bertrand, and Terrell (3) explain it and provide a case study of its use in an HIV/AIDS initiative in Central America. Others cite it as an example of a program pathway (4) or a model for developing large-scale intervention and assessment programs (5).

Results frameworks can serve a range of social programs, agencies, and levels. Here we restrict our attention to a “6-box” version commonly used by health programmers (Fig. 1) (6). Each box reflects a specific category of programmatic result, all of which contribute to the overall goal of improved health status, typically reduced mortality, morbidity, or fertility. The framework serves 2 critical programmatic functions. By describing the steps toward a program’s ultimate goal, it indicates the general programmatic direction to achieve that goal through stepwise results. Second, by outlining a hierarchy of results, it prompts managers to specify indicators to track progress toward achieving them.

The framework has 3 levels. At the top is the program goal: improved health status. Goals are usually expensive to measure, requiring large population-based surveys carried out at infrequent intervals by large organizations. Beneath the goal, in the middle tier of the framework, is the strategic objective (SO). Most health programs define their SO as increased use of key services or behaviors. Effective programs select “evidence-based interventions” (7,8), that is, services or behaviors known to have an impact on health status. The SO is the most ambitious result that programs can reasonably effect and for which implementing agencies are willing to be held accountable.

The bottom tier of the framework has 4 intermediate results (IR), which are essential steps toward achieving the SO. They are: IR-1, access to and availability of health services increased; IR-2, quality of health services improved; IR-3, demand for health services or behaviors increased; and IR-4, social and policy environment enabled. Access and availability refer to the likelihood that a person and a service will actually meet (through the person going to the service, the service going to the person, or meeting midway) and that the necessary supplies and services will be available. Access refers not only to geography but also to socioeconomic factors that can serve as barriers to care. The concept of quality is complex and refers not only to the technical quality of the services but also to clients’ perceptions of quality and acceptability—not always the same as technical quality. Nonetheless, good perceived quality facilitates use. The first 2 IR...
The results framework melds 2 disciplines: discovery science and delivery science. That is, the association between the SO and the goal rests on the efficacy studies of “discovery scientists” to assess the impact of an intervention on a health outcome. On the other hand, the association between the IR and the SO rests on common-sense assumptions that “delivery scientists” need to test and quantify. Below we propose 3 sets of questions arising from the framework: 1) from the framework itself, 2) from the framework applied to a specific program, and 3) from the framework applied to a research initiative.

**The framework itself.** Two general questions arise directly from the generic framework, regardless of the specific program setting. First, what factors might modify the expected association between supply and demand versus use or between use versus health? Victora and coworkers have proposed groups of such effect modifiers: 1) behavioral (provider, beneficiary, others); 2) biological (antagonism, synergy, curvilinear dose-response, limited scope for improvement, missing cofactor); and 3) contextual (socioeconomic, demographic, environmental, baseline health, health system, other programs) (9,10).

The second general question relates to optimizing the mix of strategic investments to strengthen supply, demand, and enabling environment. An exclusive investment in only 1 strategy may not increase use. For example, investing only in IR-1 might yield highly available interventions, poorly delivered, with incomplete demand, whereas investments limited to IR-2 could yield incompletely available interventions, well delivered (where available), with incomplete demand. Programmers must address gaps in IR, although all may not need attention. So how should effort be apportioned, and what is the evidence?

**Program application.** An example of a health program designed for an extremely impoverished setting, Southern Sudan, highlights other questions. A situation analysis (qualitative and quantitative) characterized the environment, including the health system; the causes of illness and death; and the levels of indicators of use, availability, demand, and so on—and the gaps. This analysis, in turn, directly informed the selection of lifesaving interventions (7,8). More difficult was the question of how to deliver these interventions, especially because the health system was almost nonexistent.

We drafted menus of strategies to achieve each IR (Table 1), with substrategies for each (not shown; available on request). For example, improving access through facility-based services includes renovating existing facilities; constructing new facilities; expanding service packages; integrating new interventions into existing packages; staffing up to or beyond plan; increasing service hours per day or days per week; implementing insurance, loan, or waiver schemes; using local language; respecting cultural norms; installing radio or telephone; using, controlling, and maintaining an ambulance; engaging community-based providers; and eliminating other barriers, such as removing mines or unexploded ordinance and repairing roads or bridges.

**A framework for program design and research.**

Sudan is not unique. Other examples range from complex humanitarians’ emergencies to sparse, remote, nomadic populations.
TABLE 1 Strategies to achieve intermediate results

<table>
<thead>
<tr>
<th>Intermediate result</th>
<th>Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR1: Access</td>
<td>Home-based case management</td>
</tr>
<tr>
<td></td>
<td>Community-based services: informal providers</td>
</tr>
<tr>
<td></td>
<td>Community-based services: formal providers</td>
</tr>
<tr>
<td></td>
<td>Outreach services</td>
</tr>
<tr>
<td></td>
<td>Facility-based services</td>
</tr>
<tr>
<td></td>
<td>Nontraditional service delivery points</td>
</tr>
<tr>
<td></td>
<td>Intersectoral service delivery points</td>
</tr>
<tr>
<td>IR2: Quality</td>
<td>Improve provider capacity</td>
</tr>
<tr>
<td></td>
<td>Strengthen service delivery point management</td>
</tr>
<tr>
<td></td>
<td>Link community with health system</td>
</tr>
<tr>
<td></td>
<td>Develop culture of quality</td>
</tr>
<tr>
<td></td>
<td>Ensure acceptability and respect for cultural norms</td>
</tr>
<tr>
<td>IR3: Demand</td>
<td>Community mobilization</td>
</tr>
<tr>
<td></td>
<td>Behavior change communication</td>
</tr>
<tr>
<td></td>
<td>Branded approach</td>
</tr>
<tr>
<td>IR4: Enabling environment</td>
<td>Social transformation</td>
</tr>
<tr>
<td></td>
<td>Advocacy</td>
</tr>
<tr>
<td></td>
<td>Policy-informing evidence</td>
</tr>
<tr>
<td></td>
<td>Standards</td>
</tr>
<tr>
<td></td>
<td>Partnering</td>
</tr>
<tr>
<td></td>
<td>Community capacity</td>
</tr>
<tr>
<td></td>
<td>Institutional/organizational development</td>
</tr>
<tr>
<td></td>
<td>Leveraging resources</td>
</tr>
</tbody>
</table>

In most settings, working through existing systems is a sensible starting point, at least to reach some of the population. But most countries, even those with adequate systems, have geographic areas or populations beyond their system’s reach. What exactly are the options for increasing access, availability, quality, etc. to deliver interventions to those outside the existing system? Examining the status of the “fifth quintile” (11) is an increasingly common way to quantify this inequity. Reaching the bottom quintile (or tertile or decile) requires fresh thinking and local adaptation.

A catalogue of major and minor delivery strategy options is helpful, but it forces additional questions: How complete is the list? Which strategies are best for a given epidemiological setting (12)? Ecologic setting? Health systems setting? Cultural setting? Socioeconomic setting? Which strategies are best for given interventions or packages of interventions? Which strategies best serve equity? At what cost? At what cost/benefit?

Refining research. The generic results framework is a useful point of departure for both programmers and researchers. We recently began examining IR-4, ultimately to characterize the interaction between increasing the use of health interventions and increasing community capacity. To stimulate thinking, we placed 2 very simplified frameworks side by side: a results framework for health outcomes and a social change model (13) placed 2 very simplified frameworks side by side: a results framework for health outcomes and a social change model (13). The following questions arose relating to both processes and the outcomes of community capacity. What activities result in the most community capacity? How much do community capacity activities influence individual health outcomes? How do community capacity activities modify the effect of health activities on achieving individual outcomes? How does community capacity itself modify the effect of health program activities on achieving individual health outcomes? Additional questions, although not informed by Figure 2, derived from the foregoing

Comment
Development is complex, even when the interventions are “silver bullets” such as vitamin A supplementation or measles vaccination. A results framework, admittedly (and fortunately) a gross oversimplification of reality, brings advantages. It poses key questions for programmers and researchers by an orderly arrangement of critical health system issues and even by what it omits (Fig. 3)! Thus, a thoughtful consideration of the framework can inform a response to Scheirer, who argued that the performance-monitoring literature, including that related to results frameworks, failed to assess relations among performance measures, i.e., among IR or between inputs and IR (14). Similarly it can inform a response to Bollen, Paxton, and Morishima, who urged USAID and others to carefully characterize confounders before drawing conclusions from the more commonly measured inputs and results (15).

In addition, the results framework holds researchers accountable for all 4 IR, including measuring how changes in the enabling environment influence other results. This type of inquiry may need qualitative methods or the development, testing, and validation of new quantitative indicators and methods. Also, the results framework can be adapted to permit examining interactions between health programming and other or external phenomena. Finally, if it is more widely used, the results framework could bring some standardization to terminology.

The global health community has mature, productive, respected research institutions and teams to develop new interventions, sometimes called “discovery science.” On the other hand, the discipline to inform how to bring these discoveries to families who need them, sometimes called “delivery science,” is far less advanced. Too few people in developing countries currently receive the evidence-based health interventions that could save their lives (16). Indeed, at the 2005 World Health Assembly,
Bill Gates prioritized the need for “more thinking and funding to delivering interventions—not just discovering them” (17). A results framework can convene delivery science partners (programmers and researchers) as well as advocates, policy makers, and donors around a simple model that asks tough questions. The answers to those questions may bring the Millennium Development Goals within reach.

Acknowledgments
The authors recognize the encouragement and insights from David Oot, Mary Beth Powers, Eric Swedberg, Eric Starbuck, and Kathryn Bolles at Save the Children (U.S.A.).

Literature Cited