

99587

# COST-EFFECTIVENESS ANALYSIS OF RESULTS-BASED FINANCING PROGRAMS: A TOOLKIT

DISCUSSION PAPER

MAY 2015

*Donald Shepard*

*Wu Zeng*

*Ha Thi Hong Nguyen*



**WORLD BANK GROUP**

Health, Nutrition & Population



# **COST-EFFECTIVENESS ANALYSIS OF RESULTS-BASED FINANCING PROGRAMS**

## ***A Toolkit***

Donald Shepard, Wu Zeng, Ha Thi Hong Nguyen

**May 2015**

## HEALTH, NUTRITION AND POPULATION (HNP) DISCUSSION PAPER

This series is produced by the Health, Nutrition, and Population Global Practice. The papers in this series aim to provide a vehicle for publishing preliminary results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character.

For information regarding the HNP Discussion Paper Series, please contact Martin Lutalo at [mlutalo@worldbank.org](mailto:mlutalo@worldbank.org) or Erika Yanick at [eyanick@worldbank.org](mailto:eyanick@worldbank.org).

© 2015 The International Bank for Reconstruction and Development / The World Bank  
1818 H Street, NW Washington, DC 20433

All rights reserved.

# Health, Nutrition and Population (HNP) Discussion Paper

## Cost-Effectiveness Analysis of Results-based Financing Programs

### *A Guidance Note*

Donald Shepard<sup>a</sup> Wu Zeng<sup>b</sup> Ha Thi Hong Nguyen<sup>c</sup>

<sup>a</sup> Schneider Institutes for Health Policy, Heller School, Brandeis University, Boston, USA

<sup>b</sup> Schneider Institutes for Health Policy, Heller School, Brandeis University, Boston, USA

<sup>c</sup> Health, Nutrition, and Population Global Practice, The World Bank Group, Washington DC, USA

This work was financed by the Health Results Innovation Trust Fund, funded by the governments of Norway and the United Kingdom and managed by the World Bank.

**Abstract:** Results-Based Financing (RBF), which rewards providers, users, or administrators of services upon achieving a set of verified results, has been gaining attraction in global health as a prominent approach to gain value for money. With a large number of countries adopting RBF in the recent years, evidence starts to emerge which points to the effectiveness of RBF in improving coverage and quality of important services, such as maternal and child health and reproductive health. A question remains, however, if RBF is more cost-effective than alternative interventions which aim to improve similar outcomes. Given the limited resources that countries have for health, the answer to this question is particularly important as it guides policy makers on which programs to invest to maximize health benefits for the population. A second question is how the cost-effectiveness varies across different settings and RBF program features.

The current toolkit aims to support country programs to assess the cost-effectiveness of RBF interventions and to facilitate cross-country comparisons of RBF programs. The toolkit is specifically tailored to supply-side RBF but its general principles apply to most health systems interventions directed at the health-related Millennium Development Goals (MDGs)— 4 and 5. The development of the toolkit was based on actual experience of conducting a costeffectiveness analysis of Zambia's RBF program and an extensive review of RBF programs features across the Health Results Innovation Trust Fund (HRITF) portfolio. Given that RBF programs in the HRITF portfolio are typically complex health system interventions, the toolkit recommends a practical approach of adopting a program implementer's perspective and presents different options for cost-effectiveness analysis (for selected key indicators or for an entire package of services). It also provides guidance on incorporating quality of care, which is strongly emphasized across many RBF programs.

**Keywords:** Results-Based Financing, cost-effectiveness analysis, health MDG

**Disclaimer:** The findings, interpretations and conclusions expressed in the paper are entirely those of the authors, and do not represent the views of the World Bank, its Executive Directors, or the countries they represent.

**Correspondence details:** Donald Shepard, Schneider Institutes for Health Policy, Heller School, Brandeis University, Boston, USA. Tel: +1-781 736-3975; Fax: +1-888 429-2672; Email: shepard@brandeis.edu; <http://www.brandeis.edu/~shepard>

## TABLE OF CONTENTS

ACRONYMS .....	VI
ACKNOWLEDGEMENTS.....	VII
INTRODUCTION .....	1
CHAPTER 1. A GENERAL FRAMEWORK FOR COST-EFFECTIVENESS ANALYSIS .....	3
CHAPTER 2. COST ANALYSIS.....	5
<i>Key concepts and approaches in costing.....</i>	5
<i>Conduct cost analysis .....</i>	7
CHAPTER 3. ASSESSING PROGRAM EFFECTIVENESS .....	14
<i>Estimating RBF program impact on service utilization and quality of care .....</i>	14
<i>Selecting indicators for cost-effectiveness analysis.....</i>	16
<i>Converting coverage of services to lives saved .....</i>	17
<i>Include quality of care and other potential system effects in the effectiveness estimation.....</i>	31
<i>Convert lives saved to Quality Adjust Life Years (QALYs) or Disability Adjusted Lived Years (DALYs)</i> .....	34
CHAPTER 4. PUTTING IT TOGETHER: GENERATING ICERS.....	39
<i>Sensitivity analysis in CEA.....</i>	41
CHAPTER 5. DISCUSSION AND INTERPRETATION .....	44
<i>Interpretation of impact evaluation and CEA results.....</i>	44
<i>Limitations of the toolkit.....</i>	44
<i>Strengths of the toolkit .....</i>	45
<i>Resources for conducting a CEA of a RBF program.....</i>	45
<i>Implications and conclusions.....</i>	45
APPENDIXES .....	47
<i>Appendix 1. Delphi questionnaire for evaluating the impact of quality of care.....</i>	47
<i>Appendix 2. Delphi questionnaire for quantifying relative importance of generic vs service-specific quality indicators .....</i>	49
<i>Appendix 3. Delphi questionnaire for quantifying relative importance for components within the service-specific quality indicator .....</i>	51
REFERENCES .....	57

### List of tables

TABLE 1: EFFECTIVENESS ESTIMATES FOR CEA AND SUGGESTED DATA SOURCES .....	15
TABLE 2. COMMON INCENTIVIZED INDICATORS IN RBF PROGRAMS .....	16
TABLE 3. EXAMPLE OF IMPACT EVALUATION .....	20
TABLE 4. COVERAGE OF INSTITUTIONAL DELIVERY UNDER TWO SCENARIOS .....	20
TABLE 5. COVERAGE USING LINEAR INTERPOLATION .....	20
TABLE 6. COVERAGE USING LOGISTIC INTERPOLATION .....	20
TABLE 7. KEY SERVICES INCLUDED IN LIST.....	23
TABLE 8. INTERVENTIONS THAT ARE ASSOCIATED WITH INSTITUTIONAL DELIVERY AND THEIR IMPACT ON CAUSES OF MATERNAL DEATH.....	23
TABLE 9. NUMBER OF LIVES SAVED OVER THE THREE YEARS.....	25
TABLE 10. INSTITUTIONAL DELIVERY COVERAGE FROM 2012-2014.....	25
TABLE 11. SUMMARY OF LIVES SAVED FROM IMPROVING INSTITUTIONAL DELIVERY .....	27
TABLE 12. COMPONENTS OF GENERAL AND SERVICE-SPECIFIC QUALITY INDICATORS.....	33
TABLE 13. COMPONENTS OF GENERAL AND SERVICE-SPECIFIC QUALITY INDICATORS.....	36
TABLE 14: SELECTED TYPES OF SENSITIVITY TESTING USEFUL FOR COST-EFFECTIVENESS ANALYSIS.....	42

### List of figures

FIGURE 1. SCHEMA FOR COST-EFFECTIVENESS ANALYSIS OF HEALTH PROGRAMS .....	3
FIGURE 2: APPROACHES FOR MEASURING THE COST OF RBF INTERVENTIONS.....	6
FIGURE 3. SCHEMA OF THE LIST TOOL.....	18

FIGURE 4. LIST TOOL INTERFACE IN SPECTRUM.....	18
FIGURE 5. TABS FOR CHANGING POPULATION PROJECTION.....	21
FIGURE 6. PARAMETERS FOR BASELINE CHILD HEALTH STATUS.....	22
FIGURE 7. PARAMETERS FOR MORTALITY RATES.....	22
FIGURE 8. PARAMETERS FOR RESULTS ON NEONATAL, CHILD, AND MATERNAL DEATHS.....	24
FIGURE 9. UTILIZATION OF INSTITUTIONAL DELIVERY IN LIST.....	26
FIGURE 10. UTILIZATION OF INSTITUTIONAL DELIVERY UNDER ALTERNATIVE SCENARIO.....	26
FIGURE 11. DEATHS IN CHILDREN UNDER 5 FROM THE SIMULATION OF THE TWO SCENARIOS.....	26
FIGURE 12. MATERNAL DEATHS FROM THE SIMULATION OF THE TWO SCENARIOS.....	27
FIGURE 13. COVERAGE FOR PREGNANCY-RELATED SERVICES.....	28
FIGURE 14. COVERAGE OF VACCINATIONS IN LIST.....	29
FIGURE 15. UTILIZATION OF FAMILY PLANNING.....	30
FIGURE 16. EFFECTIVENESS OF FAMILY PLANNING INTERVENTIONS.....	30
FIGURE 17. POTENTIAL RELATIONSHIP BETWEEN QUALITY OF CARE AND ITS IMPACT*.....	32

### List of boxes

Box 1. EXAMPLE.....	8
Box 2. MEASURING INFLATION WITH GDP DEFLATOR.....	10
Box 3. DISCOUNTING EXAMPLES.....	11
Box 4. THE DELPHI TECHNIQUE.....	12
Box 5. THE QUADRATIC FUNCTION.....	31
Box 6. QUALITY ADJUSTMENT IMPACT.....	33
Box 7. USING DELPHI SURVEY TO ESTIMATE QUALITY INDEX.....	34
Box 8. EXAMPLE OF QALYS SAVED FOR PREGNANT WOMEN.....	37
Box 9. EXAMPLE OF CALCULATING QALYS PER LIFE SAVED FOR CHILDREN UNDER 1 MONTH AND CHILDREN OF 1-60 MONTHS.....	37
Box 10. EXAMPLE OF CALCULATING ICER.....	39
Box 12. COST-EFFECTIVENESS ANALYSIS OF A HYPOTHETICAL PACKAGE OF RBF SERVICES, PART 2.....	40
Box 11. COST-EFFECTIVENESS ANALYSIS OF A HYPOTHETICAL PACKAGE OF RBF SERVICES, PART 1.....	40
Box 13. COST-EFFECTIVENESS ANALYSIS OF A HYPOTHETICAL PACKAGE OF RBF SERVICES, PART 3.....	41

## ACRONYMS

AIM	AIDS Impact Model
ANC	Antenatal Care
CEA	Cost-effectiveness Analysis
DALYs	Disability Adjusted Life Years
GDP	Gross Domestic Product
GNI	Gross National Income
HIV	Human Immunodeficiency Virus
HMIS	Health Management Information System
HRITF	Health Results Innovation Trust Fund
ICER	Incremental Cost-effectiveness Ratio
IPT	Intermittent Preventive Treatment
IPTp	IPT in Pregnancy
MCH	Maternal and Child Health
MDG	Millennium Development Goal
MoH	Ministry of Health
OR	Odds Ratio
PBF	Performance-based Financing
PMTCT	Prevention of Mother to Child Transmission
PNC	Postnatal Care
PVAF	Present Value of Annuity Factor
QALYs	Quality Adjusted Life Years
QoL	Quality of Life
RBF	Results-based Financing
RR	Relative Risk
SBA	Skilled Birth Attendant
UNFPA	The United Nations Population Fund
UNICEF	The United Nations Children's Fund
WHO	World Health Organization
YLD	Years of Lost Due to Disability
YLL	Years of Life Lost



## ACKNOWLEDGEMENTS

The Toolkit for Cost-Effectiveness Analysis of Results-Based Financing Programs was prepared by a team consisting of Ha Thi Hong Nguyen (Senior Economist, Task Team Leader, The World Bank), Donald Shepard (Professor, Schneider Institutes for Health Policy, Heller School for Social Policy and Management, Brandeis University), and Wu Zeng (Assistant Research Professor, Heller School for Social Policy and Management, Brandeis University and Senior Health Economist, Futures Group International). The work was financed by the Health Results Innovation Trust Fund, funded by the governments of Norway and United Kingdom and managed by the World Bank. Some material from the toolkit was drawn from an earlier draft developed by Logan Brenzel (“Cost and Cost Effectiveness Analysis for Impact Evaluation”) and we fully acknowledge that.

The development of the Toolkit significantly benefited from our experience conducting cost-effectiveness analysis of the Results-Based Financing Program in Zambia. We are grateful to the Zambia Task Team for facilitating this process and providing valuable inputs, in particular to Collins Chansa (Health Specialist), Ashis Kumar Das (Health Specialist), Jumana Qamruddin (Senior Health Specialist), and Jed Friedman (Senior Economist). We thank Andrew Chisela and Masauso Undi for helping with data collection in Zambia and thank Clare Hurley for helping with editing and formatting the Toolkit. We are also thankful for the active participation of experts from Ministry of Health and development partners in the one day workshop for Delphi consultation in Lusaka, November, 2014. The team appreciates the administrative and management support of Charity Mbangweta (Program Assistant, Lusaka) and Aissa Socorro (Program Assistant, Washington DC).

The team gratefully acknowledges inputs from internal peer reviewers, Emre Ozaltin (Senior Economist) and Edit Velenyi (Economist). Finally, we appreciate the valuable support and guidance from Nicole Klingen (Practice Manager) and Tim Evans (Senior Director, Health, Nutrition, and Population Global Practice, The World Bank).

The authors are grateful to the World Bank for publishing this report as an HNP Discussion Paper.

## INTRODUCTION

1. Over the last decade, the global health community has used Results-Based Financing (RBF) as a means to strengthen health services. In a broad sense, RBF is an incentive approach to health systems strengthening that provides financial and non-financial rewards to providers, users, or administrators of services upon achieving a set of verified results. As an approach, RBF has been receiving increasing attention in health as well as other sectors for its promise to gain value for money.

2. With the support from the governments of Norway and the United Kingdom, the World Bank is managing the Health Results Innovation Trust Fund (HRITF) to pilot RBF programs in more than thirty countries. The focus of these RBF programs is to improve maternal and child health (MCH) health outcomes, as well as other conditions of high disease burden in many developing countries such as tuberculosis, human immunodeficiency virus (HIV) infection, and malaria. A central element of the HRITF's learning agenda is to document the extent to which RBF policies are effective, are operationally feasible, and in what circumstances. As such, rigorous evaluation is essential for generating new knowledge that can inform governments and partners to effectively design and use RBF mechanisms. The eventual learning objective is not only to assess the impact of the RBF intervention(s) in individual countries, but also to compare these impacts across other countries with similar interventions, or to compare the cost-effectiveness of RBF versus alternative interventions even within the same country.

3. Given the limited resources that developing countries have for health, policy makers often need to make decisions on which programs to invest to maximize health benefits for the population. Cost-effectiveness analysis (CEA) is such a tool to provide evidence to inform policy makers associated costs and potential benefits for an intervention, and examine both economic and health consequences of alternatives, for evidence-based decision making.

4. The current toolkit aims to support country programs to assess the cost-effectiveness of RBF interventions and to facilitate cross-country comparisons of RBF programs. This is a companion tool to the Impact Evaluation Toolkit developed under HRITF's work program, the first version of which was launched in 2012 (World Bank, 2015). The purpose of the toolkit is to enable impact evaluation teams to conduct a CEA that compares the costs and health effects of a RBF program themselves when the impact evaluation results become available. Hence its primary audience is impact evaluation teams who have a high level of training in economics or public health and are already familiar with the concept of a CEA. Additionally, the toolkit also aims to contribute to global knowledge of cost-effectiveness of a RBF program if the RBF program is implemented, and the same methodology could be applied to estimate CE ratio to inform whether or not to invest in RBF if a program has not been implemented. The toolkit is specifically tailored to the most common form of RBF in the HRITF portfolio – supply-side RBF, or Performance-Based Financing (PBF). Nevertheless, the toolkit's general principles apply to most health systems interventions directed at the health-related Millennium Development Goals (MDGs)—Goals 4 and 5. RBF has similar goals and kinds of inputs (financial and managerial) with other innovations to improve maternal-child services, such as reproductive health vouchers, subsidized insurance for pregnant women, conditional cash transfers, mobile health and e-health applications, and incentives for community health workers. The toolkit was developed incorporating experience conducting a CEA for the Zambia RBF program.

5. It is important to note that this toolkit is not designed to replace valuable guidelines for the health sector as a whole (Hutubessy, Chisholm, & Edejer, 2003; World Health Organization, 2003), nor disease-specific interventions (Jamison et al., 2006). Rather, the aim here is to be concise and practical, providing hands-on experience and measurement instruments for teams to apply. As such, the theoretical discussions will be kept to a minimum. One novel aspect of this toolkit is that it explicitly incorporates quality of care and other system strengthening efforts because they are strongly emphasized in the HRITF's RBF programs. Acknowledging that RBF programs in the HRITF portfolio are typically complex health system interventions, for which a simple and straightforward CEA may not be feasible, we present options for CEA

with different levels of complexity: CEA for selected key MCH indicators and CEA for an entire package of services with or without incorporating quality of care.

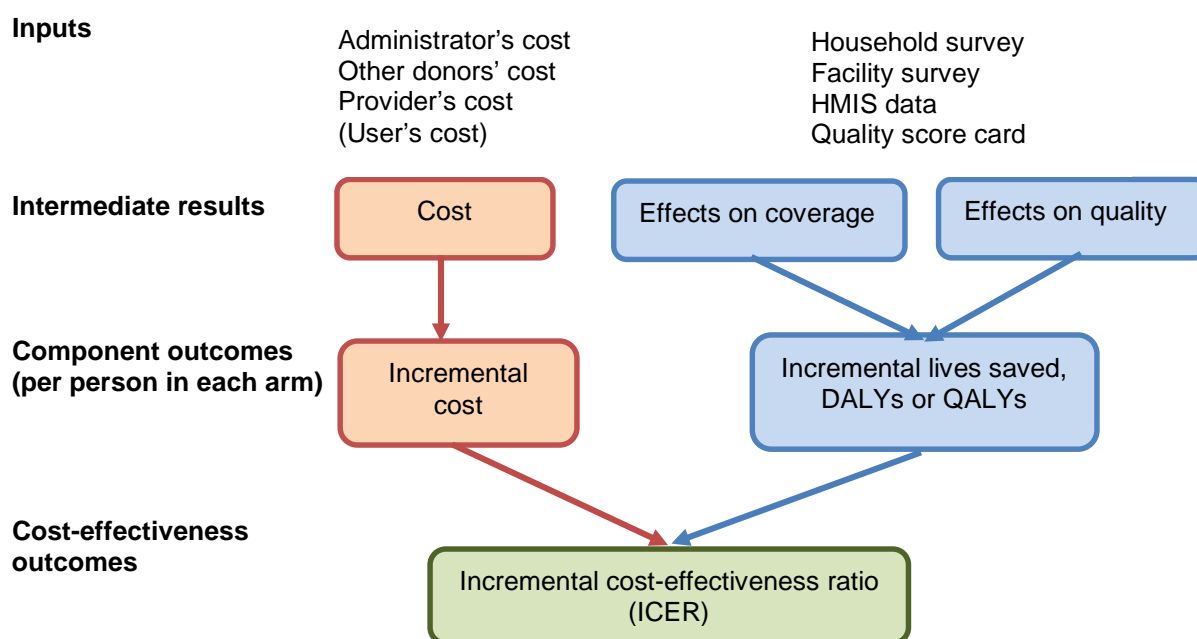
6. The production of a toolkit to facilitate CEA studies of RBF programs is inspired by the dearth of empirical evidence on RBF's value for money. A review of studies on pay-for-performance (P4P), a typical form of RBF, noted: "Only nine of [the reviewed] studies considered costs beyond the financial incentives, and only two of these studies included costs of developing and setting up the P4P scheme. The costs to providers of participating in the scheme and eventual costs to patients have also been omitted from previous analyses. Outcomes, if measured, are typically restricted to changes in incentivized services, with only four studies reporting effects on outcomes and one study reporting effects on non-incentivized services. No study to date reporting on the costs or cost-effectiveness of P4P has been published from low and lower middle income countries" (Borghi, Little, Binyaruka, Patouillard, & Kuwawenaruwa, 2015).

7. The section below provides an overall framework for incremental CEA. It is followed by detailed discussions of cost analysis and assessment of effectiveness in RBF programs, which are then combined for producing various CEA ratios. The discussions of concepts and methods are illustrated with an example from a hypothetical country. The toolkit ends with a discussion of strengths and limitations of the recommended approaches and some practical notes for researchers and evaluators.

## CHAPTER 1. A GENERAL FRAMEWORK FOR COST-EFFECTIVENESS ANALYSIS

8. Conducting a CEA requires data on a variety of inputs on costs and effectiveness. Typically, the needed data are assembled from multiple diverse sources. In the case of RBF programs, for example, the costs often include RBF program costs (i.e. incentives, costs of monitoring and evaluation), costs of consumables (i.e. costs of drugs utilized due to RBF programs), and other costs above the facility level. Effectiveness includes information on improvement of both quality and quantity of health services from the impact evaluation, which should then be translated to the number of lives saved, Quality Adjusted Life Years (QALYs) or Disability Adjusted Life Years (DALYs) for a CEA. A health CEA generally compares an intervention approach (such as RBF or another innovation) with the existing, or control approach. Figure 1 shows a general schema for a CEA of health programs in general, which can apply specifically to a CEA in RBF programs.

**Figure 1. Schema for cost-effectiveness analysis of health programs**



9. As shown in Figure 1, the final output of a CEA is typically an incremental cost-effectiveness ratio (ICER). Here we focus on ICER rather than cost-effectiveness ratio (CER), as a RBF program is built on existing health system and the complexity of tracking health system costs if CER is conducted. We also assume that the RBF program is an alternative of the business as usual model. CEA is a tool that can be used to identify which interventions achieve the greatest level of health impact per unit of investment. In CEA, the cost of a health intervention is divided by an estimate of health outcomes. Because we are interested to know whether it is better to invest an additional dollar toward one intervention compared to another, we are comparing costs and outcomes on the margin. To explain this more formally, the ICER of interventions, defined as the change in cost (with and without the intervention, or between two alternative interventions), divided by the change in effectiveness (with versus without the intervention or between two alternative interventions) is calculated for a standard population as:

$$ICER = \text{Difference in cost} / \text{Difference in outcomes}$$

10. The calculations are most clear if the standard population is one person. The difference means the value for a person in the intervention group compared to one in the control group. Thus, the equation becomes:

$$ICER = \text{Difference in cost per capita} / \text{Difference in outcomes per capita}$$

11. Cost-effectiveness ratios of alternative RBF programs can be compared to each other and to the counterfactual of a situation without the RBF intervention (control). In the context of RBF, CEA can help us understand how much it costs more for an additional unit of effect (or health outcome) averted or gained in relation to different alternatives using similar amounts of resources for implementation. It provides evidence for policy makers on which program to invest given budget constraints. If one uses DALYs averted or QALYs gained as outcome measures, GDP/capita is used as threshold to determine if the interventions under evaluation is cost-effective or not: highly cost-effective (ICER less than GDP per capita); cost-effective (ICER between one and three times GDP per capita); and not cost-effective (ICER more than three times GDP per capita) (World Health Organization, 2015), (World Health Organization, 2011).

12. The next three chapters explain the three key steps and associated concepts laid out in Figure 1: (1) how to conduct cost analysis to obtain incremental cost; (2) how to translate impact evaluation results to incremental lives saved, QALYs or DALYs; and (3) how to compute the ICER and interpret the results.

## CHAPTER 2. COST ANALYSIS

13. The first step in this manual is to apply cost analysis to estimate the costs of RBF interventions. The principles and approaches that we use reflect general practice for estimating costs of health services (World Health Organization, 2015), applied directly to RBF interventions.

14. While cost analysis can be used as a building block for estimating cost-effectiveness or performing other types of economic analyses, cost analysis in its own right can provide insight as to:

- What are the least and most costly RBF strategies?
- What components of cost account for the greatest amount of RBF resources?
- What is the cost per facility or per person served?
- What is the potential cost of scaling-up the RBF strategy?
- What is the difference in the cost of establishing the RBF mechanism compared to maintaining the program over the long-term?
- How do alternative versions of an RBF program compare in costs?
- How does the cost of a specific RBF program vary with the site in which it is implemented?

### KEY CONCEPTS AND APPROACHES IN COSTING

15. This subsection introduces key concepts commonly used in costing, as well as the main approaches for cost analysis. It then presents the method we adopt in the Toolkit for CEA of RBF programs and provides the justification for such decisions.

#### Incremental cost

16. Incremental costs are the difference in costs between an RBF program and usual care, or between two alternative versions of RBF. Incremental costs take one approach as the reference or comparator, against which the RBF program is compared. Typically, the reference is the status quo (no RBF) and the new program is the addition of the RBF. Thus, the incremental costs measure the additional costs of adding an RBF program onto the existing health system. The analysis can be done at whatever level the program is implemented—one or several facilities, a province, or an entire country.

17. Costs could be categorized as direct vs indirect costs, and fixed vs. variable costs. In the health setting, direct costs are those that result from health services, lab tests and drug therapy, while indirect costs are those forgone due to an illness. Fixed costs are those that do not vary as outputs (inpatients and outpatients) change while variable costs change as the outputs vary. In calculating incremental costs, as the fixed costs does not change with the change of the volume of health services, they can be omitted from the analysis.

#### Financial and economic costs

18. There are two types of costs to consider. First, there are actual financial outlays or expenditures made. For RBF interventions, these would include expenditures related to bonuses paid to facilities and transfers to households, training of health workers and managers, printing of specialized forms, and special initiatives for measuring quality. Expenditures for upgrading health services, developing and maintaining the verification system, and developing and maintaining other systems required for the full functioning of the program should also be included in RBF costs where applicable.

19. Financial expenditures are only part of the picture. Other inputs, such as staff time, volunteer time, and donated equipment contribute to the cost of an RBF program. These costs, often referred to as economic costs, are the second aspect of a cost analysis. Since there are limits on the availability of time and inputs, those used to produce one health intervention generally, or the RBF intervention in particular,

cannot be used to produce another intervention, and thus represent an opportunity cost that needs to be evaluated.

20. Costs of a health intervention are based on identifying the type and quantity of inputs used to provide and deliver services, multiplying each by its unit price and the percent use for a specific intervention, such as RBF, and summing the products.

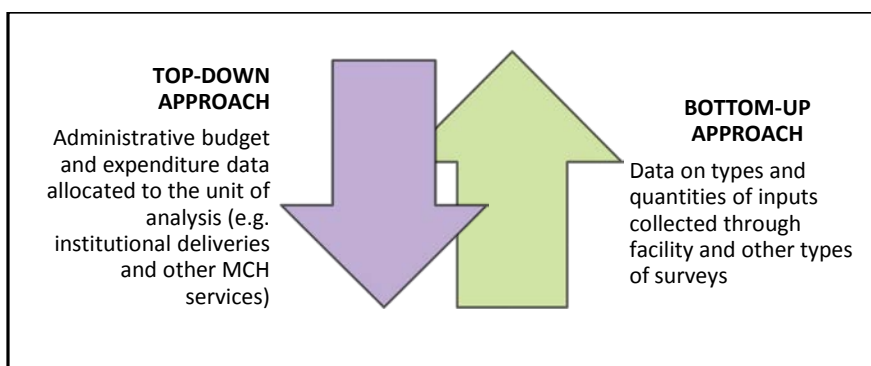
21. Conducting costing analysis from the economic perspectives is more complicated in most circumstances, as economic costs lie in different levels of the RBF program depending on how the RBF program is designed. Obtaining those costs have to track down resources used at the each level. Another challenge to obtain economic costs is that there is no system that routinely documents those costs, and primary data collection may be needed.

22. On the contrary, tracking the financial flow of the program is more feasible financially and logistically. Program operational budget and expenditure either from RBF program implementers or from funders are generally available and well documented. As the CEA focuses on incremental costs, fix costs that do not change after the introduction of the RBF program could be omitted to ease the data collection. However, it still requires tracking cost information on consumables.

### Top-down vs bottom-up approach

23. There are two basic approaches to collecting and measuring cost information (see Figure 2). A **top-down approach** to cost analysis relies on administrative records of expenditures, which are then allocated to the health intervention on the basis of a series of assumptions or proxies. In a **bottom-up approach**, data on type and quantity of inputs, as well as percent allocation of shared inputs (such as health workers) can be collected from surveys. Unit cost information may need to come from a centralized source. The bottom-up approach is commonly used in surveys of lower-level health facilities that have limited information on expenditures and budgets. It can also be used with facility-level administrative data, such as shipment of drugs and supplies from a central supply to individual health facilities or a district office. Because the impact evaluation is based on facility-specific data, this approach is recommended in this module<sup>1</sup>. Many costing studies use the combination of top-down and bottom-up approach.

**Figure 2: Approaches for Measuring the Cost of RBF Interventions**



Source: Adapted from (Bautista-Arredondo, Dmytraczenko, Kombe, & Bertozzi, 2008).

<sup>1</sup> The top-down approach can be used to allocate the costs of more complex health facilities (for example, hospitals) and administrative costs to a particular service, and is also the approach used in national health accounting. A related approach, macro-costing, allocates a hospital's costs between ambulatory and inpatient care based on the relative costs of an ambulatory visit and an inpatient bed day derived from analyses of hospital costs (D. S. Shepard, Hodgkin, & Anthony, 2000).

## Societal vs. health system perspectives

24. Both economic and financial costs can be measured from the societal or the health system's perspective. A societal perspective seeks to include all costs impacted by a program, while the health system perspective counts only the costs incurred or saved by the health system, which exclude costs from households as costs due to the loss of productivities. The societal perspective thus includes costs incurred by households, employers, non-government organizations, donors or other parties somehow impacted by an RBF intervention. As RBF programs focus more on public services, the CEA of RBF also concentrates on publicly-oriented stakeholders, which include donors, governments, and NGOs, and thus uses the health system perspective.

### CONDUCT COST ANALYSIS

25. Implementing a cost analysis for a program often involves following steps (Ozaltin & Cashin, 2014):
1. Define perspectives and timeframe
  2. Identify cost components and data collection approach
  3. Discount future costs and amortize costs if applicable
  4. Generate annual costs per capita
  5. Generate incremental costs per capita or total incremental costs

#### Define perspective and timeframe

26. For this toolkit, we determined that the health system perspective is generally the more relevant to decision makers. The toolkit is particularly designed to inform donors, their development partners, and Ministries of Health about the future implementation or expansion of RBF. The theoretical rationale for conducting CEA is the application of an objective tool for allocating scarce resources. The scarce resource in this case is the combined donor and government budget for health. Typically, government has a fixed budget for health and donors have a fixed allocation for a country. The health system perspective thus helps determine whether and how RBF helps the health system maximize the population's health within those constraints. In a country where household survey includes out of pocket expenditure of seeking care and information to derive indirect costs, such as productivity loss, a societal perspective analysis could be implemented as an option.

27. We also determined that the financial approach is more relevant than the economic perspective. The financial approach recognizes that many decisions must be made within a time-limited perspective of, say three years. During that period, within government facilities, most personnel, equipment, buildings and consumables are fixed in aggregate based on the resources available from national and donor revenues. With the exception of consumables, allocations of resources at the health-facility level also tend to be fixed over time based on historical staffing, personnel practices, and existing plant and equipment. On the other hand, quantities of consumables used in a health facility tend to vary on a quarter-by-quarter basis with demand for services at that facility. This variation occurs explicitly when a health facility can determine its own resources for purchase of consumables based on its own revenues. Under systems of cost recovery from users or insurance systems, additional users generate more revenues through payments for visits, tests, or medicines. Variation based on demand occurs implicitly if a health system allocates consumables based on need. As a facility increases services, it consumes more drugs and supplies. The combination of statistics on increased patients and consumption documents higher need, leading to higher future allocations.

28. A CEA must always have a defined time span that is applied similarly to costs and effectiveness. The time span should relate to the time span over which decisions are made. When a public law or international donor implements a program, they typically commit funding for a prescribed number of years (typically 1 to 5) of program operations. This period of funding should generally be used as the time span for the cost-effectiveness analysis. Additionally, the time span of the cost analysis needs to match of the impact evaluation, which often has a pre-post design and could help define the time span. If a time period



is not stated explicitly, it is often implicit in the lifespan of an initiative's major input, such as major equipment acquisition or the training of key staff. Implicitly, the project continues until the training becomes obsolete or the staff leave the immediate post for which they were trained. If there is no clear implicit timespan, then it is useful to choose the five-year time period used for many financial analyses. Finally, if the activity is expected to continue indefinitely in a similar steady state, then a one-year steady-state perspective is most useful.

29. If not specified, the subsection below regarding the cost analysis elaborates on our recommended approach: adopting financial cost from the health system's perspective.

### Identify cost components and data collection approach

30. From the health system perspective, the major financial costs of a RBF program often includes: (1) costs of consumables; (2) program costs of the RBF program; (3) other costs that fall outside the health system that the RBF program is targeted to. It is important for researchers to estimate cost drivers based on available data or experts' opinions, to ensure that key cost components are included in the analysis.

#### Costs of consumables

31. Costs of consumables are based on the classical principle that aggregate cost equals unit cost times quantity. Applying this principle to consumables delivered to a health facility, we estimate the quantity by the amount of consumable items that it receives – either from its own purchases or as transfers from a regional or national medical store. In principle, consumption could also be measured as the quantity of consumable items dispensed to patients or consumed in their care. If a facility maintained a comprehensive system of electronic medical records, then patient-level use could be obtained and aggregated. Without such records, it is generally not feasible to know patient-level use. The quantities received serve as a plausible proxy. If average inventories tended to be similar across years, then the approaches would be identical. While inventories of a specific product might vary from one year to the next, average inventories across all products would be expected to be similar over time.

32. The unit cost of consumables is the cost per unit delivered to a health facility. If the health facility buys the consumables from a local supplier which delivers them, the amount paid represents delivered cost. If the health facility obtains the consumable item from a central national or state warehouse, then the cost must be constructed through the steps in logistic system. It begins with the cost of the product arriving at that a national warehouse from a domestic or international supplier (i.e. c.i.f. cost, cost including insurance and freight). To this must be added the costs of the distribution system from the point of entry to the recipient facility. These include warehousing, all administration, regulation, losses (spoilage, expiration, damage, disappearance, etc.), transportation to a regional or district depot, and transportation from the region or district to the health facility. Distribution system costs are best aggregated and converted to a percentage of the cost of materials. We term the distribution cost factor as (1 + distribution cost percentage). We then apply this factor either to an individual product or to an order for a batch of products, depending on the calculation needed (see Box 1).

#### Box 1. Example.

A small country spent \$10 million to procure pharmaceuticals annually from an international supplier. The operation of the distribution system cost \$2 million. Thus, distribution adds 20 percent on average to the procurement cost of pharmaceutical products, so the distribution cost factor is 1.20. For example, a facility received a quarterly allotment of pharmaceutical products of \$10,000. The overall financial cost of that shipment is then: \$10,000 x 1.20 = 12,000.

33. Thus:

$$\text{Final cost} = [\text{acquisition cost}] \times [\text{distribution cost factor}]$$

#### RBF program costs

34. The financial costs of the RBF program are all the financial costs in running the program. These typically occur both within an RBF program office directly responsible for the program, and financial costs to health facilities for participating in the program.

35. RBF program office: Staff and other operational costs (for example, rent, utilities, transportation) for designing the system, sensitizing and updating stakeholders and participants about the program, verifying performance (for example, contracting with community associations, staff of other health facilities, and/or audit companies).

36. Participating health facilities: Their financial costs include office supplies for documenting performance, such as notebooks, added costs of electronic record systems, banking fees, and oversight mechanisms (for example, transport fees or sitting allowances for oversight or community committees). If an economic, rather than a financial perspective had been chosen, the value of time of facility staff would also be added.

37. Incentive payments: These are the payments to health facilities for their performance based on quantity and quality of agreed indicators. They are considered an added financial cost on the assumption that they are paid on top of resources that would otherwise have gone to the health facility. However, over time they could become a transfer in lieu of a financial cost. The base allocation could be lowered compared to what it would have been otherwise (for example, not fully adjusting for inflation or increase in volume) on the expectation that a facility will generate sufficient resources from RBF to make up the gap. In that way, RBF allows an average facility to stay even, a below average facility is punished, and an above average facility continues to be rewarded.

38. Inputs from the sponsor (donor) at headquarters or country levels: The sponsor, such as a donor organization, will also devote resources to the design of an RBF program and monitoring and, if needed, advising on its implementation. This information can be obtained from the donor if available, or approximated, based on the amount of staff time and travel involved, and typical unit costs for staff days and travel at comparable levels of experience.

#### *Other costs*

39. If an economic, rather than a financial, cost analysis were being conducted, a variety of induced costs would also need to be included. Most important would be changes in referrals to other health facilities. If RBF improved the capacity and utilization of health centers, they might reduce the number of normal births at district hospitals. On the other hand, if they improved prenatal care, the identification of risk conditions, and transport, they would increase the number of complicated deliveries or C-sections at hospitals.

## Discount future costs, amortize and allocate costs if applicable

### *Discount future costs*

40. When costs occur in future years, general economic principles of CEA indicate that they should be discounted to the present value as of the start of the program. In general, economic guidelines recommend that the costs be expressed each year in constant prices (those of the year in which the program began) and then discounted with a real discount rate of 3 percent per year. We use an inflation index to convert current costs into ones with constant prices, or so-called real costs. Since major inputs to health services are personnel, buildings, equipment, transport, and drugs, an overall indicator of inflation is recommended—the GDP deflator. This is typically reported by a central statistics agency. Box 2 shows it can also be derived by comparing the change in a country's GDP in constant prices with the values in current prices.

41. For discounting, typical conventions assign capital costs to the beginning of each year and personnel and operating costs to the end of each project year. Constant prices can be obtained by using the salary schedule and pharmaceutical prices of the base year. If costs are available in current prices, then costs in subsequent years can be converted to costs at constant prices by dividing by the cumulative inflation factor. Box 3 provides examples.

### **Box 2. Measuring inflation with GDP deflator**

For example, suppose a country reported its GDP in both current and constant (2010) local currency units (LCUs). We wished to measure inflation from 2012 to 2014. GDP grew from 7.5 to 8.1 billion LCU in constant prices and from 10 to 12 billion LCU in current prices.

<u>GDP (billion LCU)</u>			
Year	Constant prices	Current prices	Price index
2012	7.5	10	1.333
2014	8.1	12	1.481
Ratio			1.111

This example gives a price ratio of 1.111 or cumulative inflation of 11.1%.

**Box 3. Discounting EXAMPLES.**

Example 1 (Costs available in constant prices): An RBF program has capital costs of \$50,000, year 1 costs of \$100,000 and year 2 costs of \$200,000 (based on tariffs and incentive payments from the start of the project). The discounted costs are calculated and added as follows:

Item	Actual cost	Discount factor	Present value
Initial capital cost	\$50,000	1.000	\$50,000
Year 1 operations	\$100,000	0.971	\$97,087
Year 2 operations (constant prices)	\$200,000	0.943	<u>\$188,519</u>
Total			\$335,607

Example 2 (Costs available in current prices): An RBF program has capital costs of \$50,000, year 1 costs of \$105,000 in current prices and year 2 costs of \$210,500 in current prices. The inflation rate has been 5% per year for both years. The costs are converted to constant prices, discounted, and added as follows:

Item	Actual cost	Cumulative inflation factor	Cost in constant prices	Discount factor	Present value
Initial capital cost	\$50,000	1.000	\$50,000	1.000	\$50,000
Year 1 operations	\$105,000	1.050	\$100,000	0.971	\$97,087
Year 2 operations (current prices)	\$220,500	1.103	\$200,000	0.943	<u>\$188,519</u>
Total					\$335,607

*Annualize capital costs*

42. As RBF programs are often implemented for several years, it is likely that some capital inputs are increased due to the increased both supply or demand resulting from the program. The additional resources invested in the capital inputs need to be captured. Capital inputs, such as vehicles, motorcycles, computers, and medical equipment have a durable life of more than one year. In a cost analysis based on annual costs, there is a need to transform capital expenditures into their annual equivalents. The value of the capital equipment is divided by the relevant present value (PV). This present value is based on a discount rate and years of useful life. The years of useful life will be context specific, but some helpful guidance is that vehicles last between 3-5 years; computers have a useful life of approximately 5 years; refrigerators have a durable life of 10 years.

43. Values for the PV can be computed in Excel using the function PV. A discount rate of 3 percent per year is generally used for health projects (Jamison et al., 2006). The settings are PV(3%,N,-1), where 3 percent is the annual discount rate, N is the number of years for the asset, and -1 is used because of the convention in Excel that payments and present value are of opposite signs. The total value of an input should be divided by the PV to obtain its annualized value. For instance, if \$250,000 worth of computers was purchased for an RBF intervention, and computers have an average life of 5 years, the function PV (3%,5,-1) has a value of 4.5797. The annualized value of computers would be \$250,000 divided by 4.5797, resulting in a cost of \$54,589 per year. To make this process more convenient, the authors have created a costing template on the web. To retrieve it, open a browser and search for the following address: [www.brandeis.edu/~shepard/pv-calc.xlsx](http://www.brandeis.edu/~shepard/pv-calc.xlsx)

44. Save the resulting Excel file (denoted by .xlsx) on the user's desktop or any convenient location that will be remembered. Entering in the useful life in the yellow box, the PV appears in the green box.

45. However, if there is no change of some capital inputs, such as the cost of building, before and after the RBF program, the costs of those capital inputs could be removed from the analysis when estimating incremental costs.

#### *Allocate costs among target services*

46. Sometimes we are interested in the cost of only one or a subset of the program outputs. Examples, such as maternal child services, facility-based deliveries, vaccinations, and family planning are especially salient. For such analyses, it is necessary to allocate costs to one or more of these target services. Since quantity-based incentive payments are tied to each output indicator, their costs can be linked directly to each output indicator. Likewise, consumables can, in principle, be tied to each target service by knowing the service to which the consumable applies. Ones that apply to multiple services should, in theory, be allocated based on their shares of use. However, if such data are not available, they could be allocated in equal shares among the services to which they apply. Costs of general RBF program operations and costs of the donor organization in contributing to the design and monitoring are a joint cost. As a first approximation, these should be allocated in equal shares among all the output indicators. Another option for allocating joint cost could be based on the share of quantity-based incentive payment (or tariff) to each output indicator. However, many RBF programs purposely use the tariff as an instrument for strategic purchasing to encourage the provision of certain high priority services. Therefore, the tariff is not strictly correlated with the unit cost of service, and using the share of service-specific quantity payment as the basis for allocation joint cost can be misleading<sup>2</sup>.

47. However, there could be economies of scope, such that the first service would consume a disproportionate share of resources. As an illustration, a study of vaccination estimated the cost of a one-dose measles-only program instead of a comprehensive program entailing nine vaccination doses (D S Shepard, Sanoh, & Coffi, 1986). Because that one vaccine would require the full cold chain, registration, logistics, and administrative structure, its cost was estimated at 75 percent of the cost of the full program. If a more precise allocation is important, analysts may wish to convene a Delphi panel to examine the costs of a specific subset of indicators more precisely. See Box 4.

#### **Box 4. The Delphi technique.**

The Delphi technique is one approach for obtaining missing data or for corroborating uncertain information needed to estimate the burden of disease for the cost-effectiveness analysis. Using this method, experts are surveyed (independently and anonymously) for their estimates of the missing information, as well as the rationale behind their recommendations. These experts can also rate the quality of information they are using to generate their estimates. Individual opinions are collected, summarized, and returned to each expert for a second opinion. The process is repeated until a consensus is reached or until a certain number of rounds (three, for instance) are completed. The theory behind the technique is that the sharing of information and reasoning and repetition will improve the quality of estimates and reduce individual biases. If disagreement persists after several rounds of the exercise, the range of values may be used in a sensitivity analysis.

#### **Generate annual costs per capita**

48. To account for the difference of the population sizes in the control and intervention group, the costs per capital for each group needs to be calculated. In the intervention group, this indicator is calculated as the total cost of the RBF intervention divided by the population served by the intervention. Similarly, the

---

<sup>2</sup> For example, if family planning coverage is low, the program administrator may decide to put a high tariff to this service to encourage health workers to go out of their way to encourage family planning acceptance, whereas the true cost of providing this service is rather small relative to other services.

costs per capital in the control group(s) can be estimated using the corresponding total costs and population size.

### **Calculate incremental costs per capita or total incremental costs**

49. As noted, incremental costs are the difference between an RBF program and a control or comparison program. Because the populations served by each program are typically not identical, these are best both expressed in per capita terms. If the control is usual care, it is the difference between RBF costs per capita and control costs per capita. The RBF program in Zambia compared three alternative models: an output-based RBF program with incentive payments based on outputs in quality and quantity, an input-based financing alternative (in which facilities received extra resources averaging those of the RBF program, termed Control 1), and pure control equal to usual care (termed Control 2). In this case there are three incremental costs, each with its own interpretation:

- Output-based RBF vs. Control 2: Added cost of output-based RBF compared to usual care
- Output-based RBF vs. Control 1: Added cost of output-based RBF compared to input-based financing
- Control 1 vs. Control 2: Added cost of output-based RBF compared to input-based financing

50. Note that the three incremental costs are interrelated. Any one of these incremental costs can be derived by adding or subtracting the other two. For example, the first is the sum of the second and third. They can also be related to the incremental effectiveness, as discussed in the subsequent section.

51. When quantifying the incremental costs and effectiveness, the counterfactual scenario is trying to answer the question of “what if there is no intervention?” it implicitly wants to quantify the cost or effectiveness for a scenario for the same size of the population as the intervention, but without intervention. Therefore, if one would like to estimate the total incremental costs, it can be calculated as the incremental costs per capita multiplied by the population size in the intervention group.

52. One should be cautious about the issue of double counting. When examining financial costs, the focus is on the cash flow and the additional resources resulting from RBF. From example, sometimes, overhead costs are funded through receipts of incentives. In other cases, overhead costs do not differ between the two arms. In both circumstances, the overhead costs should not be counted.

## CHAPTER 3. ASSESSING PROGRAM EFFECTIVENESS

53. This chapter discusses the estimation of effectiveness. It uses the same perspective and timeframe as specified in the cost section, and it begins with the estimate of outputs and uses a model to estimate the change in outcomes.

54. All RBF programs measure effectiveness as the change of utilization of health services, which is a foundation of the incentive payments by the program, and is the focus of most impact evaluations (Basinga, Mayaka, & Condo, 2011; Falisse, Ndayishimiye, Kamenyero, & Bossuyt, 2014; Janssen, Ngirabega Jde, Matungwa, & Van Bastelaere, 2015; Soeters, Peerenboom, Mushagalusa, & Kimanuka, 2011; Zeng, Cros, Wright, & Shepard, 2013). However, the change of utilization of services is generally not the effectiveness measure in a classic CEA, where the effectiveness is often quantified in terms of lives saved, DALYs, or QALYs. As more health system initiatives arise and policy makers are increasingly interested in how RBF programs are compared to other initiatives, there is an increasing need to estimate the effectiveness using standardized outcome measures. To obtain the standard measure requires the conversion of utilization of health services to lives saved or QALYs/DALYs, involving several steps:

- (1) Estimate RBF programs' impact on service utilization and quality of care
- (2) Select indicators for cost-effectiveness evaluation
- (3) Convert coverage of health services and quality of care achievement to reduction of mortality, specifically the number of lives saved
- (4) Convert reduction of mortality, i.e., numbers of lives saved to QALYs/DALYs

### ESTIMATING RBF PROGRAM IMPACT ON SERVICE UTILIZATION AND QUALITY OF CARE

55. Virtually all RBF programs in the HIRTF's portfolio incentivize a package of services that incorporates both quantity and quality performance. The package can have anywhere from below 10 services offered in the health centers, mostly maternal and child health, HIV/AIDS, and family planning, to a wider spectrum that includes also TB, malaria, curative care, and surgery, to be provided in both health centers and hospitals. For quality, RBF incentivizes performance as verified and recorded during quarterly supervisory visits using a balanced score card approach. The score card covers primarily structural quality of care, although programs are making efforts to include more measurable clinical aspects. Items in the score card can be either service specific (for example, focusing on structural and clinical aspects of childbirth delivery), or cross cutting (for example, looking at the quality of health management information systems (HMIS) and facility cleanliness)<sup>3</sup>.

56. The effectiveness evaluation is heavily linked to the impact evaluation (IE) of RBF programs, the methodology of which is covered extensively in the existing World Bank Impact Evaluation Toolkit. The Toolkit was premised on a ground that that evaluations adopt a randomized controlled trial (RCT) or a quasi-experiment design and rely on primary data from household or facility surveys (World Bank, 2015). While such designs and data assure rigor in determining a RBF program's impact, they may be inadequate for the purpose of conducting a CEA. In particular, household surveys, which typically select households that have a pregnancy or childbirth in the last 2 years, do not capture program's effects on many commonly incentivized services, such as outpatient curative care, TB, prevention of mother-to-child transmission (PMTCT), and family planning. Analyses of facility surveys are typically done by questionnaire modules (such as human resource for health, lab equipment, and drugs), thus providing impact results on cross-cutting quality rather than service specific, making it difficult to convert the results to health effects.

57. It is recommended that for evaluating effectiveness, CEA should rely primarily on the most rigorous impact results from the household and facility surveys, while also make use of other data sources. Because

---

<sup>3</sup> This section describes a typical PBF program in the HRITF portfolio and does not cover the whole universe of supply side RBF design.



the cost side considers all associated costs of a full RBF program, ideally the effectiveness side takes into account all possible outcomes on both quality and quantity fronts and other system strengthening aspects, such as enhanced supportive supervision and autonomy. Two additional sources of data are available to expand the effectiveness estimate beyond what is typically produced in an impact evaluation report: the quarterly quality checklist and routine HMIS data maintained by the Ministry of Health (MoH). In addition, a CEA team is encouraged to perform additional analyses of the facility survey to obtain service specific impact information. When multiple sources of data are available, researchers need to carefully examine the accuracy and representativeness of the data. Table 1 lists sources to assess effectiveness.

**Table 1: Effectiveness estimates for CEA and suggested data sources**

Effectiveness estimates needed for CEA	Suggested data source
Coverage of key MCH services: <ul style="list-style-type: none"> <li>• <b>Antenatal care (ANC)</b></li> <li>• <b>Institutional delivery/ delivery with skilled birth attendance (SBA)</b></li> <li>• <b>Postnatal (PNC)</b></li> <li>• <b>Vaccination</b></li> </ul>	Household survey
Quality of key MCH services: <ul style="list-style-type: none"> <li>• <b>ANC</b></li> <li>• <b>Institutional delivery/ delivery with SBA</b></li> <li>• <b>PNC</b></li> </ul>	Household survey
Quality of services: service specific indicators and cross-cutting quality	Facility survey and Quarterly checklist
Other potential system effect, e.g., enhanced supportive supervision	Facility survey
Coverage or quantity of services not included in the key MCH list, such as: <ul style="list-style-type: none"> <li>• <b>Outpatient curative care</b></li> <li>• <b>Family planning</b></li> <li>• <b>TB screening and treatment</b></li> <li>• <b>Malaria treatment</b></li> <li>• <b>Treatment of acute malnutrition</b></li> <li>• <b>Minor surgery</b></li> </ul>	HMIS

Source: Authors Note: “suggested data source” means that other sources are available but the one proposed is considered to be best in depicting causal effectiveness of RBF.

58. It is clear from Table 1 that the most robust effectiveness evidence the team can obtain is on the key set of maternal and child health (MCH) services, such as ANC, institutional delivery, postnatal care PNC, and vaccination. This provides additional justification for focusing the CEA on a small package or single indicators (delivery, vaccination) rather than looking at the whole comprehensive package of services. In some countries, the household survey also provides information on quality of care in this package, such as whether the pregnant woman had blood and urine tests during her ANC checkup.

59. Given that a household survey covers only a small set of services, use of HMIS data is inevitable if teams want to perform a CEA of a full incentivized package. A major advantage of HMIS data is that they are available (in many countries) for both treatment and control facilities and they allow for the examination of trends over a long period before, during, and after the program. Unfortunately, HMIS data are known for having sub-standard quality. Furthermore, the patterns of error may be different between treatment and control facilities with direction of error being unknown, making it problematic to compare between the two using HMIS data. While countries are making great efforts to improve the quality of their HMIS data, for the purpose of the exercise at hand, i.e., conducting CEA of RBF programs, it is recommended that the facility surveys collect full HMIS data from the surveyed facilities for at least one quarter. The full HMIS data are to be collected from original registries of different services to improve the precision of data recorded.



60. Data from the quarterly quality checklist, while fairly detailed, are typically only available for treatment facilities. It is recommended that they are used to calibrate the trend line and triangulate with survey data. Currently, it is very difficult to map the items in the quarterly checklist with the items in the survey questionnaire. It is recommended that going forward, facility surveys include at least major items in the checklist, so that data are available for both baseline and end line periods in both intervention and control facilities.

### SELECTING INDICATORS FOR COST-EFFECTIVENESS ANALYSIS

61. Depending on research design of RBF programs, the evaluation of RBF on incentivized indicators could take various forms. Here we use the conventional experimental research design as an example to assess the impact of RBF on utilization of health services. As mentioned above, most RBF programs incentivize a package of services and not a few individual service indicators. Table 2 provides a summary of indicators that are commonly used for receiving incentives.

**Table 2. Common incentivized indicators in RBF programs**

Child health	Maternal health	General population
Full immunization	Antenatal visits	Hospitalization
Growth monitoring	Postnatal visits	Surgery
Curative care	Institutional delivery	Curative care
Treatment of acute malnutrition	Use of modern family planning methods	Screening and treatment of TB
	PMTCT: counseling and testing	Screening and treatment of malaria
	PMTCT: antiretroviral treatment (ART)	HIV VCT
	Intermittent preventive treatment (IPT)	
	Curative care	
	Tetanus vaccinations	

Source: Authors

62. Given the comprehensiveness of the incentivized benefit package in many RBF programs, a CEA taking into account the whole package is very challenging. In such cases, one option is to select one or several services, or a smaller and more focused package of services, to perform CEA. We have shown how to allocate overall program's cost to specific services. On the effectiveness side, the selection of which services to perform CEA can be guided by several factors:

1. RBF's impact on utilization: It is important to include services on which RBF is expected to have a large impact. However, the choice should be made before the data are available. Otherwise, the results would be a form of cherry picking, such as picking a winning lottery ticket after all the entry numbers have been exposed.
2. Program's focus: what is the key priority of the program? In the case of the World Bank's RBF programs, a clear focus is maternal and child health. So selected services could be a package of MCH services only, or could be institutional delivery and vaccination only.
3. Effectiveness: It is also important to include services that have a large impact on mortality. This could be done by reviewing literature or conducting expert consultations.
4. Cost: Similarly important is to examine the cost of RBF programs. Services that consume a large amount of incentives should be included in the evaluation.

## CONVERTING COVERAGE OF SERVICES TO LIVES SAVED

63. Effectiveness depends on (1) coverage of the interventions; (2) size of affected population—the absolute population of the intervention area; (3) proportion of the population affected by the intervention of interest; (4) initial risk of the underlying population (which may depend on demographics, location, poverty, and other factors); (5) the effect size of the intervention in terms of preventing occurrence of illnesses or deaths, which is often measured in terms of relative risks (RR) or odds ratio (OR).

64. Converting service coverage to lives saved is the core component of the effectiveness evaluation in the CEA of RBF. It is assumed that the increase of utilization of services will bring about health benefits to target populations through (1) preventing occurrence of diseases (i.e. vaccination); (2) effectively managing illnesses to improve the quality of life, and (3) reducing risks of death. So far, the software Lives Saved Tool (LiST), discussed below, is the most comprehensive tool available to convert coverage of a wide variety of health services to health outcomes (that is, reduction of mortality/the number of lives saved), and is employed in the toolkit to estimate the effectiveness of the CEA of RBF programs.

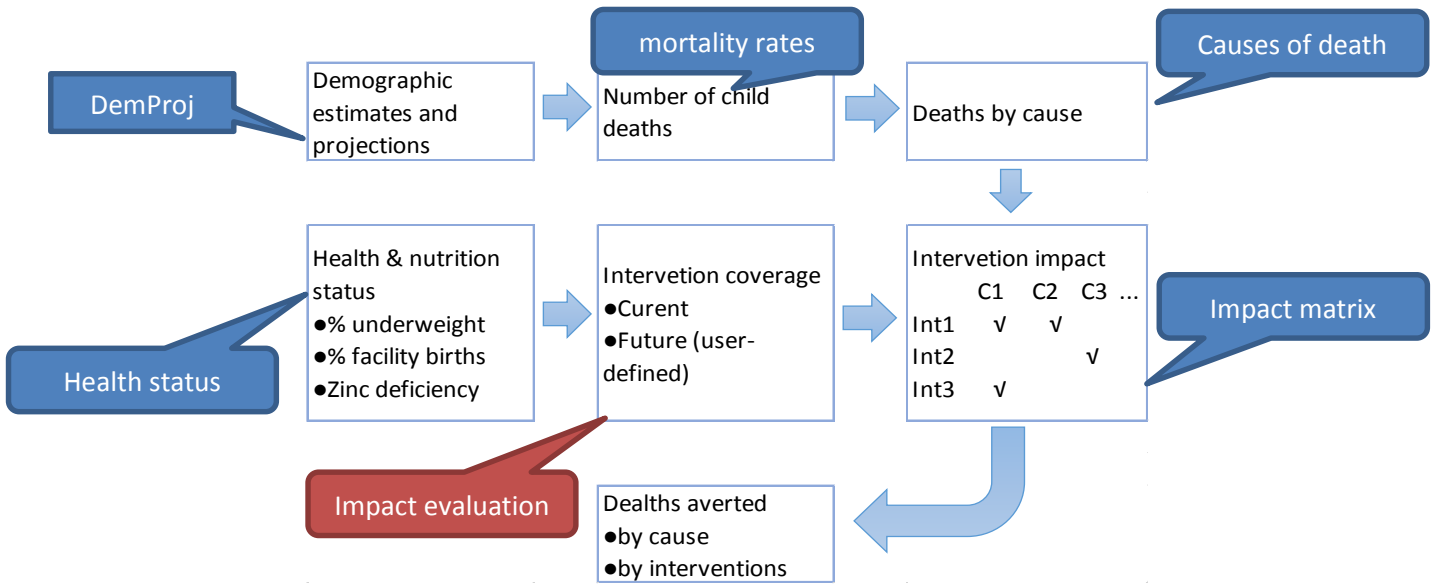
### Basics of LiST tool

65. The LiST tool was developed by researchers at Johns Hopkins University (Baltimore, MD) and was incorporated into the software of SPECTRUM, which can be downloaded from <http://www.avenirhealth.org/software-spectrum>. The LiST tool aims to estimate the impact of different interventions on MCH outcomes, which is consistent to the goal of RBF programs. The LiST tool has been widely applied in projecting the health impact of interventions and is advocated by the United Nations Children's Fund (UNICEF) for decision making (Boschi-Pinto, Young, & Black, 2010; Stenberg et al., 2014). The LiST tool requires numerous inputs to produce desired results on health outcomes. Figure 3 shows the inputs to the LiST tool. One set of key inputs for the LiST tool will come from results from the impact evaluation to provide information on baseline and end line coverage of the services. Other inputs are often preloaded based on literature and country-specific statistics. We will use LiST in this toolkit to obtain lives saved from interventions<sup>4</sup>.

---

<sup>4</sup> The detailed manual of SPECTRUM can be found at <http://avenirhealth.org/Download/Spectrum/Manuals/SpectrumManualE.pdf>.

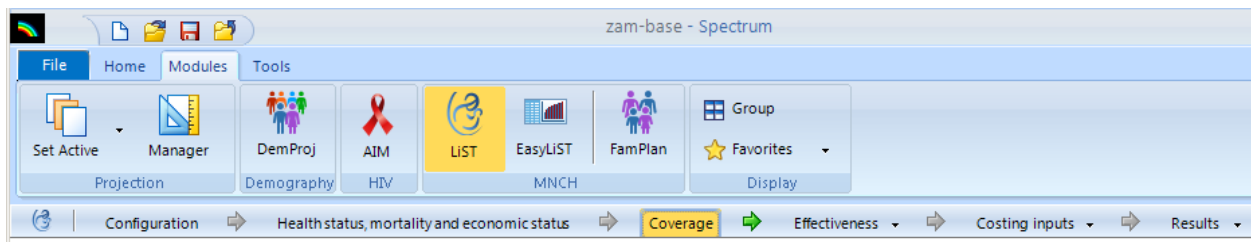
**Figure 3. Schema of the LiST tool**



Source: (Boschi-Pinto et al., 2010)

66. There are several issues to be aware of when using LiST. First, LiST does not work alone, and it is based on the module of DemProj, the primary purpose of which is to project population size at the national or subnational level. In RBF program evaluation, DemProj will project the population size in RBF areas. Second, if the CEA includes interventions of family planning, the module of FamPlan will be used. Third, LiST can also be used for HIV/AIDS interventions, in conjunction with AIDS impact model (AIM), to estimate the impact of AIDS treatment and PMTCT services on health outcomes. It should be noted that LiST cannot handle RBF's impact on adult health (except for health among pregnant women). Assumptions are required if the impact of an RBF program on adult health is included in a CEA. Fourthly, in addition to the inputs of current and future coverage of health services, there are other parameters that should be adjusted before implementing the software. Figure 4 shows the LiST tool interface in Spectrum.

**Figure 4. LiST tool interface in SPECTRUM**



Source: <http://www.avenirhealth.org/software-spectrum>

67. The key formula (Avenir Health, 2015) to estimate the impact in the LiST is

$$\%RedMort^{ij}_t = \{I^{ij} \times (P^j_t - P^j_o) \times AF^{ij}\} / (1 - I^{ij} \times P^j_o)$$

for  $J$ th intervention,  $i$ th causes.

where, %RedMort<sup>ij</sup> represents the reduction in mortality (i) before and after an intervention (j) implemented over time, t

I<sup>ij</sup> = Effectiveness of intervention j in reducing mortality from cause i

P<sub>0</sub><sup>j</sup> = Baseline coverage of the intervention j

P<sub>t</sub><sup>j</sup> = End line coverage for the intervention j

AF<sup>ij</sup> = Affected Fraction; proportion of deaths from the cause that is due to the specific condition addressed by the intervention

i = mortality for specific cause of death

j = intervention that reduces mortality I,

With multiple interventions, the model uses a multiplicative approach to estimate the combined effect.

$$\%RedMort_t = 1 - \prod_{i=1}^n \prod_{j=1}^k (1 - \%redMort_t^{ij})$$

68. Thus converting proportional coverage of services (P) to impact on mortality (I) requires obtaining parameters either from country statistics or international literature. The key parameters are: (1) distribution of causes of death for neonatal, post-neonatal, as well as pregnant women; (2) the change in coverage of various interventions; (3) effectiveness of the interventions in terms of reduction of mortality.

### Steps to using LiST for estimating number of lives saved

69. Once researchers have set up population and health status parameters, the LiST tool could be applied to model the number of lives saved from interventions. Here we focus on the number of lives saved because most interventions in RBF programs focus on preventing or treating acute illnesses, where the impact on improving morbidity, in comparison with improving mortality, is likely to be small.

70. Examining the relationship between disease burden from mortality and that from morbidity can help improve the accuracy of the estimation. According to WHO, about 75 percent of total disease burden in Africa was due to mortality in 2004 (World Health Organization, 2008). Assuming this pattern is applicable to the intervention of interest in terms of gaining QALYs, the results derived from the reduction of mortality can be multiplied by a factor of 1.33 (calculated as 1/0.75) to obtain the gains from the reduction of both mortality and morbidity. Similar data for 2010 and 2013 are available from the Institute of Health Metrics and Evaluation (IHME) at the University of Washington ([www.healthdata.org](http://www.healthdata.org)). This calculation uses the simplifying assumption that the effect of an intervention on morbidity is proportional to its effect on mortality. Refining this assumption, such as might be needed for many forms of curative care which address health problems that are uncomfortable or even disabling, but not fatal, would require substantial additional modeling that is beyond the scope of this toolkit. Unfortunately, we are not aware of another tool as powerful and user friendly as LiST to address morbidity impacts.

71. To use the LiST tool, it is required to:

1. Calculate baseline and end line coverage of services, as well the coverage of service in the interim.
2. Adjust key parameters in LiST tool
3. Set up the baseline and alternative scenario: The baseline scenario tries modeling the situation on what the service utilization would be if there were no RBF. The information from the control group will be used to generate such a scenario. The alternative scenario provides utilization of service under an RBF program. If an RBF program has an arm of input financing, that could be included as an alternative scenario as well.
4. Implement LiST tool and compare outcomes: each scenario will generate a set of health outcomes regarding mortality. We are able to compare the outcome directly.

72. Each of these steps is explained in detail in the following sections.

### Calculate baseline and end line coverage of services

73. In the baseline scenario, as counterfactual of “what would happen if there were no interventions,” the time trend in the control group(s) should be factored in the scenario for the intervention group. In the example below (Table 3), the net impact of the RBF program should consider the time trend of 20 percent (50-30 percent) in the control group.

**Table 3. Example of impact evaluation**

Group	Baseline	End line	Difference
Intervention group	40%	75%	35%
Control group	30%	50%	20%

74. The difference in difference analysis shows the RBF program improved the institutional delivery by 15% (35%-20%). Assuming the baseline year is 2012 and the end line year is 2014, the coverage of institutional delivery for the baseline scenario without RBF and the scenario with RBF is presented below (Table 4).

**Table 4. Coverage of institutional delivery under two scenarios**

	2012	2013	2014
Baseline scenario	40%	?	60%
Scenario with RBF	40%	?	75%

75. However, the coverage of institutional delivery for the year 2013 is unknown in Table 4. Assumptions will be needed to estimate it if a country does not conduct an annual assessment of the RBF program. Here we provide two examples to conduct the estimation, one using a linear assumption and the other using an “S” shaped assumption for the coverage of health services.

76. With the linear assumption, Table 5 provides the estimates for the coverage of institutional delivery in 2013 for both baseline and RBF scenarios.

**Table 5. Coverage using linear interpolation**

	2012	2013	2014
Baseline scenario	40%	50%	60%
Scenario with RBF	40%	57.5%	75%

77. If we assume the increase of the institutional delivery follows an “S” shape, then we can use logistic function to estimate the coverage for 2013.

$$\ln\left(\frac{coverage}{1 - coverage}\right) = a + b \text{ year}$$

78. Under the “S” shape assumption, Table 6 shows the coverage for the baseline and RBF scenarios

**Table 6. Coverage using logistic interpolation**

	2012	2013	2014
Base scenario	40.0%	50.0%	60.0%
Scenario with RBF	40.0%	58.6%	75.0%

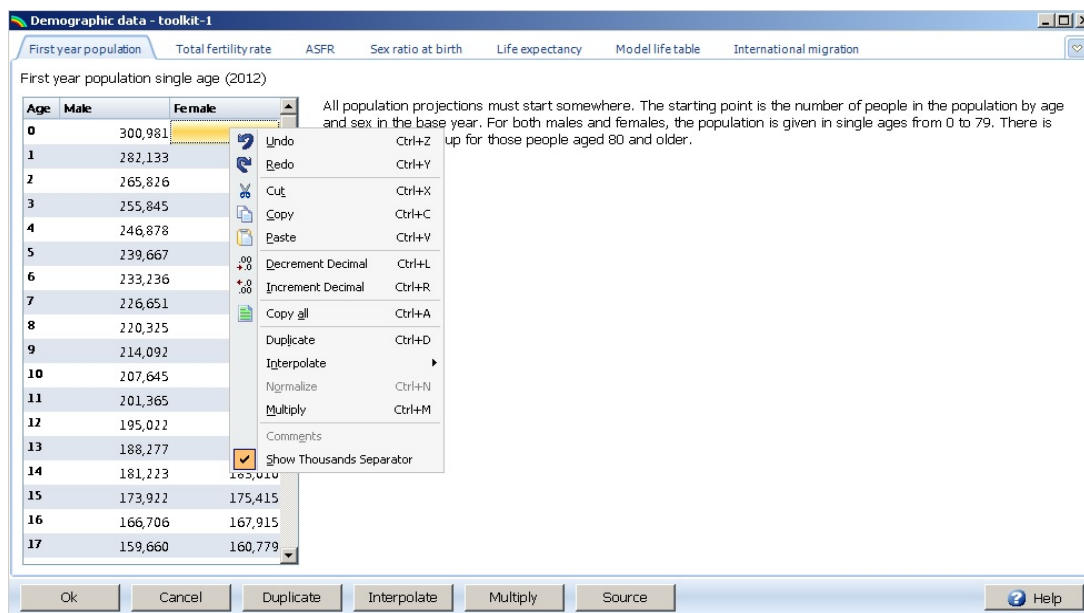
79. We will then use the coverage information to feed to the LiST tool to convert the coverage of services into lives saved.

### Adjustment of key parameters in LiST tool

80. The key parameters of the LiST, as shown in Figure 3, will need to be adjusted to reflect the situation of the RBF program. The adjustments are explained below.

81. **Demographics:** The LiST tool starts with demography projection, which is loaded in a separate module named “DemProj.” In this module, researchers can select a country to load its national population data from the United Nations Population Fund (UNFPA). If an RBF program operates at the subnational level, such as the regional or district level, or in areas where demographic characteristics are so different from national characteristics, parameters for projecting population, such as total fertility rate, age specific fertility rate, sex ratio at birth, life expectancy, immigration, and baseline population size, should be adjusted based on known statistics. If one assumes that all parameters for population projection are similar to national statistics and the only difference is the size of baseline population, researchers can apply a fraction factor to the national population to obtain the sub-national population, using the function of “Multiply” in the figure below. Figure 5 shows the tabs where parameters could be changed to modify the population projection.

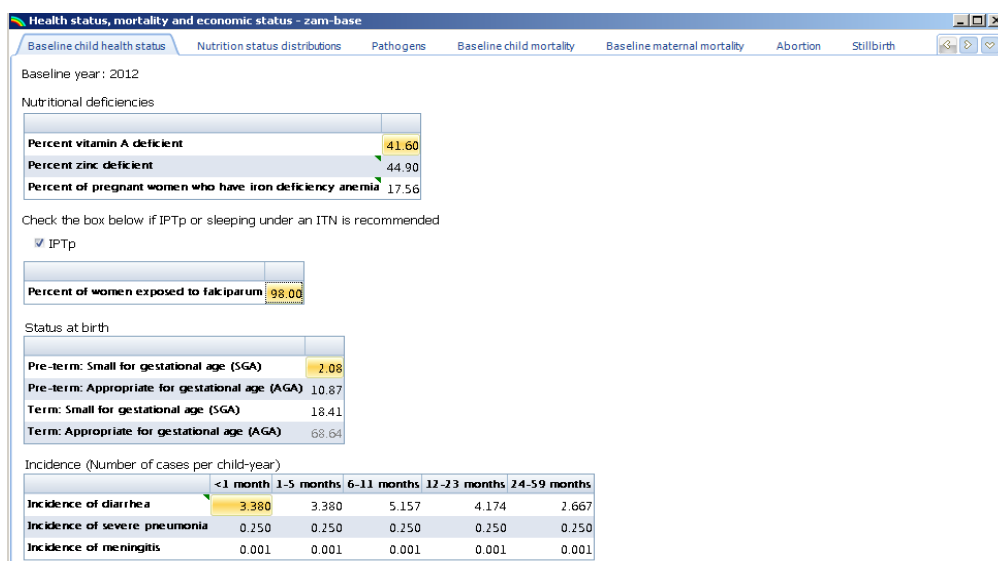
Figure 5. Tabs for changing population projection



Source: <http://www.avenirhealth.org/software-spectrum>

82. **Health Status:** When the country population data are loaded, the associated data on nutrition status, health status, abortion rate and still birth rate and its causes are also loaded. The default values are drawn mainly from literature and may not be the data from the country of the study. If the CEA team has country-specific data, those numbers could be modified to better fit the model. Figure 6 show the parameters on baseline child health status for adjustment. As some interventions are effective for a particular cause(s) of death, setting these parameters appropriately would ensure the accuracy of the estimation. Other tabs, such as pathogens for diarrhea, pneumonia and meningitis, can be changed too if necessary.

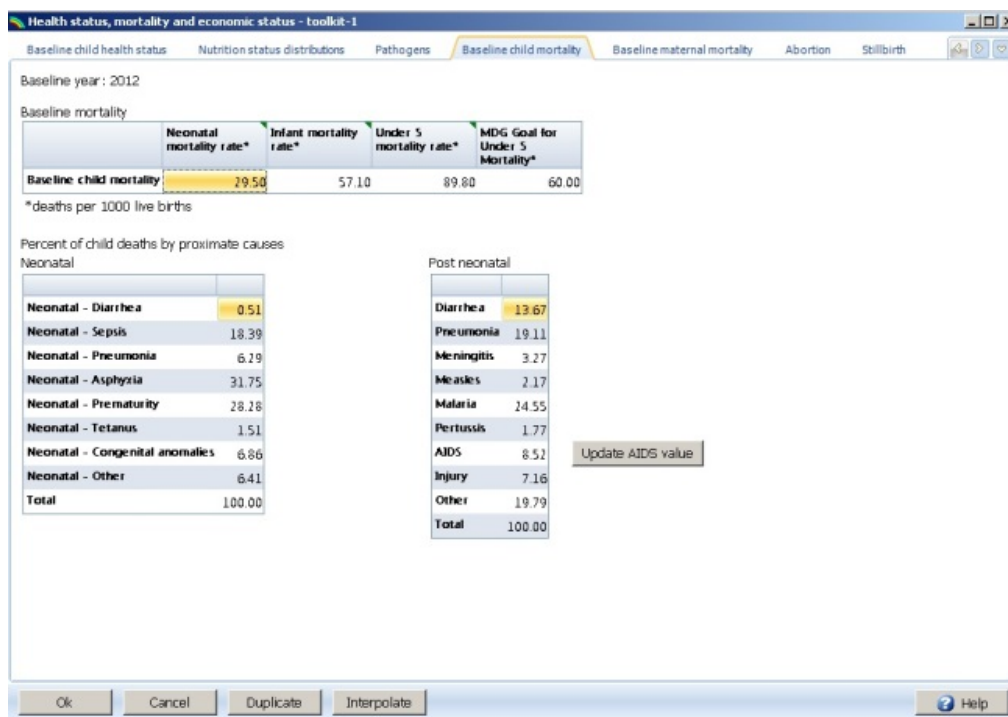
Figure 6. Parameters for baseline child health status



Source: <http://www.avenirhealth.org/software-spectrum>

83. **Mortality and causes of death:** Before estimating the impact of the maternal and child interventions, it is important to ensure mortality parameters are correctly applied in the model, which include baseline maternal mortality and child mortality and causes of death. In the tabs of baseline child mortality and baseline maternal mortality (Figure 7), researchers can review and update associated parameters if needed.

Figure 7. Parameters for mortality rates



Source: <http://www.avenirhealth.org/software-spectrum>

84. **Impact matrix:** LiST tool contains a wide variety of maternal child health service interventions. Table 7 provides key services that could be modelled in the tool, which is consistent with incentivized services under RBF.

**Table 7. Key services included in LiST**

Category	Services
Pregnancy	Antenatal care
	Tetanus toxoid vaccination
	IPT in pregnancy (IPTp)
	Calcium supplementation
	Iron folate supplementation
Childbirth	Skilled birth attendance
	Institutional delivery
Preventive	Postnatal care
	Vitamin A supplementation
	Zinc supplementation
Vaccine	BCG, polio, pentavalent, pneumococcal, rotavirus, measles
Curative care	Maternal sepsis case management
	Diarrhea treatment
	Antimalarial treatment

85. The impact of these services is modelled through direct or indirect channels. For example, institutional delivery/skilled birth delivery saves lives indirectly through its influence on the distribution on basic emergency obstetric care, comprehensive emergency obstetric care, essential care for pregnant women, and clean practices of delivery, etc. (Table 8), while the impact of IPTp is directly modelled using the efficacy of IPTp.

**Table 8. Interventions that are associated with institutional delivery and their impact on causes of maternal death**

Interventions associated with institutional Delivery	Effect size* on				
	Antepartum hemorrhage	Postpartum hemorrhage	Sepsis	Obstructed labor	Hyper-tension
Basic Emergency Obstetric Care	0.2	0.65	0.5	0.38	
Comprehensive Emergency Obstetric Care	0.8	0.95	0.7	0.99	0.96
Active management of the third state of labor		0.27			
Antibiotic for pPRoM			0.26		
Clean practices and immediate essential newborn care(home)			0.1		
Essential care for all women and immediate essential newborn care (facility)			0.1		

Source: Spectrum Manual (Avenir Health, 2015)



\*Effect size shows the proportion of deaths due to a specific cause that are reduced by the intervention. For example, Basic Emergency Obstetric Care reduces by 0.2 (that is, 20 percent) the number of deaths due to ante-partum hemorrhage.

86. Under the menu of “effectiveness,” researchers can modify the effectiveness size of health interventions, if they have recent country-specific data. For example for maternal interventions, the LiST tool provides default values for the effectiveness of basic emergency obstetrical care or (BEmOC) and comprehensive emergency obstetrical care (CEmOC) on antepartum hemorrhage. Parameters can also be modified for other causes of maternal death, including postpartum hemorrhage, hypertensive diseases of pregnancy, sepsis, abortion, obstructed labor, ectopic, malaria, and other direct and indirect causes.

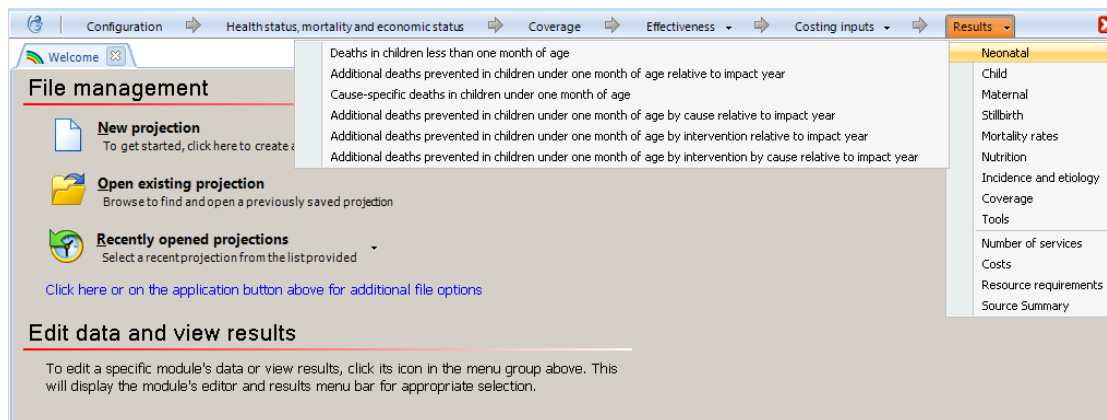
### Set up baseline and alternative scenario for key services

87. With the adjustment of the key parameters and the baseline coverage of the services of interest to be modelled, a baseline scenario will be fully established. Researchers will have to save the LiST project for the baseline scenario and repeat the same process to establish an alternative scenario by changing the coverage of services of interest and save the LiST project in another name. In Spectrum, more than two scenarios can be generated, and each scenario should be saved separately in order to compare results.

### Implement LiST tool and generate results

88. In order to compare the results from different scenarios, LiST allows users to load multiple projects simultaneously and calculate results for each project, such as neonatal, child, and maternal deaths, mortality rates, nutrition status and incidence of illness, as shown in Figure 8. For neonatal, child and maternal deaths, LiST provides not only the overall number of deaths under each scenario, but also the number of cause-specific deaths.

**Figure 8. Parameters for results on neonatal, child, and maternal deaths**



Source: <http://www.avenirhealth.org/software-spectrum>

89. Table 9 shows the number of deaths among pregnant women and children under five for two scenarios (baseline scenarios and a scenario with increased coverage of institutional delivery) in a hypothetical country. An impact evaluation typically has data for each year from the intervention (RBF) areas and the control areas. The LiST tool is set up to compare a “Baseline” or current policy with an “Alternative” policy. As discussed under “Calculate baseline and end line coverage of services,” to adapt the LiST tool for the needed calculation, some adjustments were needed. Specifically, we adjusted the Baseline to reflect the situation that would have occurred in the intervention area if the intervention had not been implemented. In that case, it would have started with the actual value in the intervention area prior to the start of the RBF program, and it would have evolved according to the subsequent year-to-year changes observed in the control area. Thus, our baseline is set to the population size and distribution of the

intervention area. The baseline initial coverage is based on the actual initial coverage in the intervention area. The baseline coverage in each subsequent year is based on the baseline coverage in the intervention area plus the change in the control area from the first year to that year. The coverage in the final year under baseline thus reflects the change in coverage in the control area. The lives saved under “Baseline” indicate the number of lives that would have been saved in the intervention area due solely to trends in coverage observed in the control area.

90. The “Alternative” shows the actual lives saved in the intervention area due to increased coverage in that area. With LiST and a little more calculation using Excel, the net number of deaths averted due specifically to RBF is estimated as the “Difference” in Table 9 and subsequent tables. Adding these difference amounts in Table 9 for the years 2011, 2012 and 2013 shows that 33 maternal deaths and 442 child deaths were averted over the three years.

**Table 9. Number of lives saved over the three years**

Year	Baseline	Alternative	Difference
<b>Projected maternal deaths</b>			
2011	165	165	0
2012	168	163	5
2013	171	143	28
<b>Projected deaths in children under 5</b>			
2011	4787	4787	0
2012	5528	5454	74
2013	6005	5637	368

Source: Authors

### Detailed illustration of core MCH services

#### *Institutional delivery*

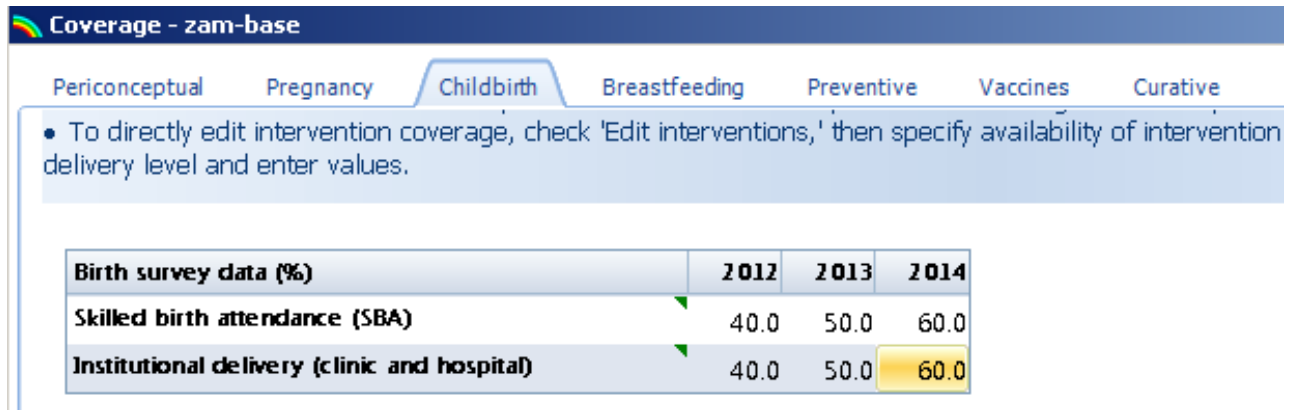
91. Let us first compare the baseline scenario and the scenario where institutional delivery is improved. We use parameters from Table 10 to simulate the analysis.

**Table 10. Institutional delivery coverage from 2012-2014**

	2012	2013	2014
<b>Baseline scenario</b>	40.0%	50.0%	60.0%
<b>Scenario with RBF</b>	40.0%	58.6%	75.0%

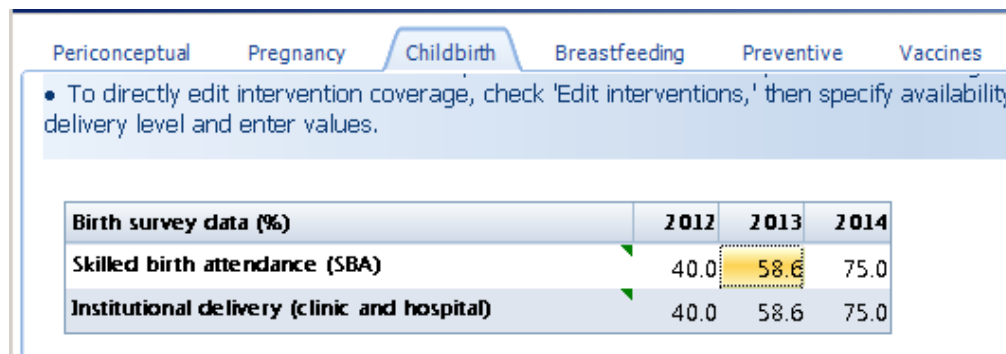
92. Setup the baseline scenario: In this example, we focus on institutional delivery/skilled birth attendance. In the LiST tool, under the LiST main menu, click “coverage” → “Childbirth”, where users can change the coverage of skilled birth attendance (SBA)/institutional delivery). Under the baseline scenario, enter the values estimated from Table 10, as shown in Figure 9, and save the LiST project.

**Figure 9. Utilization of institutional delivery in LiST**



93. Under the alternative scenario, users need to change the coverage of institutional delivery under the same tab as the one in the baseline scenario (Figure 10), and save the project in another name.

**Figure 10. Utilization of institutional delivery under alternative scenario**



94. Assuming that other parameters are appropriate, after running the model, the results on child and maternal deaths are shown in Figures 11 and 12, respectively.

**Figure 11. Deaths in children under 5 from the simulation of the two scenarios**

Deaths in children under-five years of age (Total (0-60 months))

	2012	2013	2014
<b>zam-base</b>			
Total (0-60 months)	4,800	4,541	4,676
<1 month	1,621	1,348	1,263
1-59 months	3,179	3,193	3,413
<b>zam-ins</b>			
Total (0-60 months)	4,800	4,449	4,523
<1 month	1,621	1,255	1,106
1-59 months	3,179	3,194	3,417

**Figure 12. Maternal deaths from the simulation of the two scenarios**

Maternal deaths

	zam-base	zam-ins
<b>2012</b>	164	164
<b>2013</b>	144	137
<b>2014</b>	138	127

95. The number of lives saved for pregnant women and children combined can be readily calculated by summing the respective results, which is shown in Table 11.

**Table 11. Summary of lives saved from improving institutional delivery**

	Baseline scenario	Alternative scenario	Lives saved
<b>Pregnant women</b>			
2012	164	164	0
2013	144	137	7
2014	138	127	11
Total	446	428	<b>18</b>
<b>Children &lt;1 month</b>			
2012	1621	1621	0
2013	1348	1255	93
2014	1263	1106	157
Total	4232	3982	<b>250</b>
<b>Children of 1-59 months</b>			
2012	3179	3179	0
2013	3193	3194	-1
2014	3413	3417	-4
Total	9785	9790	<b>-5</b>

Source: Authors

96. In total, 18 pregnant women and 250 children under 1 month are saved from improved institutional delivery. Due to more children surviving the first month of birth under the scenario with improved institutional delivery, 5 more children ages 1-60 months die.

*Antenatal care, IPTp, and post-natal care*

97. Similarly, if one would like to estimate the impact of ANC, IPTp and PNC on reduction of deaths, corresponding parameters should be changed. The parameters for ANC and IPTp are under the tab of "Coverage" → "Pregnancy" (Figure 13), while the parameters for PNC are under the tab of "Coverage" → "Preventive".

**Figure 13. Coverage for pregnancy-related services**

Periconceptual **Pregnancy** Childbirth Breastfeeding Preventive Vaccines Curative

**Pregnancy options**

- Check 'Antenatal care...' and enter % coverage for '\*Antenatal care.' Antenatal care interventions
- Uncheck 'Antenatal care...' to directly enter coverage for Antenatal interventions.

Intervention (%)	2012	2013	2014
<b>*Antenatal care</b>	60.3	60.3	60.3
<b>Routine</b>			
TT - Tetanus toxoid vaccination	81.0	81.0	81.0
IPTp - Pregnant women protected via intermittent preventiv...	70.2	70.2	70.2
*Syphilis detection and treatment	30.2	30.2	30.2
<b>Nutritional</b>			
*Calcium supplementation	3.0	3.0	3.0
Micronutrient supplementation (multiple micronutrients + iro...	-	-	-
Iron folate supplementation	0.0	0.0	0.0
Multiple micronutrient supplementation	0.0	0.0	0.0
Balanced energy supplementation	0.0	0.0	0.0
<b>Case management</b>			
*Hypertensive disease case management	0.0	0.0	0.0
*Diabetes case management	3.0	3.0	3.0
Malaria case management	0.0	0.0	0.0
*MgSO4 - Management of pre-eclampsia	3.0	3.0	3.0
<b>Other</b>			
*FGR - Fetal growth restriction detection and management	3.0	3.0	3.0
<b>HIV</b>			
PMTCT - Prevention of mother to child transmission of HIV (i...			

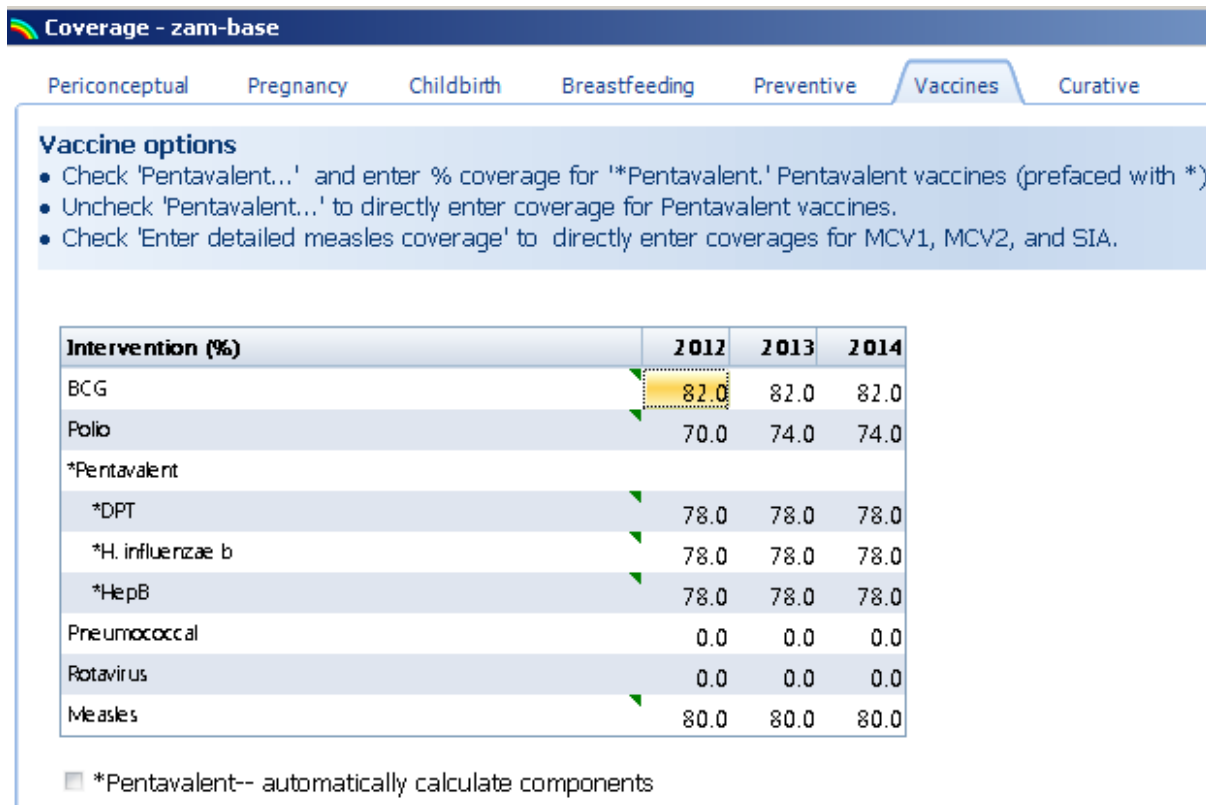
\*Antenatal care-- automatically calculate components

Source: <http://www.avenirhealth.org/software-spectrum>

**Vaccination**

98. Under “Coverage” → “Vaccines”, the coverage of vaccination could be changed and modelled for BCG, polio, pentavalent, pneumococcal, rotavirus and measles (Figure 14).

**Figure 14. Coverage of vaccinations in LiST**



Source: LiST

### *Family planning*

99. Family planning plays an important role in reducing maternal and child mortality. To model the impact of family planning, a separate module, named “FamPlan” needs to be used.

100. Under the FamPlan, users are able to modify the coverage of various family planning approaches (under tab of “Family planning” → “Method mix,” Figure 15), and effectiveness of family approaches (under the tab of “Family Planning” → “Effectiveness,” Figure 16).

**Figure 15. Utilization of family planning**

Method Mix			
	2012	2013	2014
<b>Condom</b>			
Female condom	0.2	0.2	0.2
Male condom	11.5	11.5	11.5
<b>Sterilization</b>			
Female sterilization	4.6	4.6	4.6
<b>Injectable</b>			
3 month (Depo Provera)	20.7	20.7	20.7
<b>Implant</b>			
Jadelle (5 years)	1.0	1.0	1.0
<b>IUD</b>			
Copper-T 380-A IUD (10 year)	0.2	0.2	0.2
<b>Pill</b>			
Standard Daily regimen	26.8	26.8	26.8
<b>Traditional</b>			

**Figure 16. Effectiveness of family planning interventions**

Effectiveness			
	2012	2013	2014
Female condom	81	81	81
Male condom	81	81	81
Female sterilization	100	100	100
Withdrawal	50	50	50
Periodic abstinence	50	50	50
Traditional (not specified)	50	50	50
3 month (Depo Provera)	100	100	100
Jadelle (5 years)	100	100	100
Standard Daily regimen	92	92	92
Copper-T 380-A IUD (10 year)	96	96	96
LAM	96	96	96

### Other considerations

101. For services consisting of many disease categories, such as curative care for adults and minor surgeries, and services that are not included in the LiST, the LiST tool is not able to estimate their impact. A few aspects should be considered:

1. Distribution of curative visits and minor surgeries. Users need to select diseases that have a high frequency of health facility visits or surgical cases.
2. The treatment effect. Users need to consider including the diseases that a treatment can effectively reduce mortality and morbidity.

102. Once diseases are selected, researchers would need to review literature to estimate the efficacy of the treatment/surgery.

103. There are several approaches that could be adopted to deal with these services:

1. If literature has DALYs averted or QALYs gained per case for treatment, those values could be used directly
2. If literature does not have information on DALYs or QALYs associated with a particular treatment, and if the treatment of illness among children could be modelled with LiST, such as the treatment of malaria, the results from children could be used to estimate the treatment effect among adults, with proper assumptions.
3. If literature does not have information on DALYs or QALYs for a particular treatment, and no effect could be estimated from children to derive the impact on adults, then a comprehensive estimate should be conducted based on disease distribution, efficacy of treatment, and population size affected, a similar process as LiST does for maternal and child services.

#### **INCLUDE QUALITY OF CARE AND OTHER POTENTIAL SYSTEM EFFECTS IN THE EFFECTIVENESS ESTIMATION**

104. Many RBF programs pay for improved quality of care and encourage other system strengthening efforts, such as enhanced supportive supervision and better decision making power for frontline workers.

Without considering such aspects, the effectiveness of the program is likely to be underestimated. Most notably, a well-designed impact evaluation will estimate the magnitude of the improvement of quality of care, using such tools as exit surveys of patients at health facilities, observations at facility surveys, and scenarios with health workers. However, the way these measures of quality of care translate to DALYs or QALYs is not directly observed. A Delphi process, as introduced above, would help to obtain experts' opinion to address this issue.

105. Appendix 1 provides an example of a Delphi questionnaire to obtain expert opinion on the impact of quality of care on health outcomes. In the survey, we try to quantify, using a quadratic function, to estimate the relationship between quality of care and its health impact. The quadratic function is used because of its flexibility to accommodate different scenarios on their relationship (for example, containing both linear or non-linear relationship). See Box 5 and Figure 17.

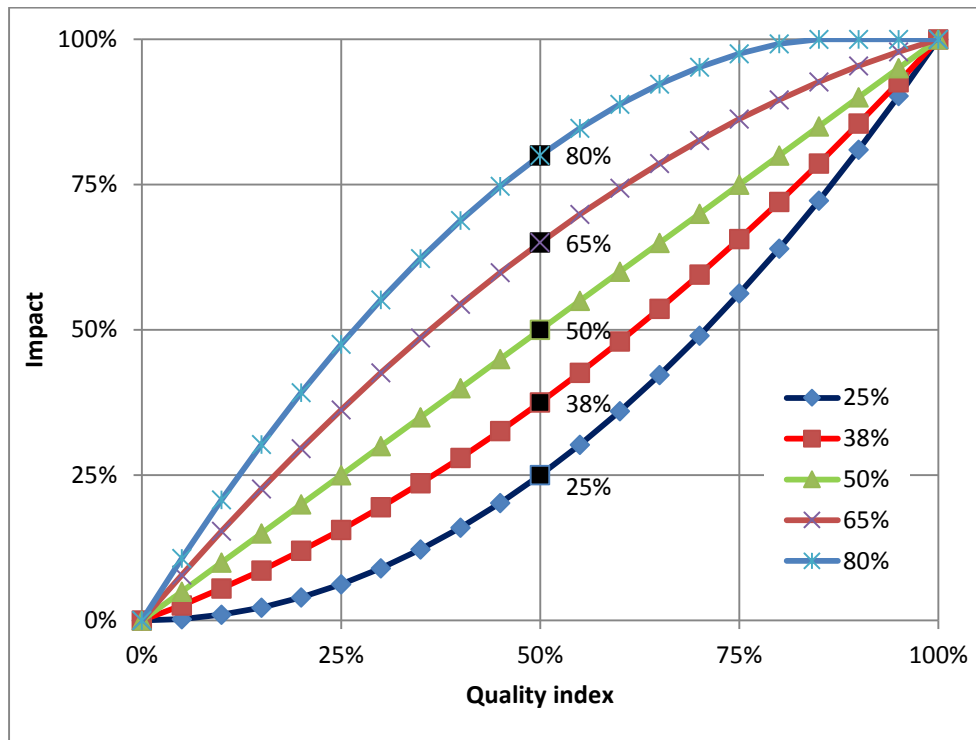
#### **Box 5. The quadratic function**

A quadratic function takes the form  $y=ax^2 + bx + c$ , where  $y$  is health impact,  $x$  is quality of care, and  $a$ ,  $b$ , and  $c$  are coefficients to be estimated. Assuming that an intervention with quality index of 0 is equivalent to no intervention and does not bring side effects then the function has to go through the point (0, 0). An intervention with quality index of 1 means the intervention can fulfill its potential to reduce the mortality rate equivalent to that of efficacy studies in the literature. Thus, the function ought to go through the point (1,1). Then a point estimated from the Delphi survey will determine the shape of the curve.

The linear function is one special case of this family of quadratic functions. It occurs when the intervention is equally sensitive to quality throughout its full range of application. The Excel template for these quadratic formula is available on the web at [www.brandeis.edu/~shepard/quadratic.xlsx](http://www.brandeis.edu/~shepard/quadratic.xlsx)



**Figure 17. Potential relationship between quality of care and its impact\***



Source: Authors

\* Impact is quantified as the percentage of full health benefits achieved from the intervention.

106. Assuming the result from the Delphi survey shows that 50 percent of quality leads to 20 percent health impact, then the function will go through the point of (0.5, 0.2). With three points, the three coefficients, a, b and c, are initially unknowns that will be determined.

$$0 = a * 0^2 + b * 0 + c$$

$$1 = a * 1^2 + b * 1 + c$$

$$0.2 = a * 0.5^2 + b * 0.5 + c$$

we can obtain a = 1.2, b=-0.2 and c = 0.

107. We assume that the results from LiST take perfect quality, meaning having the quality index of 100 percent. If the baseline and end line quality of care is not 100 percent, then the impact on lives saved from LiST should be adjusted according to the quadratic function. Here we use quality-adjusted coverage approach to correct the results from LiST (See Box 6).

### Box 6. Quality adjustment impact

Case 1. This example assumes that the Delphi survey reveals that 50 percent quality leads to 20 percent health impact (measured as number of deaths averted) for vaccination among children. The quality of care remains constant during the RBF implementation period (1 year), as 50 percent, while the coverage has been improved, increasing from 40 percent to 60 percent. The LiST tool estimates 1000 would die without RBF, while 900 would die with RBF, with 100 lives saved from RBF under perfect quality of care. . Due to low quality of care, instead of saving 100 children, the RBF saves 20 children (100\*20%).

Case 2. This example has all the assumptions as in case 1, plus the assumption that the quality of care is increased from 50% to 60% in a year. Instead of saving 20 children, this scenario shows that the improved quality of care saved 33 more children, with a total of 53 children saved by the RBF program (100/20%\*11%), as calculated below. Please note that this is a dramatic example. A 20 percent increase of quality of care doubles health gains.

		Quality of care	Health impact	Coverage	Quality-adjusted coverage	Difference	Life saved
LiST	Baseline	100%	100%	40%	40%	20%	100
	End line	100%	100%	60%	60%		
Adjustment	Baseline	50%	20%	40%	8.0%	10.6%	<b>53</b>
	End line	60%	31%	60%	18.6%		

108. A Delphi survey would also help generate the quality index from primary survey results. For services included in RBF programs, the quality could be measured with several components, including general quality indicators (that is, administration and management, and HMIS) and service-specific quality indicators (i.e., supply chain, clinical process, and availability of equipment), and these components play different roles in determining the overall quality of a particular service. For example, the supply chain weights more in quantifying quality of service for vaccination, while qualification and clinical process probably play a more important role in estimating the quality of prenatal care. Table 12 shows an example of various quality components. Some are service-specific while others are general.

**Table 12. Components of general and service-specific quality indicators**

General quality indicators	Service-specific quality indicators
Infrastructure	Clinical processes
Administration and management	Drugs and supplies
Human resource for health	Equipment
HMIS	Staff with training
Leadership and autonomy	
General equipment	

Source: authors, based on multiple health facility survey questionnaires

109. Appendixes 2 and 3 provide an example of Delphi questionnaire to estimate the relative importance of generic and specific quality indicator, and the components included in the service-specific quality indicators in determining the overall quality of care for each service. See Box 7.

110. If quality survey already has the quality index in percentage scale for each service, then a Delphi survey is not needed.

### Box 7. Using Delphi survey to estimate quality index.

The table below provides quality scores of a health facility from the quality assessment, as well as the weights obtained from Delphi survey. Assuming relative weights between general indicators and service specific indicators are 0.21 and 0.79 respectively, and the relative weight for drugs and supplies, equipment, and staff are 0.51, 0.33 and 0.16, it is estimated that the quality index is 0.79.

Category	Component	Score (1)	Total available score (2)	Standard score (3)=(1)/(2)	Component Weight (4)	Category weight (5)	Weighted standard score (6)=(3)*(4)*(5)
<b>General</b>	General indicators	34	50	0.68	1.00	0.21	0.143
	Clinical processes	-	-	-	-	-	-
<b>Service-specific</b>	Drugs and supplies	80	100	0.8	0.51	0.79	0.319
	Equipment	45	50	0.9	0.33	0.79	0.236
	Staff with training	35	50	0.7	0.16	0.79	0.090
<b>Quality index</b>							<b>0.788</b>

Source: Authors

111. It is recognized that the weights for each components would be different from countries to countries, and thus it is important to ask similar questions during the Delphi survey. An international Delphi panel can be built up with pooled results from countries where the Delphi survey is conducted to estimate standardized weight for each component.

### CONVERT LIVES SAVED TO QUALITY ADJUST LIFE YEARS (QALYs) OR DISABILITY ADJUSTED LIVED YEARS (DALYs)

112. The LiST tool could produce numbers of lives saved from interventions for both pregnant women and children. The number of lives saved could be used as outcome measures in the cost-effectiveness analysis to estimate the cost per live saved. As the RBF program targets to different populations (children vs. pregnant women), often it is more informative to estimate cost per QALY gained or DALY averted. However, the LiST does not generate the results on QALYs gained or DALYs averted, which have to be calculated manually. One would first estimate DALYs averted or QALYs gained per life saved for distinct population groups (i.e. pregnant women and children under 5) , and then multiply them by the number of lives saved to obtain the total QALYs gained or DALYs averted due to the program.

113. QALYs are calculated by adjusting the years of life gained by quality of life (QoL). QoL is measured on a scale of 0 to 1, where a value of 1 usually represents perfect health, and 0, death. Quality of life is multiplied by the years of life gained to yield a measure that incorporates both quantity and quality. DALYs are calculated by summing the morbidity and mortality averted by an intervention. Averted years of life lost are added to the years of life that would have been spent in disability in the absence of the intervention.

114. DALYs measure the burden of disease and incorporate both the years of life lost (YLL) due to premature mortality and those lost from varying degrees of disability associated with disease (YLD) in a population:

$$\text{DALY} = \text{YLL} + \text{YLD}$$

115. Methods and tools for calculating DALYs are available from the WHO Burden of Disease website, including detailed guidance from the WHO National Burden of Disease Manual (Mathers, Vos, Lopez, Salomon, & Ezzati, 2001).

116. The DALY is based on loss of time in various states of ill health. Time lost due to premature mortality (YLL) is a function of the death rate and the duration of life lost due to a death at each age. Deaths at all ages contribute to the calculation of the burden of disease, and deaths at the same age contribute equally to the burden of disease.

117. The time lost due to a disability (YLD) is estimated from the incidence of disabilities, the average duration of each disability, and a disability weight. The basic formula for calculating YLD is: <sup>5</sup>

$$\text{YLD} = I \times \text{DW} \times L$$

118. where I is the number of incident cases in the reference period, DW is the disability weight (in the range 0-1, 0 means perfect health while 1 means death, the reverse of the scale used in QALYs, and L is the average duration of disability (measured in years).

119. Estimating YLL requires information on total deaths for the condition (by age group and gender) or death rates for the condition per 1,000, and the average age at death for each age group. For estimating YLD, the worksheet requires information on incidence rates of disease/condition, age at onset, duration of the disease or condition, and a disability weight. Total DALYs for the disease/condition is the sum of YLL and YLD. The model also discounts future healthy years of life at 3 percent and the default does not include age weighting, though this could be changed in the model. If the country team does not wish to factor in the disability (morbidity component) of a disease or condition, the measure of effectiveness could focus on the YLL estimation of the model only, which is the case used in this toolkit.

120. Similarly, in order to convert the lives saved to QALYs, the information that is needed is the (1) average age at death, (2) life expectancy at the average of death; (3) quality of life if one survived. With this information, QALYs gained from each saved life could be estimated manually using the formula for fatal cases (Sassi, 2006):

$$\text{QALYs gained} = Q * \frac{1 - e^{-r(L-a)}}{r}$$

or

$$\text{QALYs gained} = \sum_{t=a}^{a+L} \frac{Q_t}{(1+r)^{t-a}}$$

121. Here Q is the average quality of life over if one survived,  $Q_t$  is quality of life at age of  $t$ ,  $e$  is Napier's mathematical constant (2.718...),  $r$  is discount rate, which is often 3%, and  $L$  is life expectancy at the age of death ( $a$ ).

122. In the case of an RBF program, using an estimate of the number of lives saved for both children and pregnant women, we are able to apply the information to estimate the total DALYs averted or QALYs gained due to the RBF program. Taking QALYs as an example, Table 13 shows how QALYs are estimated from saving one pregnant woman's life, assuming that the average age of program women is 25 years old,

---

<sup>5</sup> This is the formula with zero discounting and uniform age weights.

the life expectancy for women at age of 25 is 58 years old, and the quality of life (QoL) from age 25-58 follows the patterns specified in the column of  $Q_t$ . Using the second formula mentioned above, the total QALYs gained is calculated as 18.33 QALYs/life saved. Since the information on  $Q_t$  is not likely to be available in the country, researchers could also conduct a Delphi survey to obtain it or search for literature from comparable countries with available information, or conduct a survey on quality of life from a representative sample to estimate it.

**Table 13. Components of general and service-specific quality indicators**

Period(t-a)	Age(t)	Year interval	$Q_t$	QALYs
0	25	1	0.89	0.89
1	26	1	0.89	0.86
2	27	1	0.89	0.84
3	28	1	0.89	0.81
4	29	1	0.89	0.79
5	30	1	0.84	0.72
6	31	1	0.84	0.70
7	32	1	0.84	0.68
8	33	1	0.84	0.66
9	34	1	0.84	0.64
10	35	1	0.84	0.63
11	36	1	0.84	0.61
12	37	1	0.84	0.59
13	38	1	0.84	0.57
14	39	1	0.84	0.56
15	40	1	0.84	0.54
16	41	1	0.84	0.52
17	42	1	0.84	0.51
18	43	1	0.84	0.49
19	44	1	0.84	0.48
20	45	1	0.84	0.47
21	46	1	0.84	0.45
22	47	1	0.84	0.44
23	48	1	0.84	0.43
24	49	1	0.84	0.41
25	50	1	0.79	0.38
26	51	1	0.79	0.37
27	52	1	0.79	0.36
28	53	1	0.79	0.35
29	54	1	0.79	0.34
30	55	1	0.79	0.33
31	56	1	0.79	0.32
32	57	1	0.79	0.31

Period(t-a)	Age(t)	Year interval	Q <sub>t</sub> *	QALYs
33	58	1	0.79	0.30
<b>Total</b>				<b>18.33</b>

Source: <http://www.avenirhealth.org/software-spectrum>

\*Q<sub>t</sub> is quality of life at time t.

123. Then total QALYs gained for pregnant women (Box 8) are calculated as:

Number of lives saved for pregnant women X QALYs gained per life saved for pregnant women

124. For lives saved for children, LiST estimates the number of lives saved for children under 1 month and children of 1-60 months old. QALYs gained per life saved needs to be calculated separately for each of the two child age groups for the interest of accuracy. Similarly, the formula for estimating total QALYs gained for children is:

Number of lives saved for children X QALYs gained/life saved for children (See Box 9).

### Box 8. Example of QALYs saved for pregnant women.

Using the results from Table 11 where LiST estimates that there are 18 pregnant women saved from improved coverage of institutional delivery due to the RBF program, we calculated that QALYs saved for pregnant women from improved institutional delivery is  $18 \times 18.33 = 329.94$  QALYs. The template for this calculation is available at [www.brandeis.edu/~shepard/QALYs.xlsx](http://www.brandeis.edu/~shepard/QALYs.xlsx)

### Box 9. Example of calculating QALYs per life saved for children under 1 month and children of 1-60 months

Calculation for children under 1 month of age

Average age of children under 1 who died (Years) (a)	0.04
Life expectancy for children at age of 0.04 year old (L)	54.00
Discount rate (r)	3%
Average quality of life if children survived (Q)	0.82
<b>QALYs saved per life saved</b>	<b>21.92</b>

Table 11 shows that there are 250 children under 1 month saved from improved institutional delivery. The total QALYs saved for children under 1 month would be  $250 \times 21.92 = 5480.00$  QALYs.

Average age of children of 1-60 months who died (Years) (a)	2.8
Life expectancy for children at age of 2.8 year old (L)	55.00
Discount rate (r)	3%
Average quality of life if children survived (Q)	0.82
<b>QALYs saved per life saved</b>	<b>21.62</b>

According to Table 11, five more children died in the alternative scenario. QALYs saved from children of 1-60 months would be a negative value, which is  $-5 \times 21.62 = -108.10$  QALYs

125. We then combine QALYs gained from pregnant women and children together to generate the total QALYs gained due to the loss of lives. Combining the results from Boxes 8 and 9, we can estimate the total QALYs gained from improved institutional delivery due to the RBF program is 5701.84, which is calculated as the sum of 329.94, 5480.00, and -108.10.

126. The calculation of DALYs averted per life saved is very similar to that of QALYs (Gold, Stevenson, & Fryback, 2002), with two differences: (1) DALY uses disability weight (DW) rather than quality of life in the calculation; and (2) DALY includes the age weight factor/function in the calculation. Assuming uniform age weights, DALYs averted per life saved could be calculated as:

$$DALYs\ Averted = (1 - DW) * \frac{1 - e^{-r(L-a)}}{r}$$

where DW is the average disability weights if those who die survived the rest of their life. DW could be derived from the global burden disease study implemented by Institute of Health Matrix and Evaluation, using DALYs due to morbidity divided by life years in the target population in the same period.

127. Both QALYs and DALYs are widely used in cost-effectiveness analysis. There are some differences between them. For example, all QALYs have the same value, be it a QALY gained by a young person or by an elderly person. DALYs, as originally defined, gave greater weight per year alive to people in prime ages of 20 to 50, and lesser weight outside those age ranges. However, the version of DALYs used in the 2010 Global Burden of Disease and subsequently treats all ages the same (see [www.healthdata.org/gbd](http://www.healthdata.org/gbd)). The construction of both QALYs and DALYs also assumes a small gain is the same for many people as a large gain for few people. While the absolute levels of DALYS and QALYs differ, changes due to any health intervention are the same in magnitude but opposite in sign. A beneficial intervention would increase the QALYs but decrease the DALYs by the same amount. Neither QALYs nor DALYs directly address income levels directly, but both are tools to address equity. As they count a year of life the same whether enjoyed by a rich or poor person, they count everyone the same. However, if a country does not have accurate estimate of  $Q_t$  and DW and life expectancy at the age of death, estimation of DALYs averted or QALYs gained could be optional.

## CHAPTER 4. PUTTING IT TOGETHER: GENERATING ICERS

129. Based on the cost estimation and outcome measures, such as the number of lives saved, the QALYs gained or DALYs averted, we should be able to generate an incremental cost-effectiveness ratio (ICER).

$$ICER = \frac{\text{Incremental costs of RBF program}}{\text{Number of lives saved, QALYs saved or DALYs averted}}$$

130. Depending on researchers' interests, one could generate the ICER for the overall RBF program, for a particular service, or for a package of services. Box 10 provides an example on calculating ICER for institutional delivery under the RBF program using both the number of lives saved and QALYs gained as measures of outcomes. For conducting a CEA of a package of services, researchers can repeat the same process as that for the institutional delivery, modify multiple parameters in the LiST tool to estimate the combined impact from a package of services, and estimate the ICER at the end.

### Box 10. Example of calculating ICER

In a hypothetical country, assuming the total RBF operational costs is \$10 million over 2012-2014, of which 15% is devoted to institutional delivery, incremental costs for consumables for institutional delivery are estimated at \$3 million, and the coverage of institutional delivery is improved by 15% as shown in Table 10. LiST estimates the number of lives saved of 263 (18+250-5) from institutional delivery, which is shown in Table 11. We have estimated that 5701.84 QALYs are saved, based on the results from Boxes 6 and 7. Then

ICER for institutional delivery could be calculated as

$$ICER_{\text{Institutional delivery}} = \frac{15\% * 10,000,000 + 3,000,000}{263}$$
$$= \$17,110/\text{life}$$

or

$$ICER_{\text{Institutional delivery}} = \frac{15\% * 10,000,000 + 3,000,000}{5701.84}$$
$$= \$789/\text{QALY}$$

Note: This example does not adjust for discounting and annualization for simplicity.

131. To illustrate this process, Boxes 11, 12, and 13 work through the calculation for a hypothetical RBF program for a package of services.



### Box 11. Cost-effectiveness analysis of a hypothetical package of RBF services, part 1

**Description.** A hypothetical country has implemented an RBF program with an experimental design supported by the World Bank from 2011 to 2013. The RBF program operates in 10 districts with a combined population of 1.5 million persons. Comparative data are available from 10 matched control (usual care) districts with a combined population of 1.7 million persons. For illustration, we assume that the country has demographic, epidemiologic, and economic conditions similar to those of Zambia. In 2013, Zambia had a population of 14.53 million and a per capita GDP of \$1845.

**Cost analysis.** Data from the country's national pharmaceutical service shows that medications and supplies with a value of \$15 million (\$8 million in 2012 and \$7 million in 2013) were shipped to district offices and health facilities in the RBF districts, compared to \$10 million (\$5 million in 2012 and 2013 respectively) in the control districts. In addition, RBF facilities made local purchases (funded through user fees and government funding) of \$2 million (1 million in 2012 and 2013 respectively) compared to \$1 million (0.5 million in 2012 and 2013 respectively) in control facilities. The RBF districts received \$4 million (\$1.5 million and 2.5 million in 2012 and 2013 respectively) in incentive payments, while the control districts received no such payments. An in-country program office provided or arranged training, program management, service verification, and administration of incentive payments at a cost of \$0.8 million per year. In addition, the World Bank spent \$0.3 million (\$0.2 million and 0.1 million in 2012 and 2013 respectively) on staff time and travel to oversee the experimental area. In addition, \$250,000 worth of computers was purchased for an RBF intervention, and computers have an average life of 5 years.

Note: The example does not adjust for discounting, annualization for simplicity.

### Box 12. Cost-effectiveness analysis of a hypothetical package of RBF services, part 2

To analyze costs, we compute the costs per capita in each arm as follows:

Item	RBF			Control		
	Year 0	Year 1	PV	Year 0	Year 1	PV
Medications and supplies (central)	\$8.00	\$7.00	\$14.80	\$5.00	\$5.00	\$9.85
Medications and supplies (local)	\$1.00	\$1.00	\$1.97	\$0.50	\$0.50	\$0.99
Incentive payments	\$1.50	\$2.50	\$3.93	\$0.00	\$0.00	\$0.00
In-country office	\$0.80	\$0.80	\$1.58	\$0.00	\$0.00	\$0.00
World Bank staff	\$0.20	\$0.10	\$0.30	\$0.00	\$0.00	\$0.00
Computers	\$0.055	\$0.055	<u>\$0.11</u>	\$0.00	\$0.00	<u>\$0.00</u>
Total			\$22.68			\$10.84
Population (million)			1.5			1.7
Per capita			\$15.12			\$6.38

These calculations show that the incremental cost per capita is \$8.74 (i.e., \$15.12 - \$6.38) over 2 years.

### **Box 13. Cost-effectiveness analysis of a hypothetical package of RBF services, part 3**

Effectiveness. The evaluation conducted baseline and end line household surveys in 2011 and 2013, which showed greater improvements in the RBF area compared to the control area on numerous indicators. These improvements in the quantity and quality of services led to the projected saving, when applied to the intervention area of 1.5 million people saved the lives of 45 pregnant women and 450 children. Each maternal death averted generates 18.33 QALYs. Each child under 5 saved (for example, from vaccination against childhood diseases) generates 21.62 QALYs. These results generate 824.9 QALYs in pregnant women and 9,729.0 in children for a total of 10,553.9 gained in the intervention area population over 2 years. To make this result consistent with the cost data, we convert to a per capita basis, getting 0.007035 QALYs per capita. This was calculated as 10,553.9 divided by 1,500,000.

Cost-effectiveness. Dividing the cost by effectiveness we have  $\$8.74 / 0.007035$  equals \$1,242 per QALY gained through the RBF program. To interpret this, we note that the cost-effectiveness ratio is substantially below the per capita GDP of Zambia (\$1,845). Thus, RBF is highly cost-effective in this context. In using these results to think about the future evolution of this hypothetical RBF program, we note that vaccinations are already doing very well and have little scope for further improvement. On the other hand, there is substantial scope for further increase in institutional deliveries, intermittent preventive treatment, and quality. Future RBF programs may wish to incentivize those indicators even more and reduce the incentive for vaccinations.

132. Please note that if one would like to examine the ICER of institutional delivery, then both effectiveness and cost measures should be limited to institutional delivery. For the effectiveness, users are only allowed to change the parameters related to institutional delivery in the LiST tool, such as baseline and end line coverage of institutional delivery. Correspondingly, the incremental costs should also be estimated for institutional delivery only. The allocation of costs for a particular service and a package of services were mentioned above.

133. The ICER can be calculated using the aggregate amount of costs and aggregate effectiveness over the study period as shown in Box 10, and it could also be calculated on an annual basis, using annual costs and annual effectiveness. As each program goes through a maturation process, we recommend using average annual costs and average annual effectiveness in the CEA and conducting sensitivity analysis.

### **SENSITIVITY ANALYSIS IN CEA**

134. Uncertainty in a CEA can arise from many different sources, including areas of methodological disagreement, parameter uncertainties (for example, unknown parameters, disagreements about appropriate values, uncertainty surrounding the estimation process, data from a specific population, sampling variability), and modeling. The objective of the sensitivity analysis is to test the robustness of the base-case results against the most important assumptions. There are different ways to perform sensitivity analysis, but the most commonly used for CEA is to investigate how the results change as a result of changing the value of each important assumption while holding all other aspects of the analysis constant. A second alternative investigates whether best- and worst-case scenarios of assumptions change the conclusions of the analysis significantly. Other alternatives involve statistical analysis of testing the effect the distributions of key parameters have on the distribution of results (i.e., Monte Carlo analysis). Table 14 highlights some of the approaches that can be taken in the analysis.

**Table 14: Selected Types of Sensitivity Testing Useful for Cost-effectiveness Analysis**

Type of Sensitivity Analysis	Question	Description
<b>One-way sensitivity analysis/ Partial sensitivity analysis</b>	<i>Does the cost-effectiveness of an intervention change when a single assumption is varied, while holding all others constant?</i>	One-way sensitivity analyses can be single or multiple and are conducted by varying assumptions – generally the most important or uncertain ones – one at a time. The limitation of this method is that it is not possible to observe how net benefits change as a result of modifying more than one variable at the same time. A tornado diagram visualizes the varying effects of different parameters.
<b>Threshold</b>	<i>Is there a particular value an assumption can have where none of the alternatives is favored?</i>	A threshold analysis varies the size of an assumption over a range to determine the ‘threshold’ point where none of the intervention alternatives is favored over others. This analysis is meaningful when the value of the parameter is known, but difficult to interpret if variables used are dependent on each other (for example, a graph of \$ per DALY averted vs. vaccine efficacy against mortality in children (%) with several vaccine dose costs plotted).
<b>Two-way sensitivity analysis</b>	<i>How do net benefits change as we vary two assumptions while holding all others constant?</i>	Two-way sensitivity analyses are similar to one-way sensitivity analyses, except two critical parameters are simultaneously varied.
<b>Worst- and best-case scenario</b>	<i>Does any combination of reasonable assumptions change the overall conclusion of the cost-effectiveness analysis?</i>	In this approach, the base-case scenario (or most likely scenario) is compared with both the best-case scenario (the assumptions which would collectively produce the highest estimate of net benefits), and the worst-case scenario (the assumptions which would collectively produce the lowest estimate for net benefits). This method may not be appropriate if there is a nonlinear relationship between net benefits and a given explanatory variable.

Adapted from (Boardman, Greenberg, Vining, & Weimer, 2011).

135. For cost-effectiveness of RBF mechanisms, critical factors to consider in a sensitivity analysis include assumptions regarding:

- Assumptions on costs: If a country has a good expenditure tracking system, economic costs may be possible and CEA using economic costs could be examined.
- Utilization of services: The variation of results of the impact evaluation can also be estimated to understand the uncertainty.
- Effectiveness of interventions, measured in reduction of mortality, from the utilization of the various maternal-child health services and quality of care performance: We assumed the effectiveness used in the list as efficacy of the intervention. This assumption could be changed.
- Epidemiological parameters (case fatality rates, incidence rates, etc.): the variation of epidemiological parameters should be investigated if the sample size for obtaining this parameters is small.

- Time horizon and maturation of RBF program: CEA could be conducted any stage of the RBF program, and the results could vary. Often it is less cost effective when a program at the start-up stage, and more cost-effective when it is matured.
- Inflation: 3 percent is often used for deriving discount factor for a CEA. In countries with much higher inflation rate, discount factor should also be changed to estimate present value of future monetary resources.

136. Country teams are encouraged to focus on variability and uncertainty in key parameters in the cost-effectiveness analysis. The results of a sensitivity analysis can provide some indication of how robust initial conclusions of cost-effectiveness are. If the cost-effectiveness results do not change very much, it can generally be concluded that results are robust. On the other hand, if small changes in the parameters for estimating costs and/or benefits significantly alter the results ( $\pm 10\%$ ), then it would be important to highlight a range of findings and discuss the implications for policy-making.

137. Regarding the sensitivity analysis on the effectiveness measures, users could modify parameters used in the LiST tool to generate a new set of results, and repeat the analysis multiple times, generating corresponding results to assess the robustness of the model.

138. Ideally, if one knows the distribution of parameters, Monte Carlo simulation could be conducted to generate confidence interval to estimate the uncertainty of the estimate, i.e. ICER. However, as the effectiveness in this toolkit primarily dependent on LiST tool, Monte Carlo simulation that repeats 1,000 more times may not be feasible as each run requires modifying parameters manually. A simulation of 20-50 repetitions is possible and could well serve the purpose.

## CHAPTER 5. DISCUSSION AND INTERPRETATION

### INTERPRETATION OF IMPACT EVALUATION AND CEA RESULTS

139. The results of ICER will be used to compare interventions. Cost-effectiveness ratios of alternative interventions may be relatively similar in value; when this occurs, neither intervention is dominant over the other. If the ICER is negative and the change in effectiveness is favorable, then the intervention dominates the alternative. If the ICER is positive and the change in effectiveness is favorable, then the intervention is dominated by the alternative. Ratios that have an order-of-magnitude difference between them demonstrate more clearly which intervention is most cost-effective and deserves priority over the alternative. The ICER will have essentially the same numerical value whether the calculations are based on DALYs or QALYs.

140. Threshold values of ICERs recommended by WHO CHOICE (Choosing Interventions that are Cost-Effective) project are categorized into three categories of cost-effectiveness: highly cost-effective (ICER less than GDP per capita); cost-effective (ICER between one and three times GDP per capita); and not cost-effective (ICER more than three times GDP per capita) (World Health Organization, 2015), (World Health Organization, 2011). ICER ratios have been calculated similarly for other MCH interventions, such as demand side voucher schemes. Both supply side and demand side RBF (PBF and voucher programs) are widely implemented in developing countries. Comparing ICER between the two types of the program, in conjunction with joint program, would better inform policy makers on where to best invest in health.

### LIMITATIONS OF THE TOOLKIT

141. This toolkit applies CEA to an RBF program and bears general limitations of the CEA for decision making in the health sector. First, CEA measures costs and effects at one point in time and at one particular level of scale. Costs and outcomes will likely change as coverage increases or declines. As a result, the relative cost-effectiveness of interventions will fluctuate over time. For this reason, it may be useful to plan to conduct CEA at more than one point in time.

142. Second, the estimation of effectiveness in this Toolkit is based on LiST. The LiST tool covers maternal-child health interventions, HIV/AIDS services, and family planning. Thus, the proposed approach should provide good estimates of the lives saved through these interventions and reasonable estimates of the morbidity averted through these interventions. It does not, however, provide estimates of lives saved from improving curative care for adults. Other health simulation models would need to be developed or adapted for this purpose.

143. Third, CEA can be demanding in terms of data requirements and analysis, and much of the analysis is based on assumptions. Sensitivity testing can provide a better picture of the possible range of cost-effectiveness results for a particular intervention. Unless specifically built into the study design, the results of CEA do not tell policy makers about the distributional impacts of a health strategy and whether benefits accrue to certain population groups, such as the poor, more than others. For this reason, it is useful to complement CEA with other types of evaluations that examine these concerns more closely. In addition, some experts argue that health effects should not be discounted as heavily as costs. However, at the extreme, this may produce a bias in favor of projects with very short time spans.

144. Fourth, although CEA is an important tool for policy making and allocating resources, it is not the only tool. Decision making must incorporate other considerations, such as priorities of populations and political commitment. For example, it may cost much more to deliver services to poor populations living in remote rural areas than to an urban population, but equity and political considerations would still require that the rural populations be served equally.

145. Besides the general limitations, there are two specific limitations pertained to this toolkit: (1) this toolkit, using the LiST tool to quantify health outcomes, only measures the effectiveness of interventions on mortality reduction, and is not able to examine their impact on improving quality of life and reducing the

duration of illness; and (2) although the toolkit is very comprehensive for addressing maternal child health services, it is not able to include services for the general population in the modelling.

#### **STRENGTHS OF THE TOOLKIT**

146. This is the first toolkit to examine the cost-effectiveness of RBF programs. Unlike many evaluation studies which only focus on outputs of RBF programs, this study incorporates both costs of RBF programs services and outcomes. Results are translated to outcomes, a more important indicator than outputs. Using standardized effectiveness measures in the cost-effectiveness analysis allows policy makers to compare RBF programs to other programs whose primary objective is to improve maternal and child health.

147. This toolkit uses a widely applied tool, LiST, to evaluate the effectiveness, which enhances the robustness of evaluation results. RBF programs provide incentives to providers on a variety of services, and this creates challenges to use one model to conduct a comprehensive evaluation of their impact. LiST, in conjunction with FamPlan, DemProj, and AIDS impact model (AIM), is able to model a majority of maternal and child health services included in RBF programs.

148. This toolkit first includes quality of care, an important component of RBF programs, in the modelling. A significant proportion of incentives are used to improve the quality of care. However, due to challenges of converting quality of care to health outcomes, quality of care is often disregarded in the effectiveness evaluation. In this toolkit, we proposed to use a Delphi survey to estimate the impact of quality of care on health outcomes to address the modelling challenges.

#### **RESOURCES FOR CONDUCTING A CEA OF A RBF PROGRAM**

149. On the top of the impact of evaluation, which involved in household and health facility survey, additional resources are needed to conduct the CEA. Creating a core technical group of 4-6 members responsible for organizing and implementing the cost analysis exercise(s) is essential. The composition of the multi-disciplinary team will depend upon the training and availability of individuals. As a whole, the team needs to have members who have economic and statistical expertise and some experience with survey analysis. The team must include members who will be able to determine the quality of information collected and whether alternative sources need to be pursued. Some of the team members should be aware of and have access to data sources, such as hospital records, or Ministry of Health and Finance documents.

150. Although it may be difficult for officials from the Ministry of Health, Ministry of Finance, or Ministry of Planning to participate in the analysis, their commitment and involvement can help ensure that the results of the exercise(s) are linked to the overall decision-making process. Local consultants, university groups, and other technical experts can conduct field research and provide valuable advice to the economic analysis country team.

151. The core team should include a lead technical person who will be responsible for taking the work forward in-country. The lead technical person should have demonstrated skills in organizing data collection and analysis efforts, as well as have training in finance, economics, or related field.

#### **IMPLICATIONS AND CONCLUSIONS**

152. Study teams are encouraged to develop a succinct presentation of the approach taken, main results, and the implications for policy decision making for busy decision makers. Study teams should remember that the results of cost- and cost-effectiveness analysis need to be interpreted within the particular country context, and that factors such as level of scale, technologies, and prices will affect cost-effectiveness results. In addition, cost-effectiveness results do not provide information on the distributional aspects of RBF.

153. Country team members should be familiar with the methods presented in this toolkit prior to embarking on data collection and analysis. A review of basic data requirements and sources for estimating costs and outcomes should be conducted initially.

154. The methods discussed in this toolkit are merely tools for providing additional information for policy and decision making regarding whether to introduce, scale-up, or alter an RBF intervention. We recommend that the analysis be presented and discussed with a range of stakeholders who may affect decisions in the country, and to involve them at various stages of the analysis, particularly when key decisions are being taken on parameter estimates.

155. All of the generic data collection instruments need to be adapted and pre-tested within each country prior to use, so that the most relevant cost and outcome information is collected. Country teams also should assess the data available from secondary information (administrative, impact evaluation data collection, etc.) early in the process in order to identify whether any prospectively designed primary data collection tools are required.

156. Depending upon the policy questions to be addressed, country teams should strive to produce the following indicators in order to facilitate comparison across countries under the HRITF:

- Total annual cost of the RBF intervention (for the sample of facilities in the RBF area)
- Cost/capita of the RBF intervention (facility and aggregated)
- Total annual RBF cost by line item, by functional classification, and by level of the health system as relevant to each country's RBF program.
- Double-difference estimator of total costs between control and RBF areas
- Incremental cost-effectiveness between control and RBF areas
- The cost-effectiveness analysis should focus on evaluating differences in costs and differences in outcomes between control and treatment areas, or in the treatment area with and without the RBF intervention. An ICER should be estimated. At the facility level, a double-difference analysis can be conducted. The module recommends use of either DALY or QALY as the outcome measure for the cost-effectiveness study, but other outcome measures can be used instead, such as the number of lives saved. Costs should be collected and reported in local currency units for the year of analysis and then converted into current US dollars to facilitate comparison among countries.
- The objective of the sensitivity analysis is to test the robustness of the base-case results against the most important assumptions. There are different ways to perform a sensitivity analysis, but the most commonly used for CEA is to investigate how the results change as a result of changing the value of each important assumption while holding all other aspects of the analysis constant.

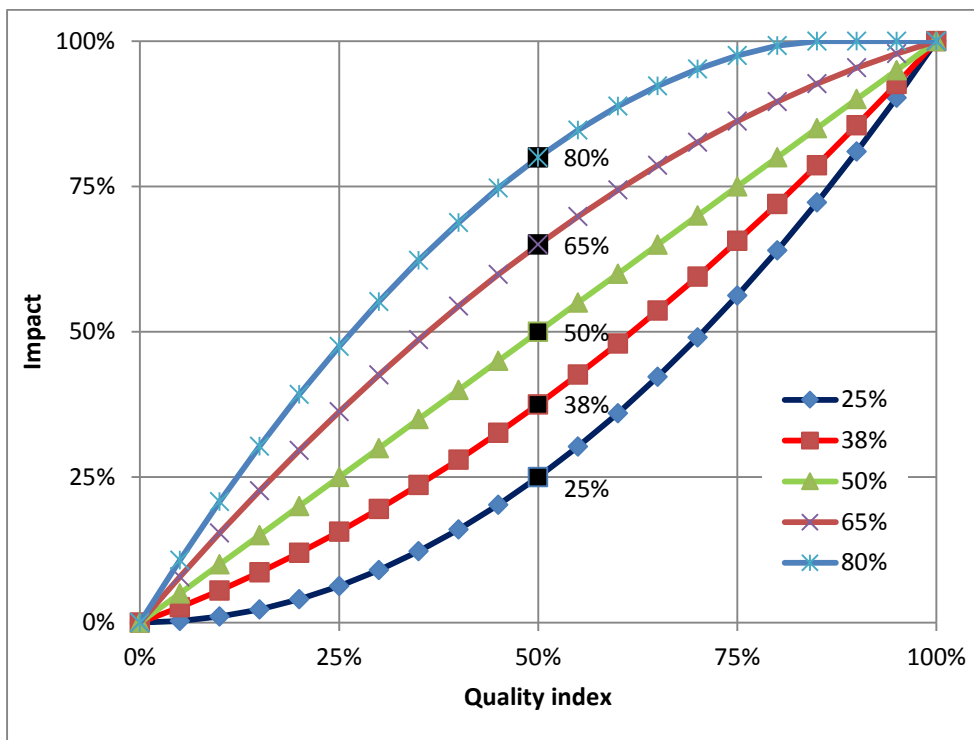
157. In conclusion, while this toolkit has been oriented specifically around RBF, the concepts of incremental cost per capita, improvements in coverage, improvements in quality, incremental health benefits per capita, the ICER are all concepts that apply to many health interventions. The technique of combining actual coverage with quality to derive adjusted coverage is a feature that may be helpful in combining improvements in both coverage and quality into other health economics evaluations.

# APPENDIXES

## APPENDIX 1. DELPHI QUESTIONNAIRE FOR EVALUATING THE IMPACT OF QUALITY OF CARE

**Purpose:** Results based financing (RBF) has been implemented for more than 2 years. The design of the RBF program aims to enhance both quantity and quality of health care for targeted services, particularly for maternal and child health services, including prenatal care, postnatal care, institutional delivery, vaccination, family planning, and curative care in health facilities. However, the quality of care on health outcomes (impact) is challenging to quantify. We would like to have your opinions to help quantify the relationship between them. Potential relationship between quality and impact could be represented below, such as 50% of quality achieves only 25% of the impact on health outcome. For each service, use your best judgment to estimate the potential impact of a compromised quality.

Figure A.1.1: Quantify quality of care on health outcomes



Source: <http://www.avenirhealth.org/software-spectrum>



**Individual background**

Organization \_\_\_\_\_; Position \_\_\_\_\_;

Expertise \_\_\_\_\_; Years of experience \_\_\_\_\_ (Years)

Having clinical or epidemiology background:      Yes      No

**Questions: Please estimate what share of the potential impact of that intervention would be achieved for each intervention, if the quality score were 50 percent.**

1. For adult curative care (malaria, upper respiratory infection, and diarrhea), if the quality score is 50 percent, what is the share of the potential impact that the curative care treatment would be achieved? (Note that answer must be between 0 and 100 percent)
  
2. For child curative care? \_\_\_\_\_
3. For family planning? \_\_\_\_\_
4. For vaccination? \_\_\_\_\_
5. For institutional delivery? \_\_\_\_\_
6. For prenatal care? \_\_\_\_\_
7. For postnatal care? \_\_\_\_\_
8. For treatment of HIV+ pregnant women? \_\_\_\_\_
9. For pregnant women HIV/AIDS counseling? \_\_\_\_\_

## APPENDIX 2. DELPHI QUESTIONNAIRE FOR QUANTIFYING RELATIVE IMPORTANCE OF GENERIC VS SERVICE-SPECIFIC QUALITY INDICATORS

Factors contributing to quality of care and health outcomes potentially include not only the service specific factors but also other factors in the facility, such as autonomy, leadership and management, infrastructure, and supportive supervision and technical support from higher level.

Question: please assign a value between 0 and 100 for disease specific versus general factors that reflect the importance in overall quality of care for that specific disease. The sum of the two should be 100 percent.

Curative care	Weight (sum=100%)
Service specific (clinical processes, drugs and supplies, equipment, staff)	_____
General (autonomy, technical support & supervision, HRH, infrastructure)	_____
Family planning	Weight (sum=100%)
Service specific (clinical processes, drugs and supplies, equipment, staff)	_____
General (autonomy, technical support & supervision, HRH, infrastructure)	_____
Vaccination	Weight (sum=100%)
Service specific (clinical processes, drugs and supplies, equipment, staff)	_____
General (autonomy, technical support & supervision, HRH, infrastructure)	_____
Institutional delivery	Weight (sum=100%)
Service specific (clinical processes, drugs and supplies, equipment, staff)	_____
General (autonomy, technical support & supervision, HRH, infrastructure)	_____
Prenatal care	Weight (sum=100%)
Service specific (clinical processes, drugs and supplies, equipment, staff)	_____
General (autonomy, technical support & supervision, HRH, infrastructure)	_____
Post natal care	Weight (sum=100%)

Service specific (clinical processes, drugs and supplies, equipment, staff) \_\_\_\_\_  
General (autonomy, technical support & supervision, HRH, infrastructure) \_\_\_\_\_

---

HIV VCT and PMTCT Weight (sum=100%)

Service specific (clinical processes, drugs and supplies, equipment, staff) \_\_\_\_\_  
General (autonomy, technical support & supervision, HRH, infrastructure) \_\_\_\_\_

---

Malaria treatment Weight (sum=100%)

Service specific (clinical processes, drugs and supplies, equipment, staff) \_\_\_\_\_  
General (autonomy, technical support & supervision, HRH, infrastructure) \_\_\_\_\_

Individual background:

Organization \_\_\_\_\_; Position: \_\_\_\_\_

Expertise: \_\_\_\_\_; years of experience: \_\_\_\_\_

Having clinical or epidemiological background:            Yes                            No

### APPENDIX 3. DELPHI QUESTIONNAIRE FOR QUANTIFYING RELATIVE IMPORTANCE FOR COMPONENTS WITHIN THE SERVICE-SPECIFIC QUALITY INDICATOR

**Purpose:** Results based financing (RBF) has been implemented for more than 2 years. The design of the RBF program aims to enhance both quantity and quality of health care for targeted services, particularly for maternal and child health services, including prenatal care, postnatal care, institutional delivery, vaccination, family planning, and curative care in health facilities. The quality of care is measured with general quality and service-specific quality. The potential dimensions of the each type of the quality include:

**Table A.3.1: Components of general and service-specific quality indicators**

General quality indicators	Service-specific quality indicators
Infrastructure	Clinical processes
Administration and management	Drugs and supplies
Human resource for health	Equipment
HMIS	Staff with training
Leadership and autonomy	
General equipment	

Source: <http://www.avenirhealth.org/software-spectrum>

Within the service-specific quality, the relative importance of each component varies, depending on which service is evaluated. As an illustration, vaccination does not need high technical skills, and the supply of vaccine is an important factor for a successful vaccination program. In this case, a higher weight will be given to the component of “Drugs and supplies” while the “clinical processes” received a smaller weight. We would like the experts help estimate the relative weight among the components within the service-specific quality for eight indicators. We will provide components and their measures within each service. Please use your best judgment to determine the relative importance of each component. Thank you!

**Individual background**

Organization \_\_\_\_\_; Position \_\_\_\_\_;

Expertise \_\_\_\_\_; Years of experience \_\_\_\_\_ (Years)

Having clinical or epidemiology background:      Yes      No

**Questions: Please assign a value between 0 to 100 to each component in the tables below. Please note the sum of the total value of the all the components should be 100.**

<b>Curative care</b>	<b>Weight (Total 100)</b>
<b>CLINICAL PROCESSES</b>	_____
Vignette for child diarrhea, fever, cough, Measure weight, height, and temperature Prescribe medication, counselling	
<b>DRUGS AND SUPPLIES</b>	_____
Tetracycline ophthalmic ointment  Paracetamol (Panadol) tabs  Amoxicillin (tabs or capsule) Amoxicillin (syrup) Oral Rehydration Solution (ORS) packets Cotrimoxazole	
<b>EQUIPMENT</b>	_____
Microscope Centrifuge Hemoglobinometer Refrigerator for storing reagents	
<b>STAFF</b>	_____
Staff received recent training	

<b>Family planning</b>	
<b>DRUGS AND SUPPLIES</b>	_____
Condoms (male or female) Oral contraceptive tablets Depot Medroxyprogesterone Acetate (DMPA) Implant jadelle Intrauterine Device (IUD)	
<b>STAFF</b>	_____
Staff received recent training	

<b>Vaccination</b>	<b>Weight (Total 100)</b>
<b>DRUGS AND SUPPLIES</b>	
Bacille Calmette-Guérin (BCG)	
Oral Polio Vaccine (OPV)	
Tetanus Toxoid (TT)	
Dyphtheria Tetanus Pertussis (DTP)	
Hepatitis B Vaccine (HBV) Tetravalent	
Measles vaccine	
HiB vaccine	
Pentavalent (DPT, Hepatitis B, Hemophilus influenzae B)	
<b>EQUIPMENT</b>	
Main vaccine thermometer	
Cold box / Vaccine carrier	
Ice packs	
Refrigerator	
<b>STAFF</b>	
Staff received recent training	

Institutional Delivery	Weight (Total 100)
<b>CLINICAL PROCESSES</b>	_____
Vignette for prolonged labor	
<b>DRUGS AND SUPPLIES</b>	_____
Magnesium Sulfate	
Diazepam Injection	
Misoprostol	
Oxytocin	
<b>EQUIPMENT</b>	_____
Delivery table/bed	
Delivery light	
Resuscitation bag, newborn	
Eye drops or ointment for newborn	
Intravenous fluids	
Vacuum extractor	
Vaginal retractor	
Bag Valve Mask (Ambu bag)	
Guedel airways-neonatal, child, and adult	
Uterine dilator	
Needles	
<b>STAFF</b>	_____
Staff received recent training	

Prenatal care and postnatal care	Weight (Total 100)
<b>CLINICAL PROCESSES</b>	
Vignette for ANC	
Iron or folate routinely prescribed for ANC mothers	
Reported having ITP for malaria	
Measuring weight, height, blood pressure, pulse,	
Check for anemia, check fetal heart,	
Counseling about warning sign, HIV, FP, etc.	
Time of first ANC	
Procedure done during a ANC visit	
Time of PNC	
Received iron supplement, vit A	
<b>DRUGS AND SUPPLIES</b>	
Pregnancy testing kit	
Rapid plasma reagent (RPR) test for syphilis	
Urine testing kit	
Folic acid tabs	
Vitamin A	
Pregnancy testing	
Iron tabs (with or without folic acid)	
<b>STAFF</b>	
Staff received recent training	

Postnatal care	Weight (Total 100)
<b>CLINICAL PROCESSES</b>	
Vignette for ANC	
Iron or folate routinely prescribed for ANC mothers	
Reported having ITP for malaria	
Measuring weight, height, blood pressure, pulse,	
Check for anemia, check fetal heart,	
Counseling about warning sign, HIV, FP, etc.	
Time of first ANC	
Procedure done during a ANC visit	



Time of PNC

Received iron supplement, vit A

**DRUGS AND SUPPLIES**

---

Pregnancy testing kit

**RAPID PLASMA REAGENT (RPR) TEST FOR SYPHILIS**

Urine testing kit

Folic acid tabs

Vitamin A

Pregnancy testing

Iron tabs (with or without folic acid)

**STAFF**

---

Staff received recent training

---

**HIV counseling and testing and Treatment of  
HIV+ pregnant women**

**Weight (total 100)**

---

**DRUGS AND SUPPLIES**

---

HIV test kit

**STAFF**

---

Staff received recent training

---

**Malaria treatment**

**Weight (Total 100)**

---

**CLINICAL PROCESSES**

---

Vignette for child diarrhea, fever, cough,

**DRUGS AND SUPPLIES**

---

Chloroquine

Quinine

Fansidar / Sulphadoxine-Pyrimethamine (SP)

Artemisinin-Based Combination Therapy ACT  
(fansidar + artesunate) / Coartem

Malaria rapid diagnostic kits

**STAFF**

---

Staff received recent training

---

## REFERENCES

- Avenir Health. 2015. *Spectrum Manual: Spectrum System of Policy Models*. Glastonbury, CT: Avenir Health.
- Basinga, P., Mayaka, S., and Condo, J. 2011. "Performance-based Financing: The Need for More Research." *Bulletin World Health Organization*, 89(9), 698-699. doi: 10.2471/BLT.11.089912
- Bautista-Arredondo, S., Dmytraczenko, T., Kombe, G., and Bertozzi, S. M. 2008. "Costing of Scaling up HIV/AIDS Treatment in Mexico." *Salud Publica Mex*, 50(Suppl 4S), :S437-444.
- Boardman, A., Greenberg, D., Vining, A., and Weimer, D. 2011. *Cost-Benefit Analysis: Concepts and Practice, 4th ed.* Boston: Prentice Hall.
- Borghini, J., Little, R., Binyaruka, P., Patouillard, E., and Kuwawenaruwa, A. 2015. "In Tanzania, the Many Costs of Pay-for-performance Leave Open to Debate Whether the Strategy is Cost-effective." *Health Aff (Millwood)*, 34(3), 406-414. doi: 10.1377/hlthaff.2014.0608
- Boschi-Pinto, C., Young, M., and Black, R. E. 2010. "The Child Health Epidemiology Reference Group Reviews of the Effectiveness of Interventions to Reduce Maternal, Neonatal and Child Mortality." *Int J Epidemiol*, 39 Suppl 1, i3-6. doi: 10.1093/ije/dyq018
- Falisse, J. B., Ndayishimiye, J., Kamenyero, V., and Bossuyt, M. 2014. "Performance-based Financing in the Context of Selective Free Health-care: An Evaluation of its Effects on the Use of Primary Health-care Services in Burundi Using Routine Data." *Health Policy Plan*. doi: 10.1093/heapol/czu132
- Gold, M. R., Stevenson, D., and Fryback, D. G. 2002. "HALYS and QALYS and DALYS, Oh My: Similarities and Differences in Summary Measures of Population Health." *Annu Rev Public Health*, 23, 115-134. doi: 10.1146/annurev.publhealth.23.100901.140513
- Hutubessy, R., Chisholm, D., and Edejer, T. T. 2003. "Generalized Cost-effectiveness Analysis for National-level Priority-setting in the Health Sector." *Cost Eff Resour Alloc*, 1(1), 8. doi: 10.1186/1478-7547-1-8
- Jamison, D. T., Breman, J. G., Measham, A. R., Alleyne, G., Claeson, M., Evans, D. B., . . . Musgrove, P. 2006. *Disease Control Priorities in Developing Countries* (2nd ed.). New York, NY: Oxford University Press.
- Janssen, W., Ngirabega Jde, D., Matungwa, M., and Van Bastelaere, S. 2015. "Improving Quality Through Performance-based Financing in District Hospitals in Rwanda Between 2006 and 2010: A 5-year Experience." *Trop Doct*, 45(1), 27-35. doi: 10.1177/0049475514554481
- Mathers, C. D., Vos, T., Lopez, A. D., Salomon, J., and Ezzati, M. 2001. *National Burden of Disease Studies: A Practical Guide*. Geneva: World Health Organization.
- Ozaltin, A., and Cashin, C. 2014. *Costing of Health Services for Provider Payment: A Practical Manual based on Country Costing Challenges, Trade-offs, and Solutions*. Washington DC: Joint Learning Network for Universal Health Coverage.
- Sassi, F. 2006. "Calculating QALYs, Comparing QALY and DALY Calculations." *Health Policy Plan*, 21 (5), 402-408. doi: 10.1093/heapol/czl018
- Shepard, D. S., Hodgkin, D., and Anthony, Y. E. 2000. *Analysis of Hospital Costs: A Manual for Managers*. Geneva, Switzerland: World Health Organization.
- Shepard, D. S., Sanoh, L., and Coffi, E. 1986. "Cost-effectiveness of the Expanded Programme on Immunization in the Ivory Coast: A Preliminary Assessment." *Social Science and Medicine*, 22(3), 369-377.
- Soeters, R., Peerenboom, P. B., Mushagalusa, P., and Kimanuka, C. 2011. "Performance-based Financing Experiment Improved Health Care in the Democratic Republic of Congo." *Health Aff (Millwood)*, 30(8), 1518-1527. doi: 10.1377/hlthaff.2009.0019
- Stenberg, K., Axelson, H., Sheehan, P., Anderson, I., Gulmezoglu, A. M., Temmerman, M. 2014. "Advancing Social and Economic Development by Investing in Women's and children's Health: A New Global Investment Framework." *Lancet*, 383(9925), 1333-1354. doi: 10.1016/S0140-6736(13)62231-X
- World Health Organization. 2003. *WHO Guide to Cost-effectiveness Analysis*. Geneva: WHO.
- \_\_\_\_\_. 2008. *The Global Burden of Disease: 2004 Update*. Geneva: WHO. Retrieved 14 Apr 2015, from [http://www.who.int/healthinfo/global\\_burden\\_disease/GBD\\_report\\_2004update\\_full.pdf](http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf)
- \_\_\_\_\_. 2011. *Commission on Macroeconomics in Health*. Geneva: World Health Organization.

- \_\_\_\_\_. 2015. Cost-effectiveness Thresholds. Retrieved April 3, 2015, from [http://www.who.int/choice/costs/CER\\_thresholds/en/](http://www.who.int/choice/costs/CER_thresholds/en/)
- Zeng, W., Cros, M., Wright, K. D., and Shepard, D. S. 2013. "Impact of Performance-based Financing on Primary Health Care Services in Haiti." *Health Policy Plan*, 28(6), 596-605. doi: 10.1093/heapol/czs099





Results-Based Financing (RBF), which rewards providers, users, or administrators of services upon achieving a set of verified results, has been gaining attraction in global health as a prominent approach to gain value for money. With a large number of countries adopting RBF in the recent years, evidence starts to emerge which points to the effectiveness of RBF in improving coverage and quality of important services, such as maternal and child health and reproductive health. A question remains, however, if RBF is more cost-effective than alternative interventions which aim to improve similar outcomes. Given the limited resources that countries have for health, the answer to this question is particularly important as it guides policy makers on which programs to invest to maximize health benefits for the population. A second question is how the cost-effectiveness varies across different settings and RBF program features.

The current toolkit aims to support country programs to assess the cost-effectiveness of RBF interventions and to facilitate cross-country comparisons of RBF programs. The toolkit is specifically tailored to supply-side RBF but its general principles apply to most health systems interventions directed at the health-related Millennium Development Goals (MDGs)— 4 and 5. The development of the toolkit was based on actual experience of conducting a costeffectiveness analysis of Zambia's RBF program and an extensive review of RBF programs features across the Health Results Innovation Trust Fund (HRITF) portfolio. Given that RBF programs in the HRITF portfolio are typically complex health system interventions, the toolkit recommends a practical approach of adopting a program implementer's perspective and presents different options for cost-effectiveness analysis (for selected key indicators or for an entire package of services). It also provides guidance on incorporating quality of care, which is strongly emphasized across many RBF programs.

## ABOUT THIS SERIES:

This series is produced by the Health, Nutrition, and Population Global Practice of the World Bank. The papers in this series aim to provide a vehicle for publishing preliminary results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character. For free copies of papers in this series please contact the individual author/s whose name appears on the paper. Enquiries about the series and submissions should be made directly to the Editor Martin Lutalo (mlutalo@worldbank.org) or HNP Advisory Service (healthpop@worldbank.org, tel 202 473-2256).

For more information, see also [www.worldbank.org/hnppublications](http://www.worldbank.org/hnppublications).



1818 H Street, NW  
Washington, DC USA 20433

Telephone: 202 473 1000  
Facsimile: 202 477 6391  
Internet: [www.worldbank.org](http://www.worldbank.org)  
E-mail: [feedback@worldbank.org](mailto:feedback@worldbank.org)