



A Guide to the Fundamentals of Economic Evaluation in Public Health

April 2019



A Guide to the Fundamentals of Economic Evaluation in Public Health

Scott Moreland, PhD, MEASURE Evaluation, Palladium

Shaylen Foley, MPH, MEASURE Evaluation, Palladium

Lauren Morris, PhD, MEASURE Evaluation, Palladium

April 2019

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. MS-19-162
ISBN: 978-1-64232-135-7



ACKNOWLEDGMENTS

This document has benefited from the input and suggestions of several people. First, we acknowledge the early contributions of John Fieno (formerly with MEASURE Evaluation, Palladium) to the chapter on cost-effectiveness. Second, Stacie Gobin (MEASURE Evaluation, Palladium) provided valuable oversight and guidance in the early stages of the project and useful suggestions for the final draft. Zulfiya Charyeva and Lisa-Marie Albert (MEASURE Evaluation, Palladium) provided a valuable review of the guide. Peter Lance and Gustavo Angeles (MEASURE Evaluation, University of North Carolina at Chapel Hill [UNC]) also offered important guidance.

We acknowledge the financial support of the United States Agency for International Development under the MEASURE Evaluation project.

We thank the knowledge management team at MEASURE Evaluation, UNC, for editorial, design, and production services.

CONTENTS

Figures	viii
Tables	viii
Abbreviations	x
Preface	xiii
1. Introduction	1
Approaches to Economic Evaluation in Public Health	3
Decision Making in Public Health	5
2. Cost Analysis	7
Introduction	7
What Is Cost Analysis?	7
Study Design and Scope	8
Data Collection Method	15
Analyzing and Presenting Cost Data	26
Cost Analysis of Complex Programs	30
Further Reading	35
3. Measuring Impact on Lives	37
Disability-Adjusted Life Years	37
Quality-Adjusted Life-Years	39
Health-Related Quality of Life	41
Value of a Statistical Life	41
Value of a Year of Healthy Life	43
Further Reading	47
4. Evaluating Impact	48
Using the Results of an Evaluation	49
Types of Evaluations	50
Impact Evaluations	51
Further Reading	61
5. Cost-Effectiveness Analysis	63
Summary Measures	65
Decision Rules	66

Other Considerations.....	70
Further Reading.....	77
6. Cost-Benefit Analysis.....	79
Introduction	79
Measuring Costs and Benefits.....	80
Discounting	82
Uncertainty	82
Conducting Cost-Benefit Analysis	83
Criticisms of Cost-Benefit Analysis	84
Related Concepts	85
Further Reading.....	87
Readers' Discussion Guide.....	88
References.....	90

FIGURES

Figure 1. Systems approach to understanding cost analysis.....	8
Figure 2. Cost elements to consider for different study perspectives.....	10
Figure 3. Breakdown of potential unit costs.....	14
Figure 4. Direct and indirect costs in health service delivery programs.....	17
Figure 5. Step-down cost accounting steps.....	19
Figure 6. Cost estimation method matrix.....	21
Figure 7. Relationships among measurement methods, data collection time, and data sources.....	23
Figure 8. Injectable cost per visit (2010).....	30
Figure 9. Distribution of cost elements, adults, 12 sites.....	35
Figure 10. Comparison of DALY and QALY calculations.....	40
Figure 11. Logic model linking inputs to impacts.....	52
Figure 12. Triangulation framework.....	57
Figure 13. Conceptual representation of measuring impact of an intervention from a baseline.....	63
Figure 14. A simple theory of change linking cost inputs to outcomes.....	65
Figure 15. Number line to represent ICER threshold to save one DALY by unit of per capita income.....	67
Figure 16. Decision or events tree.....	69
Figure 17. Frontier graph.....	76

TABLES

Table 1. Methods used in economic evaluation of public health programs.....	4
Table 2. VSL calculated using different income elasticities.....	43
Table 3. VSLY calculated using different income elasticities.....	44
Table 4. Disability weights for cleft lip and palette.....	46
Table 5. Estimates of potential DALYs averted through cleft lip and palette repair in SSA.....	46
Table 6. Reasons to conduct an evaluation.....	48
Table 7. Tips for ensuring the use of evaluations.....	50
Table 8. Common issues studied using quantitative and qualitative data.....	55
Table 9. Data collection techniques for evaluations.....	56
Table 10. Example of cost-effectiveness assessment.....	66

Table 11. Decision rules in cost-effectiveness analysis	66
Table 12. One-way sensitivity analysis of VMMC	70
Table 13. Two-way sensitivity analysis of VMMC.....	71
Table 14. League table for malaria interventions	72
Table 15. Simple DEA example	75
Table 16. Two dimensional DEA	76

ABBREVIATIONS

ABC	activity-based costing
ACCM	all-cause child mortality
ACT	artemisinin-based combination therapy
ANC	antenatal care
ART	antiretroviral therapy
ARV	antiretroviral
CBA	cost-benefit analysis
CDC	Centers for Disease Control and Prevention
CEA	cost-effectiveness analysis
CER	comparative effectiveness research
CRS	constant returns to scale
CUA	cost-utility analysis
CYP	couple-year of protection
DALY	disability-adjusted life-year
DCEA	distributional cost-effectiveness analysis
DEA	data envelope analysis
DFID	Department for International Development
DHS	Demographic and Health Survey
DW	disability weight
ECEA	extended cost-effectiveness analysis
GDP	gross domestic product
GNI	gross national income
HAART	highly active antiretroviral therapy
HHS	Department of Health and Human Services
HMIS	health management information system
HRQOL	health-related quality of life
HTC	HIV testing and counseling
ICER	incremental cost-effectiveness ratio
IPTp	intermittent preventive treatment in pregnancy

IUD	intrauterine device
ITN	insecticide-treated bed nets
LMIC	low- and middle-income country
M&E	monitoring and evaluation
MIS	Malaria Indicator Survey
MIS	management information system
MOH	Ministry of Health
MSC	most significant change
MWRA	married woman of reproductive age
NGO	nongovernmental organization
OMB	Office of Management and Budget
OVC	orphans and vulnerable children
PHC	public health center
PMTCT	prevention of mother-to-child transmission
PPP	purchasing power parity
QALE	quality-adjusted life expectancy
QALY	quality-adjusted life-year
RCT	randomized controlled trial
ROI	return on investment
SDCA	step-down cost accounting
SSA	sub-Saharan Africa
SSF	Suraj Social Franchise [voucher program]
TB	tuberculosis
UGX	Ugandan shillings
UNAIDS	Joint United Nations Programme on HIV/AIDS
US	United States
USAID	United States Agency for International Development
USEPA	United States Environmental Protection Agency
VFM	value for money
VMMC	voluntary male medical circumcision

VRS	variable returns to scale
VSL	value of a statistical life
VSLY	value of a statistical life year
WHO	World Health Organization
WTP	willingness to pay
YLD	years of life disabled
YLL	years of life lost

PREFACE

Managers and decision makers in public health face choices in a world with limited resources. This *Guide to the Fundamentals of Economic Evaluation in Public Health* presents an overview of methods and tools that can help to inform public health decisions based on economic principles. Although the guide's perspective is economic, the principles address health outcomes. The guide is intended for public health program planners, managers, and funders who are not familiar with economic evaluation but want to become familiar with its fundamentals. Some chapters provide more detail than others, especially in areas that may not be well known to many audiences: costing, cost-effectiveness analysis, and cost-benefit analyses. This is not a textbook; the intention is for readers to become familiar with the basic principles of economic evaluation; know when to use economic evaluation methods; be able to read an economic evaluation report; and be able to understand, at a general level, the approaches used and the conclusions reached. The guide will help program managers/decision makers know which approach is the most appropriate to use.

Evaluations of public health interventions, programs, or strategies look at their effectiveness in achieving the intended goal of improved health outcomes. Economic evaluation provides an additional element of understanding the *cost factors* of an intervention. Economic evaluations answer such questions as these:

- What is the magnitude of the effect that a program will achieve (or has achieved) for a given level of resources?
- Which activities are the most effective for a given level of resources?
- What is the optimal mix of health interventions?
- How can we make the best use of limited public health resources?
- What are the benefits achieved for each dollar spent on a given public health intervention?
- How can a specified public health objective be achieved at the least possible cost?
- What will it cost to scale up an intervention?

The guide presents the main tools and approaches used in economic evaluation. These are as follows:

- **Cost studies or cost analyses** (often referred to colloquially as “costing”) examine the costs of the inputs of an intervention or a series of interventions in a program (or project).
- **Impact evaluations** provide the denominators for cost effectiveness and cost-benefit analyses. Impact evaluations aim to provide an estimate of the impact of a health intervention on a specified health outcome, such as years of life gained or reductions in the length of an illness.
- **Cost-effectiveness analysis** is the measurement of the cost to achieve one unit of the desired health effect. The desired health effect can be an event in nature or the prevention of an event: for example, discrete events (a safe birth) and those that are averted (HIV, malaria, or tuberculosis [TB] infections averted, or an unwanted pregnancy avoided). The desired effect can also be defined as incremental outcomes: for example, the degree to which stunting decreased or the degree of improvement in early childhood development
- **Cost-utility analysis** is the identification of the incremental cost to gain one “quality-adjusted life-year” (QALY) or to avert one disability-adjusted life-year (DALY). DALYs and QALYs are measures

of disease burden on one year of life. Conversion tables permit QALYs to be converted into DALYs and vice versa.

- **Cost-benefit analysis** measures all effects in monetary terms at the level of society. One advantage of CBA is that all kinds of interventions—health and non-health—are comparable, because they share a common denominator (monetary terms).

Throughout the guide, the authors have strived to provide concrete, real-world examples and case studies. Where judged to be appropriate, mathematical examples of the required calculations are given. The limitations of each approach as well as any ethical considerations are also discussed.

1. INTRODUCTION

John F. Kennedy once said, “Our responsibility is one of decision, for to govern is to choose.”¹ Although he was speaking as the President of the United States, in many ways this statement describes the responsibility of any manager or decision maker in a position of authority in public health. Because people in these roles face many choices in a world with limited resources, they must often make difficult decisions. The choices may involve alternative ways of achieving the same health objective. For example, a manager may want to choose between alternative ways of delivering the same service (e.g., a fixed versus a mobile clinic). Or, the choice may be between expanding treatment for a disease or expanding prevention, or between a malaria eradication program or a childhood vaccination program.

This *Guide to the Fundamentals of Economic Evaluation in Public Health* presents an overview of methods and tools that can direct public health decisions based on economic principles. Although the guide’s perspective is economic, the principles address health outcomes. This document is intended for program planners, managers, and funders who are not familiar with economic evaluation but want to be oriented on its fundamentals. Some chapters provide more detail than others, especially in areas that may not be well known to many audiences: costing, cost-effectiveness analysis, and cost-benefit analyses. We do not expect readers to be able to conduct an economic evaluation after reading this guide; it is not a text book. The intention is that readers become familiar with the basic principles, know when to use the economic evaluation methods, be able to read an economic evaluation report, and be able to understand, at a general level, the approaches used and the conclusions. Program managers/decision makers are also expected to know which approach is the most appropriate to employ.

Public health interventions and programs often have the maximization of health outcomes as their goal. However, public health interventions, programs, and policies have costs. Costs are important when it comes to making decisions about starting, continuing, or scaling up a specific intervention or program. Of course, if resources for public health were unlimited, the costs of public health programs would not be an issue. But the reality is different and, therefore, decision makers need to understand the value a program provides in return for the costs incurred. Budgetary pressures in the public sector continue to squeeze programs, especially in developing countries, because long-term population growth and general cost inflation impact the health field. Cost-effectiveness analysis, for example, provides information that supports decision making in public health because it “helps to determine how to maximize the quality and quantity of life in a particular society that is constrained by a particular budget” (Muennig & Bounthavong, 2016).

Learning Objectives

What questions do economic evaluations address?

What are the primary uses of economic evaluations?

What are the main approaches to economic evaluations?

¹ Address at Independence Hall, Philadelphia, Pennsylvania, July 4, 1962, Available at https://www.jfklibrary.org/Research/Research-Aids/JFK-Speeches/Philadelphia-PA_19620704.aspx.

Evaluations of interventions, programs, or strategies look at the effectiveness of these efforts in achieving the intended goal of improved health outcomes. Economic evaluation provides an additional element of understanding the *cost factors* of an intervention.

Economic evaluations answer such questions as:

- What is the magnitude of the effect that a program will achieve (or has achieved) for a given level of resources?
- Which activities are the most effective for a given level of resources?
- What is the optimal mix of health interventions?
- How can we make the best use of limited public health resources?
- What are the benefits achieved for each dollar spent on a given public health intervention?
- How can a specified public health objective be achieved at the least possible cost?
- What will it cost to scale up an intervention?

There are several uses of economic evaluations:

- *Evidence for decision making.* Economic evaluations offer the researcher or analyst an additional piece of evidence to demonstrate to a decision maker that an activity is worth continuing, scaling up, or eliminating. Budgets are not limitless; economic evaluation allows the researcher or analyst to understand the relative value-for-money of a given activity in comparison with other activities (Drummond, Schupler, Claxton, Stoddart, & Torrance, 2015).
- *Evidence for stakeholders in the donor community.* As national budgets for foreign assistance decline, they are under increasing pressure to show results for their investments in global health (Shillcutt, Walker, Goodman, & Mills, 2009). Economic evaluations provide an opportunity for obtaining convincing evidence for those who are not well-versed in global health.
- *Assistance to scale-up successful programs.* Economic evaluations provide cost data that are helpful for planning or scaling up a specific activity. It is rare that the unit cost of an intervention does not change when it is scaled up.
- *Finding efficiency.* Economic evaluations help identify efficiencies in service delivery by comparing one clinic with another or one program with another. It also finds areas of cost overruns.
- *Allocation.* Economic evaluations assist decision makers to formulate their budgets to maximize impact on health outcomes.

Economic evaluations of public health programs vary in how they deal with the costs and outcomes of a given study. In economic evaluation, efficiency is viewed from two perspectives: technical or allocative efficiency.

Technical efficiency refers to the degree to which a program's activities produce the intended outcome at a specific cost. Technical efficiency sheds light on the most economical way to implement a program or intervention to gain the same level of impact. For example, is a vaccination campaign that reaches the same level of coverage cheaper with a community-based approach or through fixed service delivery points?

Allocative efficiency is a concept that refers to the optimal choice of what program(s) or intervention(s) achieve the best results for society. Allocative efficiency is about the choice of what or how much to do rather than the best way to do it. In public health, “allocative efficiency is achieved when it is not possible to increase the overall benefits produced by the health system by reallocating resources between programs. This occurs where the ratio of marginal benefits to marginal costs is equal across all health care programs in the system” (Shiell, Donaldson, Mitton, & Currie, 2002). For example, should we invest in a preventive campaign to reduce HIV infections by encouraging consistent condom use, or should we invest in more antiretroviral therapy (ART)?

Approaches to Economic Evaluation in Public Health

For technical efficiency, the main economic evaluation methods are (1) cost study or cost analysis; (2) cost-effectiveness analysis (CEA); (3) cost-utility analysis (CUA); (4) cost-benefit analysis (CBA); and (5) data envelope analysis (DEA) (Table 1).

A **cost study or cost analysis** (often referred to colloquially as “costing”) is an examination of the costs of the inputs of a specific intervention. This type of analysis does not consider health outcomes. One use of costing data is to plan for the scale-up of a health intervention. Economies of scale can lower the unit cost, but if scale-up requires coverage in remote areas, the unit cost could increase. Another reason for this type of analysis is to identify the cost drivers of a program. In the early days of HIV testing and counseling (HTC), doctors in some settings were performing counseling services pre- and post-test. When a costing study was performed and it was revealed that a highly paid doctor was performing a service that a lower paid person could do, it became common practice to have lower paid health cadres (sometimes community health workers or expert patients) perform the counseling duties. Cost studies are independent of the associated outcome, impact, or effect.

Cost-effectiveness analysis is the measurement of the cost to achieve one unit of the desired health effect. The desired health effect can be an event in nature or the prevention of an event, for example, discrete events (a safe birth) and those that are averted (HIV, malaria, or tuberculosis [TB] infections averted, or an unwanted pregnancy avoided). The desired effect can also be defined as incremental outcomes (the degree to which stunting decreased or the degree of improvement in early childhood development). The CEA often compares an intervention with another program or the status quo. A researcher can answer simple questions, such as whether an intervention achieves an effect at a cost, or whether the program offers value for money compared with another intervention. One limitation of the CEA is that although it is possible to compare interventions in which the effects are similar (i.e., HIV infections averted), a comparison across diseases requires a common denominator.

Cost-utility analysis is the identification of the incremental cost to gain one quality-adjusted life-year (QALY) or to avert one disability-adjusted life-year (DALY). DALYs and QALYs are measures of disease burden on one year of life. Conversion tables permit QALYs to be converted into DALYs and vice versa. The main benefits of CUA is that it uses a uniform metric and allows for comparison within and across disease areas, permitting decision makers to compare programs for different diseases and providing a basis for decisions.

Cost-benefit analysis measures all effects in monetary terms at the level of society. One advantage of CBA is that all kinds of interventions, health and non-health, are comparable because they share a common

denominator (monetary terms). Like CUA, it may not be programmatically relevant for researchers or evaluators to convert the outcomes into monetary terms because the calculation of the economic value of a person's life can quickly turn into an empirical and ethical predicament.

Table 1. Methods used in economic evaluation of public health programs

Type of analysis	Sample question	Unit of effect	Calculation	Example
Cost analysis	What does it cost for a client to be on highly active antiretroviral therapy (HAART) for one year?	Not applicable	Calculation of costs; no information on health outcomes	Costing the scale-up of HAART
Cost-effectiveness analysis	What is the cost per HIV infection averted of a condom distribution program?	Natural events (cases of disease and/or death averted)	Calculation of costs and health effects	CEA of HIV infections averted by a female condom distribution program
Cost-utility analysis	Should we invest United States dollars (\$) 10 million in prevention of mother-to-child transmission (PMTCT) of HIV or in ART?	QALYs/DALYs	Calculation of costs and health effects; the latter are converted into QALYs/DALYs	CUA of QALYs gained from a PMTCT versus an ART prevention program
Cost-benefit analysis	What is the return on investment of \$10 million family planning program?	Monetary value, often in US dollars	Calculation of costs and health effects; the latter are converted into a monetary value	CBA of a family planning program
Data envelopment analysis	Which service delivery points are the most efficient?	Output per input	Calculate efficiency scores using data on outputs and inputs	Ranking of HIV clinics supported by the United States President's Emergency Plan for AIDS Relief by level of efficiency

Sources: Baltussen, Adam, Tan-Torres Edejer, Hutubessy, Acharya, Evans, ... & World Health Organization, 2003; Boardman, Greenberg, Vining, & Weimer, 2011; Jamison, 2009.

The **data envelopment analysis** (DEA) method (sometimes called frontier analysis) is used to calculate technical and allocative efficiency at the service delivery level. DEA employs linear programming in a statistical approach to estimating efficiency (Akazili, Adjuik, Jehu-Appiah, & Zere, 2008). Linear programming is a mathematical technique that involves optimization (in this case, maximizing efficiency) subject to constraints (such as relationships between inputs and outputs).

Decision-Making Perspective

Which economic evaluation approach to use depends on the question(s) being asked and the level(s) of the decision maker(s) in the public health system. At the **national level**, the president or finance minister may need to make decisions about expanding primary education or improving public health infrastructure. Because the outcomes are different (increased educational achievement of the school age population versus reduced mortality of the general population), a CBA that converts outcomes into dollars is the most appropriate economic evaluation method for decisions at that level.

At the **sector level**, a minister of health typically makes decisions about interventions that affect different health programs, for example, voluntary male medical circumcision (VMMC) versus early childhood vaccinations. The outcomes are different—reduced HIV versus reduced early childhood mortality and morbidity; however, both decisions are health-related and, therefore, a CUA, which includes only benefits that affect different aspects of health is the most appropriate economic evaluation method.

At the **program level**, the director of an HIV prevention program may need to decide between programs that affect the same health outcome, for example, two interventions to improve HTC that use different approaches. In this case, CEA is the most appropriate economic evaluation method.

Decision Making in Public Health

Economic evaluations provide information to inform the decision-making process in public health.

Drummond, et al. (2015) argue for a “systematic approach” to public health decisions (including economic evaluation) and highlight four advantages:

1. Helping identify clear alternatives
2. Providing information from different perspectives
3. Quantifying the benefits and costs of alternatives instead of anecdotal evidence
4. Increasing the “explicitness” and accountability of decision making

But it would be naïve to assume that the results of economic evaluations are the only criteria that influence decisions. Other factors are:

- *Equity and social justice.* Decision makers may be concerned about specific subpopulations (e.g., rural populations) as beneficiaries of a program or intervention, even if reaching them may be more expensive.
- *Scale of impact.* Politicians and health ministers may want to show that they are reaching high levels of impact by benefiting a large number of people.
- *Feasibility.* Economic evaluations do not typically concern themselves with the administrative and social feasibility/acceptability of some programs, even if they are the most cost-effective.
- *Political.* Public health decisions are always made in the context of the public health system, which at some point is governed by politicians, not technically savvy civil servants. Politicians often weigh several factors to reach a decision; economic effectiveness is only one of them.

This guide is organized as follows: Chapter 2 discusses the methods and principles of estimating the costs of services and programs. Most economic evaluations depend on some form of costing estimates. Chapter 3 addresses measuring the impact of health events and conditions on human lives. This chapter has a discussion of measuring the impact of disease and disability on human lives. Chapter 4 is an overview of the fundamentals of evaluating public health interventions, with an emphasis on impact evaluation. Because evaluation is such a large area, the aim is to focus on some of the more salient issues and approaches. Chapter 5 covers the fundamentals of CEA and Chapter 6 describes CBA.

2. COST ANALYSIS

Introduction

All economic evaluations require estimates of the costs of delivering services. Cost analysis involves examining the **inputs** of an intervention. In the context of global health, researchers conducting a cost analysis collect and categorize costs that relate to a health program or intervention.

Compared with other forms of public health-related evaluations, the arithmetic underlying cost analysis is quite simple: determine the sum of the costs of an intervention, service, or program of interest, and if desired, divide the total amount by service delivery or beneficiary information. Although outwardly simple, there are several important factors that should be considered when requesting, funding, or conducting a cost analysis. The objective of this chapter is to review the basic concepts and processes when assessing costs in an economic evaluation.

Because this guide is intended for program implementers and monitoring and evaluation (M&E) specialists who have not had economics training, it takes a programmatic perspective. This means that there is a focus on understanding and categorizing costs incurred by a program implementer as opposed to a patient, government, or society as whole.

Costs are broadly defined as “the value of resources used to produce a good or service” (Joint United Nations Programme on HIV/AIDS [UNAIDS], 2000). In the context of a health program, resources include people, facilities, equipment, and supplies. The costs associated with these resources are the focus of the programmatic cost analysis.

Costing studies address such questions as these: What resources does the program, intervention, service, or facility use? Are there “big-ticket” costs, such as buildings, vehicles, computers, or medical technology? Are commodities or other supplies used? Who is involved, how much time do they spend, and how much is their time worth? Are there investments in technology, training, or capacity building?

What Is Cost Analysis?

Cost analysis is the first half of a full economic evaluation, in which the analyst estimates and categorizes the cost of a health intervention or service. This process is also called cost identification, programmatic cost analysis, cost minimization analysis, or cost consequence analysis. A “systems framework,” such as in Figure 1, illustrates the process by which the program inputs/costs are linked to intermediary processes or actions that ultimately result in outputs and, eventually, long-term changes in health outcomes (Janowitz & Bratt, 1994). The goal is to link the costs of inputs to the related output or outcome.

Learning Objectives

What is cost analysis?

What is the importance of perspective?

What are the main costing frameworks?

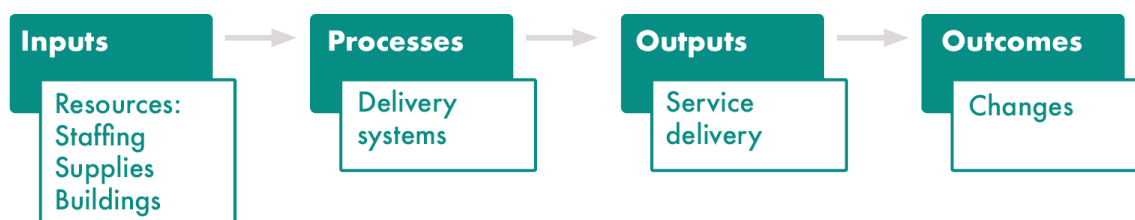
What are the main types of costs?

How are costs estimated?

What are the main sources of cost data?

What should be included in a costing report?

Figure 1. Systems approach to understanding cost analysis



Source: Janowitz & Bratt, 1994

“**Costing**” is a common shorthand for referring to the process of cost estimation that takes place in an economic evaluation. The terms “costing” and “cost analysis” are used interchangeably in this guide. The term is not used interchangeably with CEA, CBA, or other forms of economic evaluation that consider costs compared with effects.

Cost analysis is a valuable evaluation tool that is used to:

- *Compare the cost of services, sites, or providers.* Cost comparison studies assess and compare the unit cost per a determined output. A **unit cost** is the cost of providing a single good or service (Conteh & Walker, 2004). These studies are useful for programmatic decision making and improving **allocative efficiency** or setting user or reimbursement fees. They can also be used to identify **cost drivers** and compare costs across programs, locations, or services (technical efficiency).
- *Assess the effect of programmatic changes on cost.* Sometimes implementors or donors are interested in understanding whether **changes** should be made to the service delivery system or programming. These types of decisions may involve removing, adding, or supplementing program activities, staff, or services.
- *Understand the resources needed to deliver a service.* Knowing the resources that are needed for a specific service or program component enables implementors to set priorities when allocating resources or making budgets.

Cost analysis is feasible in low- and middle-income country (LMIC) settings. It can be quick, simple, and affordable when compared with other forms of economic evaluation. It is also feasible to perform cost analysis retrospectively in a LMIC setting, depending on the data systems in place.

Study Design and Scope

The first component of any research study is the establishment of the scope and methods. A cost analysis is no different. Several decisions need to be made when determining how to approach data collection for a costing activity. Many of these decisions are driven by the research question to be answered. This section is organized to reflect the order in which the researcher generally wants to make decisions.

- First, determine the *perspective*.

- Second, establish the primary *purpose, population, and focus* of the cost estimation activity.
- Third, decide whether the analysis will be limited to *financial costs* or will include *economic costs*. Will the study address all costs of a program or only additive (incremental) costs? What costs will be included and excluded?
- Fourth, define the *unit of measurement*.
- Fifth, set the *period* for when data are collected.

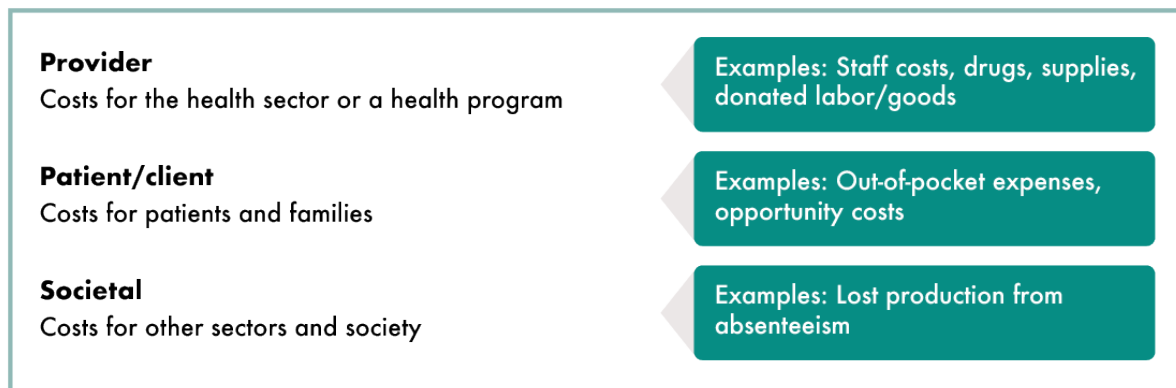
Cost Perspective

Define the **perspective** of the study or “the viewpoint from which it is conducted” (United States Agency for International Development [USAID] Global Health eLearning Center, 2011). A study’s perspective is determined at the outset because it affects the research question being addressed, method, cost elements, and statistical analysis. The potential perspectives are the patient, provider, purchaser (payer), sponsor (e.g., employer), government, and societal (Luce, Manning, Siegel, & Lipscomb, 1996). The most commonly discussed are provider, patient, and society. Figure 2 presents the costs that may be associated with these three major perspectives.

As seen in Figure 2, programs and providers are concerned about the costs of staff, supplies, and equipment. By contrast, costs considered by a study that takes the perspective of a patient involve assessing the patient’s out-of-pocket expenses, costs that are not relevant or collected when using a programmatic perspective. Similarly, the value of lost production owing to absenteeism is important when one is taking a societal perspective but is less relevant to program implementers.

The perspective chosen is often influenced by the interest of the donor or entity requesting the cost analysis. It is essential that the perspective be clearly defined and described from the start.

Figure 2. Cost elements to consider for different study perspectives



Source: Levin (n.d)

Decision Problems

Next, the researcher defines the **decision problem** or economic question to be answered when conducting the cost estimation. Like traditional evaluation research questions, there are basic contextual factors that should be outlined (Vassall, Sweeney, Kahn, Gomez, Bollinger, Marseille, ...Levin, 2017). The factors include “what,” “who,” “how,” and “why” the study is being conducted.

- What is it that the cost study is designed to consider? Define the activity, service, intervention, or output being studied.
- Who is the target population?
- How is the program implemented? Define the delivery mechanism, such as a health system, specific type of facility, or community. Researchers may also want to define the phase of the program, such as a pilot program or scale-up.
- Why is the program important? Identify the epidemiological indicators, such as incidence or prevalence of the illness being addressed and that the program seeks to influence.

Common categories for decision problems are linked to the desired goal of the study (Janowitz & Bratt, 1994). Cost comparison studies assess and compare the unit cost per a determined output. As stated above, these studies are useful for programmatic decision making and improving **allocative efficiency**. Cost comparison studies address decision problems around the denominator for unit costs, such as the cost per service, per beneficiary, or per treatment output. These unit costs can be further broken down by geographic region or by different providers, such as charity-based organizations or local implementing partners.

Implementors or donors are sometimes interested in understanding whether changes should be made to the service delivery system or programming. These types of decisions can involve removing, adding, or supplementing program activities, staff, or services. Cost information helps answer questions about when to change a program, including information on the cost savings of adding or removing a service, commodity, or program component. These changes can also include the cost savings or expense of changing the mode of delivering a service or the type of provider used. Moreover, researchers may want to know the costs of

additional training for providers to implement the change. Before scaling up a pilot program to new sites, decision makers should review local differences in the costs of providing services.

Evaluating aspects of the financial sustainability of programs is another category for cost estimation decision problems. Such costing studies may want to establish a benchmark to determine fees or reimbursement amounts for specific goods or services. The researchers may also seek to forecast the costs of meeting the demand for services so that project budgets can be determined. For example, knowing how much it costs to deliver child wellness visits at primary care centers across service delivery points helps determine the budget required for these activities.

Costing Frameworks

Once the costing study has been defined by addressing the decision problems, researchers decide what broad categories of costs to include in the analysis. The cost elements are determined by the study perspective and by the context. In this section, costing frameworks that help inform what costs to include in a study are defined.

Financial and Economic Costs

The difference between economic and financial costs is a foundational component of the cost analysis method. A study will collect and present either the financial or economic costs, and the approach used should be clearly stated.

Financial costs are direct expenditures on resources. When a payment is made for a good, service, or labor, it is considered a financial cost. The estimation of financial costs relies on knowing the price of that resource and the quantity—information most likely obtained from program financial records (electronic and paper). It also requires an understanding of who paid for the resource. If the program did not pay for the resource, deciding whether to include the value of the resource is shaped by the study's perspective. Do we care about the cost to an individual, a program, or a government? If the resources or time are donated, or if the person paying does not fit the specified point of view of the study, the value of the resources is excluded when estimating the financial costs. It is often difficult to generalize financial costs unless payment structures are similar across program sites (Vassall, Sweeney, Kahn, Gomez, Bollinger, Marseille, ...Levin, 2017).

Economic costs represent the costs “in terms of the alternative uses that have been foregone by using a resource in a particular way” (UNAIDS, 2000). Economic costs are a broader category of costs and include financial costs and **opportunity costs**. The opportunity cost of a resource, even if not paid for by a program, represents the value that the resource could have contributed if used elsewhere. Opportunity costs include the value of donated supplies, labor, or space. Sometimes called implicit costs, these costs are often more difficult to measure (Mogyorosy & Smith, 2005). Examples are free HIV test kits used at a clinic, the time of community volunteers in an orphans and vulnerable children (OVC) program, or government office space provided for free to program staff.

Whether to include economic costs is shaped by the study's perspective. For example, the value of lost labor time for a patient to travel and wait for a health service would not be included if the study looked at economic costs from the perspective of a program. By contrast, these costs would be estimated if the point of view was to assess the cost to an individual or to compare the cost of different modes of service delivery to society. Economic costing also includes **productivity costs** (sometimes called *indirect costs*), which are the costs of lost

productivity time due to death or disability. **Out-of-pocket** expenses incurred by patients for transportation or child care when seeking a service or program are likewise considered and included in an economic cost analysis.

Full and Incremental Costs

A **full cost analysis** is when researchers estimate the cost of all inputs to a program or service. An **incremental cost analysis** considers the cost of adding or expanding to an existing program or services, not considering the cost of existing programming (UNAIDS, 2000). An incremental cost analysis would not consider the following costs, unless they are needed for the added component:

- Overhead and administrative costs
- Infrastructural costs
- Program costs outside the new component being assessed

Incremental cost analysis is useful when there is a major new input or program component being added to an existing program. It is less time and resource intensive than a full cost analysis. Like most cost analysis studies, conducting an incremental cost analysis is determined by existing data structures at an organization. Incremental cost analysis often underestimates the costs of administration and is often difficult to generalize (UNAIDS, 2000).

Other Costs to Consider

There are several major categories of cost inputs that should be reviewed for applicability in cost analysis:

Above-site costs, also called upstream costs, “include various support services or activities provided by central administration” and may include office and administrative costs of head offices, training or outreach, demand generation, or centralized laboratory services (Vassall, Sweeney, Kahn, Gomez, Bollinger, Marseille, ... Levin, 2017). These costs are commonly excluded even though they constitute a large proportion of intervention costs. Nevertheless, the inclusion of above-site costs is recommended, unless their collection is not feasible, challenging, or requires significantly different measurement methods from onsite costs.

Research expenses and “**costs of supporting change**” (costs of changes due to updated guidance, reorganization of services, etc.) are two additional inputs that are often overlooked or excluded but whose contributions and costs should be closely assessed for inclusion.

Societal costs are linked to costing that takes a societal perspective. They involve a wider range of possible costs, such as loss of production from work absenteeism. Societal costs are addressed in more detail in Chapter 6 on CBA. Societal costs are more likely than other cost categories to fall outside the formal healthcare system. Obtaining such costs is often less feasible in LMICs, because of limited data availability (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014), which in turn limits the use of this perspective.

Determining the Unit of Analysis

Once the cost perspective is defined, the decision problems are addressed, and a costing framework is selected, researchers should consider the unit of analysis for the costing study. The unit of analysis generally refers to

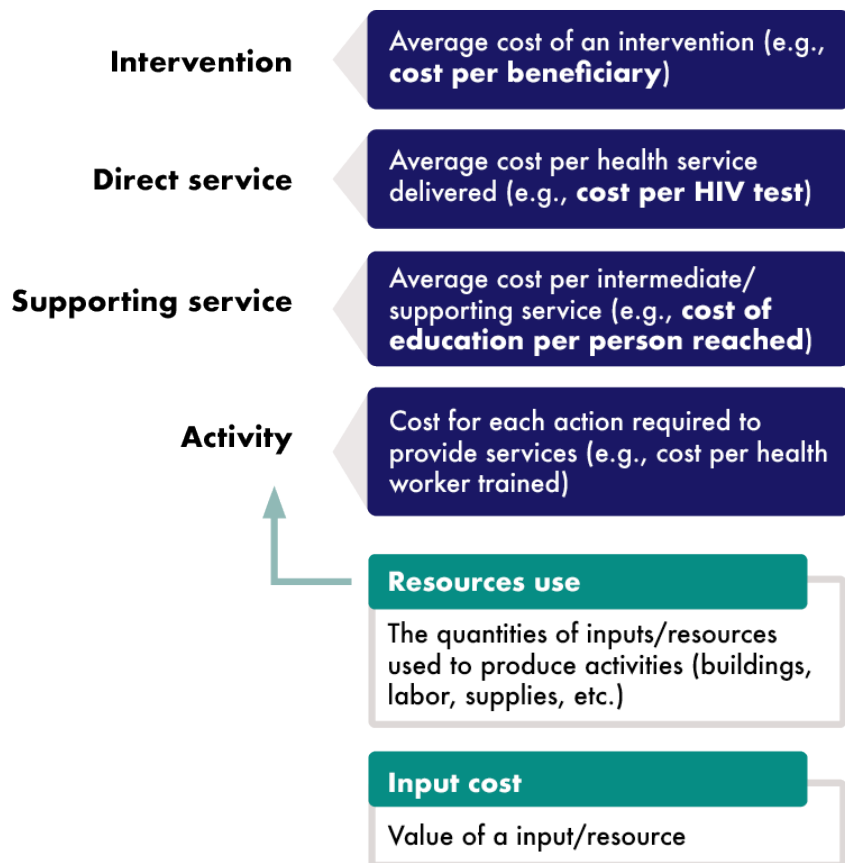
the unit cost that the study is designed to estimate. The following cost definitions should be considered as units of analysis on their own and are also helpful for defining the **unit cost**.

The **total cost** is the entire cost of producing a quantity of services or an output for a particular project or program. Total costs are the result of either a full or incremental cost analysis. An example of total cost is the entire cost of delivering HIV testing services at a public health center (PHC) in Nigeria.

Average cost is the total cost of producing a quantity of services or an output, divided by the total units of the output produced. An example is taking the total cost of delivering HIV testing services at the selected PHC in Kano, Nigeria and dividing it by the total number of HIV tests delivered during the same period. Average cost is most often used as another term for unit cost.

There are many types of outputs that can be used to determine the unit cost. There can be different levels of outputs, some that are closer to measuring an intermediate service (laboratory testing), others that measure the actual delivery of a specific service (an outpatient visit) (UNAIDS, 2000). The list of potential unit cost categories given in Figure 3 is taken from the *Reference Case for Estimating the Costs of Global Health Services and Interventions* (Vassall, Sweeney, Kahn, Gomez, Bollinger, Marseille, ...Levin, 2017). These categories should be viewed as a hierarchy, beginning at the top with the largest and most comprehensive category, which then becomes more precise toward the bottom.

Figure 3. Breakdown of potential unit costs



Source: Vassall, Sweeney, Kahn, Gomez, Bollinger, Marseille, ...Levin, 2017

Marginal cost is the cost of producing one *additional* unit of an output. An example is the cost of delivering one additional HIV test at a specific testing site. As a program expands, the marginal cost may increase, for example, if the program expands to target populations that are more difficult (expensive) to reach. When marginal costs increase, average costs also increase. Marginal costs may decrease as the volume of a service increases because there can be some “excess capacity” in terms of staffing or facilities, allowing the enterprise to reap “economies of scale.” When marginal costs go down, average costs also decrease. Marginal cost and its impact on average or unit costs is important when trying to estimate the cost of scale-up. If planners assume a constant unit (average) cost with program expansion when the average costs may change, scale-up estimates will be inaccurate.

Time Horizon

The last step in designing a costing study is to determine the “when” or the period that the study covers. The **time horizon** is the period during which costs are measured. Researchers should present a clear description of the period(s) chosen and why. The time horizon is typically selected to capture costs related to the decision problem. A sufficient length of time depends on the time required to deliver an intervention or service. Some

health programs may be continuously ongoing, making the selection of a distinct period necessary. A typical cost analysis may look at a one-year or six-month period of an ongoing program. Other interventions may have a duration of a few months to several years, allowing the researcher to collect costs for the entire duration of the intervention or planning cycle.

It may make sense to disaggregate costs into different periods to reflect the phases of an intervention and related variations in average cost. An example is to distinguish **start-up costs** (defined as all costs incurred before the start of service delivery) from implementation costs.

The time horizon also plays a role in unit cost calculations. The data collected for the two primary components of the calculation—the cost data and the count of service/output data—should share a time horizon, or phases of the time horizon, to be comparable.

Data Collection Method

After establishing the major components of the study, the next decisions concern methods for the collection of the cost data.

- First, decide on the *timing* for data collection as it aligns to program implementation.
- Second, identify the *types of costs* that will be collected.
- Third, decide from *what levels* the costs will be collected.
- Fourth, choose the *valuation method* and *analytical approach* that will be used to measure the costs.
- Fifth, outline the *sources* and *data collection tools* that will be used.
- Sixth, consider *sampling* and *data quality*.

Data Collection Timing

The timing for collecting cost data was discussed above, as part of the study design. The method for collecting cost data over time is determined by whether the cost study is prospective or retrospective. **Prospective** data collection is used when it is possible to collect data during the implementation of the health service activity. Prospective approaches involve eliciting expectations about resource use from stakeholders who are knowledgeable about the implementation of the health service activity. **Retrospective data collection** approaches use existing data and information on the resources consumed during the development and implementation of the health service activity. Although retrospective cost estimation methods are often the least costly approach, they rely on the availability and accuracy of the original database and cost recording systems, meaning that accuracy and reliability can vary widely (Luce, Manning, Siegel, & Lipscomb, 1996; Slothuus, 2000). Some studies may employ a combination approach, using retrospective data collection to assess the cost of a health program's start-up or capital costs, and concurrent or prospective approaches to collect information on time and labor or intervention-specific costs that cannot be captured retrospectively.

Types of Costs

There are many ways to define the costs that are commonly used as part of costing studies. This section provides a basic summary of the costing terms. These definitions and categories are *not mutually exclusive*. The researcher should decide how to categorize and group costs, what terms are preferred, and which terms are most appropriate given the context of the research study.

Cost elements of service costs or resource items can be numerous. An important component of the costing process is the identification of the cost elements that are relevant. Depending on the context of the cost analysis, whether it is facility-based or community-based, the cost elements a research study identifies can vary. A generic list of the main cost elements includes (Mogyorosy & Smith, 2005):

- **Human resources**
- **Commodities**, such as medical equipment, pharmaceuticals, or diagnostics
- **Capital costs**, such as buildings, land, or vehicles
- **Administrative and management costs**, including utilities, rent, security, maintenance, and cleaning of office space, information technology services, printing costs, and banking fees
- **Overhead costs**, such as negotiated indirect rates
- **Indirect costs**, including out-of-pocket expenses, lost working time or wages, or volunteer time

Cost elements can be further broken down into several categories, including direct and indirect costs, joint and non-joint costs, capital and recurrent costs, and fixed and variable costs. These cost designations can be helpful in determining the sources of costs.

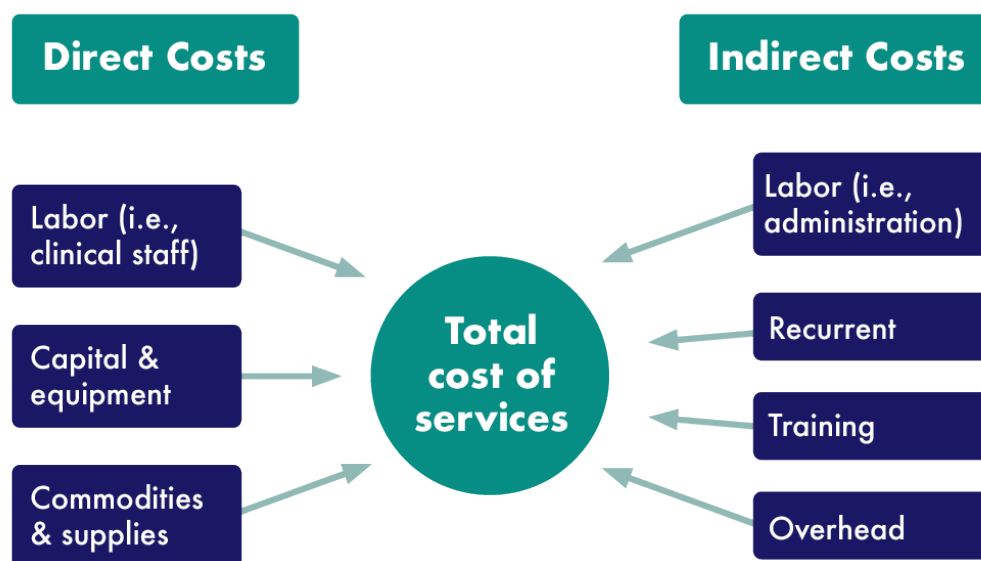
Direct and Indirect costs

Direct and **indirect costs** are costing labels that identify the relationship of a resource to a programmatic activity or health service, as summarized in Figure 4. During the cost estimation process, it is often possible to associate certain inputs with specific outputs. These terms and designations take on slightly different meanings in the context of medical services compared with health prevention or behavior change programs.

Direct costs refer to those resources that can be clearly linked to an output. For example, in the delivery of health services, the value of HIV test kits can be clearly linked to the provision of HTC services.

Indirect costs are those costs that cannot be directly identified with a service or product but are included in the costs of supporting the activities; for example, costs associated with collecting statistics, clinic administration, and office spaces or supplies. Indirect costs can also be related to production losses caused by illness or premature death (Janowitz & Bratt, 1994; Elliot & Payne, 2005).

Figure 4. Direct and indirect costs in health service delivery programs



Source: Janowitz & Bratt, 1994

Joint Costs and Non-Joint Costs

Joint costs are the costs of resources that are shared by more than one client, participant, or service. **Non-joint costs** are costs that can be completely allocated to the service or activity for which they are incurred. Similar to direct and indirect costs, this category of costs is most appropriate in the context of health clinics and service delivery. In this context, non-joint costs are the value of an HIV test kit and other disposable medical supplies used during a single patient visit for HTC. The joint costs are the value of staff time spent on the visit, equipment, and other administrative costs, such as clinic space. These joint costs cannot be allocated to a specific visit; they should be allocated across all visits.

Capital and Recurrent Costs

Recurrent costs are expenses incurred in the day-to-day provision of services, apart from personnel and commodity and pharmaceutical costs. These may include (for example) building utilities and rent, and transportation and fuel costs. Recurrent costs are relatively simple to calculate when their market price reflects their opportunity price (Janowitz & Bratt, 1994; Elliot & Payne, 2005).

Capital costs are considered investment expenses and are treated differently than operating costs. They include expenditures on durable goods or equipment. Capital costs generally have benefits longer than one year. Organizations may have their own local definitions of what is to be included in capital costs and how the rules for financial management and for assessing the current value of these costs should be applied. Capital costs are depreciated over a defined period. Although most capital costs pertain to the purchase of durable assets, the cost of pre-service or one-time in-service training should also be treated as a capital cost, because of its nonrecurrent nature.

Fixed and Variable Costs

Fixed costs are costs that remain the same regardless of the quantity of goods or services produced. This includes both recurrent and capital costs, if the value of the resources remains constant over the duration of the period being considered. These costs include but are not limited to:

- Equipment (vehicles, laboratory equipment, computers)
- Development of training or communication materials
- Overhead (building, utilities, indirect expenses)

Variable costs are the counterpart of fixed costs, constituting all costs that can change depending on the quantity of the inputs. These costs include such items as:

- Personnel allowances (travel costs and per diem)
- Supplies (printing materials, drugs, and health commodities)
- Transport costs (fuel, maintenance, taxi, public transport)

When determining the cost classifications for a study, the researcher will most likely decide between using fixed and variable cost categories OR capital and recurrent costs. This determination depends on the research question and the purpose of the study.

Costs at Different Levels of Service

Depending on the decision problem and purpose of the cost analysis, costs can be incurred at multiple levels of service delivery or programming. For example, a large-scale primary care association in Ghana delivers services at more than 50 service delivery points. There are regional offices, a central office, and related overhead costs incurred at the donor level, which in this case would be the central government ministry that supports this organization.

If the decision problem requires knowing the cost at multiple levels of implementation, where financial data can be captured separately, the *data collection activities need to be replicated at all levels of implementation* (Larson & Wambua, 2011). This is because the financial information at one level may not have the amount of detail needed for adequate disaggregation of information and allocation of costs from the lower level.

Another example is a USAID-funded HIV prevention program, where a prime organization makes lump sum payments to seven subcontracted organizations. In this context, the lump sum payments indicate the total amount spent by the subcontracted organizations, but they do not show to what the money actually contributed programmatically. To answer the research question and explore potential cost drivers, more detailed cost data need to be collected from the subcontracted organizations' records.

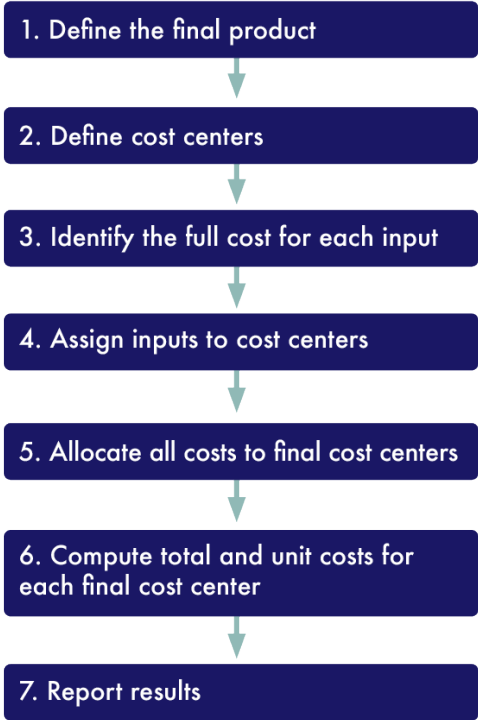
Cost Estimation Approaches

Just as costs can be broken down into several categories, there are several approaches available to estimate costs. They include cost accounting methods, such as step-down cost accounting (SDCA) and activity-based costing (ABC). Researchers should decide whether to collect costs by aggregating individual cost elements (bottom-up approaches) or by disaggregating high-level expenditures into cost categories or facilities (top-

down approaches). Mixed methods costing allows these accounting methods to be combined to suit the needs of the costing study.

Step-down cost accounting is an analytical approach to calculating unit costs that relies on a step-by-step approach. SDCA is typically broken into six or seven steps (Figure 5). (1) SDCA starts by defining the scope and perspective of the study. (2) It then requires the researcher to assess the available resources and context to determine the cost categories that will be used to identify and group costs (3) The researcher then lists all potential resources that are used in delivering the service or program. This step can also entail reviewing records or logs or conducting interviews with staff to identify costs that may not have been otherwise considered. (4) These cost line items are then grouped into the chosen cost centers; (5) depending on the nature of the line items/costs, there may be different approaches to allocating costs to the final cost centers. (6) After all costs are allocated, the final steps involve the calculation of a total cost and the related unit cost for the chosen cost center (or unit of analysis selected), (7) followed by the reporting and interpretation of results (Conteh & Walker, 2004). There are slight variations to this approach, depending on the context (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014; Larson & Wambua, 2011).

Figure 5. Step-down cost accounting steps



Source: Conteh & Walker, 2004

Activity-based costing assigns resource costs to cost objects, such as products, services, or customers, based on the activities performed. ABC is considered a better way of costing clinically-provided services compared with traditional costing approaches that measure costs at the departmental level using top-down allocation procedures. It typically does not include patient-level cost information nor is it based on service

delivery processes. As McBain, et al. (2018) point out, traditional costing “fails to capture whether, how and why clinical processes, activities and protocols vary from one patient to another, including among patients who present with the same condition. Nor does the approach give information about the actual mix of resources used to treat individual patients. Traditional cost methods simplistically assume homogeneity across patients and providers. However, evidence indicates that clinical care is highly idiosyncratic and that variation can sometimes serve a purpose, such as to customize care for a patient’s comorbidities and medical history. Equally important, such methods do not link practice variations to variation in patient outcomes” (McBain, Jerome, Leandre, Browning, Warsh, Shah, ... Kaplan, 2018).

ABC can still use a stepwise process. The primary difference is in the “cost centers” (activities) chosen and the allocation process used for indirect costs. ABC uses the observation of service delivery procedures and service provider time to calculate cost driver-specific rates that are then applied to allocate costs to the chosen cost centers (Javid, Hadian, Ghaderi, Ghaffari, & Salehi, 2016). Therefore, the primary difference from the more traditional method is in the allocation approach for indirect costs (see discussion on allocation below), which is done using personnel time spent on the direct cost centers. The advantage of this approach is the ability to use personnel interviews to determine the main activities at an organization—an approach that is practical in lower-resource settings (Conteh and Walter, 2004; Waters, Abdallah, & Santillán, 2001).

In a study of a hospital in Iran, researchers used ABC to calculate the costs of medical services using ABC and a more traditional approach (Javid, Hadian, Ghaderi, Ghaffari, & Salehi, 2016). The study revealed significant differences in the cost estimates using the two approaches. For example, emergency visit costs per patient were estimated as \$29.21 using ABC compared with \$19.20 using a more traditional approach. Similarly, radiology costs per patient were estimated at \$4.01 versus \$1.79 for ABC and a traditional approach, respectively.

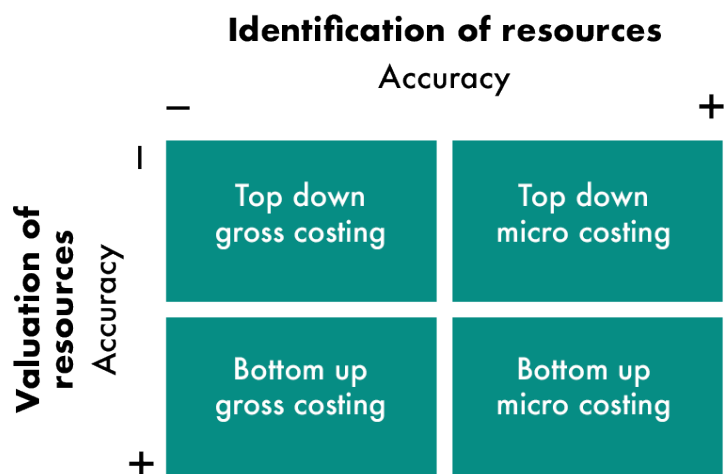
Bottom-Up versus Top-Down Approaches

Another important aspect of cost estimation is deciding whether to use a bottom-up or top-down approach for data collection and measurement. The decision on which approach to use is primarily based on what cost data are available. **Bottom-up approaches** may be either retrospective or prospective, and often lead to more detailed, accurate, and reliable cost estimates. Possible data sources for bottom-up approaches are inventories, supply lists, or use of direct observation and patient flow analyses. Bottom-up approaches to data collection and analysis are time consuming, and there are components of program or service delivery that may be missed, including start-up, social and behavior change and demand-creation efforts, training, and administrative overhead (Cunnama, Sinanovic, Ramma, Foster, Berrie, Stevens, ... Vassall, 2016). For example, when equipment costs are determined using a bottom-up approach, the equipment price, interest rate, estimate of valuable life years, maintenance costs, and a specific breakdown of the total yearly equipment utilization for all patients plus specific utilization information on patients with the disease or service of interest would be needed. The latter two components on patient utilization would need to be collected through observation, survey, or interview. This information is highly detailed and would not be available in existing information systems; it would need to be collected from scratch.

Because **top-down approaches** frequently rely on financial and accounting records and other databases, they are retrospective (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014; Slothuus, 2000). Top-down approaches rely on comprehensive data sources and aggregated cost data. Top-down approaches tend to be more efficient and less time-consuming but risk some loss of accuracy in the estimations they

provide. For example, when determining equipment costs for a top-down approach, much of the same information is needed, including the equipment price, interest rate, valuable life years estimate, and maintenance costs. Instead of patient utilization data, more readily available information on the total number of patients is used.

Figure 6. Cost estimation method matrix



Source: Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014

Sometimes used interchangeably with bottom-up and top-down costing, **micro costing** and **gross costing** methods are additional methodological approaches that can be layered on top of a top-down or bottom-up approach (Levin, n.d.; Orlewska & Mierzejewski, 2003). **Micro costing** focuses on highly detailed cost inputs. It starts with the detailed inventory and measurement of all inputs consumed in a healthcare intervention. These resources are then converted into values to produce a cost estimate. Micro costing is frequently associated with primary data collection. Existing sources can also be used and might include financial reporting data used in a top-down approach; however, because of the high level of detail available, a more granular, micro costing approach can be conducted. When done using a bottom-up approach (Vassall, Sweeney, Kahn, Gomez, Bollinger, Marseille, ... Levin, 2017), micro costing is most likely to provide an accurate identification and valuation of resources (Figure 6). The examples given above for top-down and bottom-up methods are both micro costing methods.

Gross costing approaches use aggregate information on resource use, commonly estimating total costs and dividing by the relevant unit of interest. The approach starts with the identification of a sequence of events associated with the intervention, which may include one or more of the following: hospitalizations, physicians services, and drugs. Gross cost estimation requires estimating these component event costs, then summing. The processes of measurement and valuation of resources, which are reasonably distinct in micro costing, are more blurred in gross costing. Gross costing is less likely to provide an accurate identification and valuation of resources; however, it is the simplest and least expensive approach. An example of gross costing with a top-down estimation approach is to take total expenditures or a budget and allocate them entirely to a department or service.

The **expenditure approach** is a data collection method that uses the total expenditure or budget reports from the government entity or implementing organization. Often retrospective in nature and a type of top-down costing, the expenditure approach typically results in two things: (1) the researcher takes the total cost and divides it by the chosen output measure (gross costing); and (2) the researcher goes line by line through the expenditure or budgetary reports and reassigns costs to the chosen cost categories for analysis (micro costing).

A second type of data collection method is the **ingredients approach**. The ingredients approach involves collecting information on the quantity and prices of all resources used. It is a type of bottom-up costing.

These data collection methods are often *not mutually exclusive*. Researchers often rely on a combination of both approaches to collect all data required to answer the question at hand.

Deciding which approach to use depends on:

- The decision problem and perspective of the study.
- What level of detail is required to answer the question.
- The type and complexity of the program/intervention/technology.
- The availability of data/feasibility of collecting what is needed.
- The time and resources required for both data collection and analysis.
- The need for generalizability and/or representativeness.

The methods chosen can be tailored to the context. The researcher should aim to maximize accuracy, and considering both feasibility and resource availability for conducting the cost analysis (time, money, etc). The approach can be selected depending on the availability of data for each **cost element** being considered. For example, staff time can be collected using a bottom-up, micro costing approach, whereas top-down or gross costing can be used to assess administrative costs or other activities with less readily available detail. This combination of approaches is traditionally called **mixed-methodology of costing**. In LMICs, the flexibility of this approach allows researchers to use all available data sources to maximize measurement accuracy (Levin, n.d.; Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014).

Data Sources and Measurement

Once the cost data needed for the study have been defined, researchers should determine what sources of cost data are available and how the data will be collected. Sources of cost data are extensive. In some cases, data will be collected from **existing sources**, such as financial reporting or payroll systems; records for such expenditures as utilities; reimbursements and subgrant payments; pay slips; procurement records, etc. Depending on the research question of interest and the availability (or lack thereof) of records, the researcher can collect specific cost-related data through more **direct methods or tools**, such as interviewing, observation, or surveys. The section below on data element-specific factors has a list of common cost elements considered in health service delivery or programming and their respective data collection sources or data collection methods. Figure 7 presents a basic overview of the relationships among measurement methods, the data collection time, and sources of cost data.

Figure 7. Relationships among measurement methods, data collection time, and data sources

		Relationship to time	
		Retrospective	Prospective
Data registration method	Top-down	Using hospital information system (including accounting, statistics, etc)	Not possible
	Bottom up	Using (electronic) medical records, cost accounting, billing and other information systems and/or surveys/interviews	Using (electronic) medical records, cost accounting, billing and other information systems and/or surveys/interviews

Source: Mogyorosy & Smith, 2005

Specific Sources of Cost Data

Personnel Costs

Labor is often the largest contributing cost element in health-related services or programming, making it an important component of data collection. The quantity of staff, the average salary by staff type, and an estimate of the time spent on the variable of interest that has been chosen as the unit of analysis (i.e., proportion of time spent on a specific program area, service, or activity) are needed.

Obtaining labor information often requires using a combination of existing records and direct data collection methods, including bottom-up approaches. For example, a time and labor study can be done to determine the total productive working hours and minutes spent per service. This information can be collected via interviews, activity logs, or self-administered questionnaires. The information is then used to allocate specific staff salary information to specific services/activities/outputs.

Existing records can include *salary, payroll, or pay slip information and records* for staff directly involved in the provision of services (including clinical, outreach, or laboratory staff), and administrative staff (including laboratory, secretarial, cleaning, driving, or management staff).

Capital Costs

Briefly described above, capital costs can include buildings, equipment, and even big programming costs, such as training. It is up to the researcher to gain an understanding of the program and establish an approach that adequately collects and accounts for capital costs. Capital costs are typically greater than \$1,000; however, the threshold may vary by country or organization and should be discussed before the start of data collection. Using existing records, capital costs can be collected from transportation-related records on the purchase dates and costs of vehicles. For buildings and equipment, inventory or purchase records can be used. If inventory or other records are not available, researchers can conduct the inventory themselves and estimate the market value of the items identified.

Capital costs can be direct and indirect in nature. In some cases, only a portion of the cost of a capital item, like a vehicle, is attributable to the activity or intervention of interest. The researcher should determine both the value of the capital item and the direct or indirect role the capital item plays in relation to the unit of analysis.

Because capital items are defined as those that provide a stream of benefits beyond the period in which they are purchased, the costs need to be spread out over time. This process is called annualization; it is reviewed in more detail below.

Recurrent Costs

As described above, information on recurrent costs is likely available in existing sources. Recurrent expenses can be collected from *building-related expense records*, such as rent bills, utility bills, or from *transportation-related records* for gasoline/petrol expenses or reimbursements. Recurrent costs are often indirect in nature, because they rarely contribute directly to a service or programming intervention/activity. In community-based programs, office supplies can also be considered to be recurrent costs.

Commodity Costs

Commodity costs are expenses related to health goods and supplies. They are most likely relevant to health service delivery; many disposable supplies are consumed during the delivery of health services. Commodity costs can be collected from *procurement or purchase records* for medical and laboratory supplies, pharmaceuticals, diagnostics, medical devices and instruments, and other commodities. Commodity costs can also be identified as direct and indirect. In a service delivery setting, certain supplies cannot be linked to a specific service and should be considered indirect. Others, like an intrauterine device in the context of a family planning consultation, can be directly linked to a specific service.

Training Costs

Training costs are important in many health programs. Training is commonly defined as either pre-service (before program implementation or the start of service delivery) or in-service (administered on an ongoing basis, whether it covers new technology/approaches or refreshes staff on previously trained content). The cost of training is often challenging to reconstruct, but a review of related expenses from a program financial system is possible. Training costs can also be reconstructed by asking for the number of staff trained, per diem amounts, and estimating costs related to hotels, transportation, and materials.

Overhead Costs

Overhead costs, often composed of administrative and headquarters costs, are sometimes referred to as “upstream” costs. They are frequently not captured in bottom-up data collection methods. Top-down approaches that use an entire budget or expense reports for the period of interest provide a more concrete way of estimating overhead costs. Other overhead costs can include program start-up expenses.

It can be challenging to determine the proportion of overhead costs to attribute to an activity or service.

Approaches include mark-up methods (which divide the total direct expenditure by total indirect expenditure, and then marks up each service/disease-specific cost by that percentage) (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014). Another approach for a community-based program setting is to use the average percentage of time staff spend on a specified activity to allocate overhead using the same proportion.

Data on the Unit of Analysis

Service/output **data** can be collected from *service delivery records* or from aggregate sources, such as *patient medical records* or project *management information systems* (MIS). These sources provide the count data needed on the

number of services, activities, or beneficiaries that were conducted/reached during the reporting period. The type of output measure data needed is determined by the unit of analysis.

Common Challenges with Cost Data Collection

This section describes the ways in which cost data can be collected and suggests sources for the different types of cost information. Several challenges may be encountered when selecting, collecting, and interpreting cost data. Some issues that may be faced when conducting a costing study are addressed here.

Data Availability

One major limitation to conducting cost analysis in LMIC settings is the lack of accurate and high-quality cost and program outcome data (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014). For example, it is often difficult to collect existing, top-down data reliably and confidently in settings where financial reports are spread across electronic tools in Excel, on paper, etc. It gets especially troublesome when dealing with complex health and social programs (discussed in the section below).

Choice of Cost Elements

There is wide variation in the cost elements that are included in different studies of the same topic. A common issue is the lack of headquarters costs (also called upstream costs) during the data collection and allocation processes. When cost assessments are done using a bottom-up approach, the absence of these costs leads to underestimated total costs and lower average costs. Decisions to exclude headquarters or other indirect costs are often made because of challenges in both collecting the relevant data (e.g., programs are unwilling to share indirect costs or do not know them) and issues with allocating headquarters costs (Larson & Wambua, 2011). The collection of upstream/indirect costs should be done, when possible. Allocation decisions can use an ABC approach, where resources are more limited and staff interviewing is more feasible for determining how to allocate overhead costs.

Routine Collection of Cost Data

Because capacity to capture and track data over time has increased thanks to modern technology, there is increasing demand to build capacity to track cost data routinely over time and even to integrate cost data in existing M&E systems.

Adjusting for Quality

When calculating the unit cost, it is important to remember that a lower unit cost may not necessarily imply that the service is delivered or program is provided with the same quality as a higher unit cost. Additional contextual data may be useful to understand the role of quality. (Chapter 4 includes a discussion of qualitative data in economic evaluations.)

Sampling

Depending on the nature of the costing activity taking place, there may be a need for sampling sites or clinics. A detailed discussion of sampling is not presented here. Nevertheless, it is important to consider some of the fundamental concepts of sampling for a cost study.

First, whether sampling is needed should be decided based on the number of sites, organizations, or entities from which data need to be collected. If the number is small, it may be feasible to collect information from every facility or location. If the number of sites is too large, then sampling is needed.

A **sampling framework** should be established. The framework helps limit potential bias that may be introduced due to how sites are selected. The sampling strategy selected depends on the reason for the costing activity, but the underlying goal is to develop a strategy that captures a *representative* sample of sites. Traditional resources on sampling can be applied to costing. A brief overview of a few sampling approaches follows.

- *Convenience sampling*: Sites are selected based on their convenience for data collection. Although this is quite common in costing studies, it is less than ideal and is highly likely to introduce bias and reduce generalizability.
- *Purposive sampling*: Sites are selected based on specific features, such as size, geographic location, or data quality, with the intention of accounting for factors that may impact costs. Unfortunately, a purposive sample cannot be used to make inferences or generalizations about the rest of the sites or a grouping of sites at an organization.
- *Stratified random sampling*: Sites are selected randomly from a subset of sites that meet a specified variable of interest. For example, a random selection of sites in specific regional locations are selected. This approach allows for generalization to all sites that share the chosen variable of interest.
- *Random sampling*: Sites are selected randomly in such a way as to be representative of all sites where service delivery or administration is taking place. Representative sampling ensures that sites selected randomly are like all sites at an organization based on specific attributes that can affect data quality. These attributes are like those discussed above in the purposive sampling description. The primary difference is that a random subset of sites is selected in each attribute and determined to be important to the organization and question of interest. Random sampling allows for the generalization or inference of data from samples sites to the rest of the organization. It is often quite difficult to do with costing research, but is nonetheless the gold standard.

Analyzing and Presenting Cost Data

After the cost data have been collected from a variety of sources, they should be organized and analyzed to be of use to decision makers. In this section, we present a few ways of allocating costs, including how to spread certain types of costs across activities and periods. Several considerations when presenting the results of a costing study are then described.

Cost Allocation

As previously discussed, there are some costs that are either not directly incurred in producing a health service, or the cost can be for an input for the service in question but can also be jointly incurred with another service. Examples in the first category include administrative costs incurred at a program's headquarters, ministry of health administrative costs, the cost of provider's time not associated with a specific service, and the cost of general service personnel, such as cleaners and receptionists. These costs may be referred to as overhead,

indirect, or upstream costs. The second category is often referred to as joint costs and can include construction and/or rent for physical facilities and capital goods (computers, autoclaves, buses, and cars).

Cost allocation is the assignment of such costs to the specific health service that is being costed. The terms cost distribution, cost apportionment, and cost assignment are often used interchangeably with cost allocation. The first thing to know about cost allocation is that all cost allocation rules are arbitrary. Second, cost allocation may get complicated depending on the relationship between the operating units of an organization. Large health programs that provide multiple services in multiple service delivery points fall in this second category.

For public health costing, there are three main ways that upstream and joint costs can be allocated to specific program services:

1. Allocate proportionally to the services provided. In this approach, service delivery data are used to calculate the distribution of the upstream and joint costs proportionally to the services being provided. Therefore, if an HIV program is providing multiple services, it would allocate the upstream and joint costs to PMTCT according to the proportion that PMTCT clients are of all patients served.
2. Allocate according to the proportion of direct service labor involved in the service. With this method, the researcher should know either the total level of effort (in person-days or person-hours) providing all services or their value.
3. Use an ABC approach to allocate proportionally by identified activities. ABC can more accurately assign indirect costs in cases where there may be economies of scale (Chan, 1993, Cokins 1996).

Annualizing Capital Costs

As previously discussed, **capital costs** are large expenditures on items such as buildings, vehicles, laboratory or computer equipment, or large investments at the beginning of a project. Because capital assets provide benefits and services beyond the period in which they were purchased, the value of a capital cost needs to be spread out over more than one period.

The **annualization** of capital costs requires an estimate of the good's replacement cost and an estimate of its useful life. A standard procedure is to use "straight-line" annualization. This approach simply divides the original purchase price of the capital asset by the number of useful years of life of the asset. For example, an OVC program purchases a vehicle during the first year of the program that costs \$20,000. The vehicle will last five years. This means that the annual cost of the vehicle is \$4,000.

In some cases, capital items are donated. In this case, the replacement value should be estimated. The market value of the item can be obtained from local suppliers or manufacturers or other records. The useful life of equipment or capital expenses should also be context-specific. The Choosing Interventions that are Cost-Effective project (<https://www.who.int/choice/cost-effectiveness/en/>) of the World Health Organization (WHO) provides standard life-years for general equipment. Manufacturers can provide the same information for specialized medical equipment. Program finance staff can also provide standard interest rates and life-year estimates (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014).

Costing Considerations

When presenting cost information, certain topics need to be addressed, for example, the main conclusions of the costing study, including answers to the economic questions (or decision problems) that motivated the study. One topic that may be of interest are the major factors that drive the total cost, known as **cost drivers**.

Cost drivers identify from where the bulk of expenses at a site come. Is the cost of a selected reproductive health service driven by the high cost of a medical specialist or by the need for a certain package of commodities? Does an HIV prevention program have more costs at the headquarters level than in the field where implementation happens? Is this because of staffing or the high cost of HIV testing kits purchased centrally? Cost drivers can be useful for providing an additional level of detail on top of other findings, such as the unit cost. Moreover, by presenting the unit cost, total cost, or even a cost description, a more detailed breakdown can provide additional information to inform the interpretation and understanding of the findings. Knowing the main cost drivers can also assist with understanding where potential efficiencies may be realized.

Below is a list of important concepts to remember when presenting results. These topics are common across many cost studies and should be addressed to the best of the researchers' ability. Time and budget constraints may not allow for a full analysis of each issue (such as a comprehensive sensitivity analysis), but addressing these topics in the context of a costing study can add credibility.

Transparency

Having collected, allocated, and estimated costs, it is imperative to transparently and appropriately convey results. If costs were included or excluded during the costing process, they should have been recorded and should be identified when presenting unit costs so that the reader is aware of the scope and boundaries of the analysis. If allocation decisions were made for specific cost elements, they also need to be described.

Heterogeneity

Like more traditional health research methods, it is important to consider why unit costs vary, including such drivers as the **size and scale** of the program, the characteristics of the **population** or subpopulations, service delivery **platform**, geographic **setting**, or **quality** of care. Sample size should be considered. Larger sample sizes may merit use of cost functions (a topic that is beyond the scope of this guide). To deal with potential heterogeneity in studies with smaller sample sizes, consider (1) reporting unit costs disaggregated by site together with a description of the characteristics of each site; (2) reporting average cost elements disaggregated by features identified as significant through the use of statistical testing of differences; or (3) including qualitative data to support and bolster the quantitative findings.

Dealing with Uncertainty

All cost estimates are subject to **uncertainty**. In this context, uncertainty refers to different forms of potential bias that can be introduced in a cost study. Although the ways in which the researcher deals with uncertainty falls outside the scope of this guide, brief definitions of common forms and solutions follow.

Common types of uncertainty are parameter uncertainty, model uncertainty, and generalizability uncertainty (Hendriks, Kunda, Boers, Bolarinwa, Te Pas, Akande, ... Swan Tan, 2014). Parameter uncertainty is due to variation in the estimation of variables, such as staff time. **Sensitivity analysis** is a technique that tests the robustness of the conclusions. It involves repeating the comparison between inputs and consequences while

varying the assumptions that underlie the estimates. Univariate and multivariate analysis can be used to assess the fidelity of variable estimates.

Ethical Considerations

When working with field-based organizations, it is important to consider the willingness of people in the organization to participate in the study (Larson & Wambua, 2011). Cost analyses often use many resources, including staff time. In LMICs and resource-constrained settings, the service delivery organization's cooperation in providing information and data is crucial to ensuring good data quality, timely responses from participating entities, and the eventual completion of the data collection process.

Costing work is often excluded from human subject's approval because the data are organizational or aggregate in nature. This does not mean that economic evaluators are exempt from ethical considerations. This also does not mean that local organizations (who often compete for funding against fellow development implementors) are not concerned about sharing financial information with outside researchers. Their concerns are legitimate and should be addressed as best as possible to ensure that organizational staff are convinced of the usefulness of the research and are fully willing in their participation. Reducing harm in this context also requires conscientious use of the participating organization's resources during the data collection process, balanced by the need for high quality and detailed data.

When working in these contexts, it is also the responsibility of outside researchers to be thoughtful about how **feedback on findings** is provided to the participating organizations, and how the interpretation and use of the results can be facilitated.

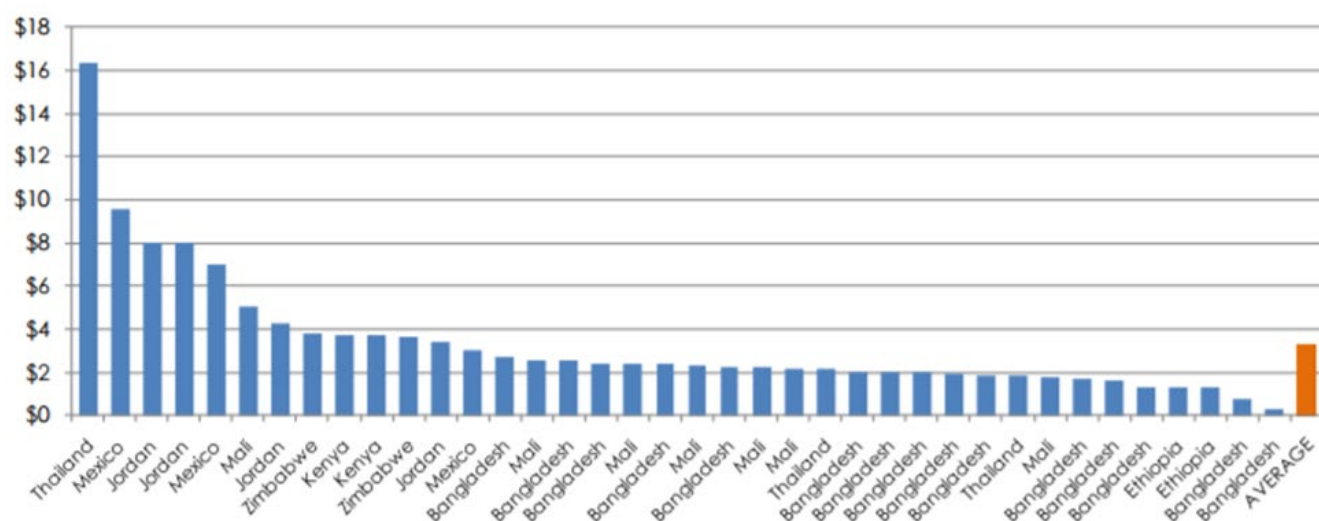
Cost Studies and Cost Systems

Cost studies are often one-off studies that differ in approach and structure based on the existing records that are in place at an organization. There is usually great variation in the cost elements included, the methods used to gather costs, and geographic coverage, making comparisons and use of the data in other settings difficult.

A **cost system** is a tool that makes use of the widespread availability of routine data at health service delivery organizations to collect cost data routinely using existing records. It treats cost data as M&E data, enabling organizations to use cost data to inform program management and decision making.

Figure 8 is an example of a comparison of family planning cost studies. The example looks at the cost per visit for injectables, a type of long-lasting reversible contraceptive. In this situation, there is an interest in understanding how the variation can be due to differences in the service delivery approach (was service delivery at a clinic or in the community?), differences in salaries for personnel, or other programmatic features. Unfortunately, conclusions about the impact of a program feature cannot be drawn, because of the wide variation in methods used for these studies. There were major differences in the types of cost components that were included. When indirect/overhead costs were included, the estimates were higher, on average. The calculation and allocation approaches used also varied greatly for such key cost elements as capital costs, personnel, and the value of donated items or time. Although still not commonly used, a cost system can help prevent some of these types of issues, either in an organization or across organizations.

Figure 8. Injectable cost per visit (2010)



Source: Moreland, 2011

The Reproductive Health Cost Reporting System is a tool developed by MEASURE Evaluation to systematize the collection of existing financial, human resource, equipment, and service delivery data. It allows organizations to use existing records to calculate the average cost of their services, also called “cost per service.” The unit costs can then be compared across sites and regions of the organization and broken down by various cost elements. The annual nature of the data enables programs to assess trends in service costs. The system has reports and graphics to illustrate and summarize these comparisons. The Reproductive Health Cost Reporting System is intended to be used as management tool by organizations in LMICs that provide family planning and reproductive health services and have multiple service delivery points.

For more information on cost systems or if you are interested in building a cost system for your organization, see Moreland, Foley, & Gobin (2018).

Cost Analysis of Complex Programs

Complex health and social care programs are interventions that “consist of a number of interconnecting elements that seem essential to the proper functioning of the intervention but the mechanism through which this is achieved is uncertain” (Byford & Sefton, 2003). Complex programs may impact many areas of a person or family’s life, have multiple goals, provide packages of care, and intervene over a long period of time. The populations targeted by these programs may be heterogeneous and they may have a varied degree of involvement in the program.

Traditional economic evaluation of complex programs falls in the realm of welfare economics, wherein the recommended perspective is societal, prospective data collection is preferred, and measurement focuses on final outcomes as opposed to intermediate outcomes or program outputs (Byford & Sefton, 2003). This guide’s focus on program evaluation renders many of these preferences unfeasible or irrelevant. Does that

mean that more complex global health programs, such as OVC or HIV prevention programs, cannot be costed?

The short answer is no! These programs can still be costed; however, the use of prospective data collection, a provider/program perspective, and output measures in lieu of outcomes restrict researchers from doing a full cost-effectiveness analysis. Instead, cost analysis or cost outcome analysis is more appropriate (Larson & Wambua, 2011). For example, when costing OVC programs, a key component of the process is eliciting a full and complete description of the program components. This description is imperative when developing the input categories into which the costs will be organized. Larson & Wambua (2011) used the following input categories for an OVC program in Kenya: nongovernmental organization (NGO) staff, NGO office, NGO other office costs/supplies, NGO travel/meetings/M&E, facilitators/peer educators, savings and loan association materials and services, mentors, and OVC education expenses. These “input categories” somewhat aligned with the initial financial report categories but drew on additional descriptive information that was available to expand the researchers’ ability to describe program costs with a more intervention-oriented approach. These efforts often require the researcher to map activities taking place in the program, linking them to costs that are already being tracked. This is challenging but necessary with complex programs. The Gobin and Foley (2019a) report on rapid costing of structural and behavioral HIV interventions showcases this type of mapping approach.

Another challenge in costing complex interventions such as OVC programs is the inability to disaggregate costs into more detailed intervention areas. For example, OVC programs provide a wide array of services, including healthcare linkage and education, school fees and clothing support for children, village savings and loan groups for caregivers, and household case management support through community case workers/volunteers, among other things. Providing these services is dependent on the needs of the child and household, and measuring the cost and effect of individual service areas has been a long-standing issue in complex program evaluation. Use of qualitative research to bolster and support cost analysis data is one potential way to expand capacity for evaluating complex programs in greater detail. MEASURE Evaluation is at the forefront of methodological approaches to integrate qualitative data in cost analysis. Gobin and Foley (2019b) used qualitative interview data with case workers to better contextualize costs related to case management delivery. Qualitative data were used to develop and finalize the cost categories in which the economic data were presented. These data were also used to explore the relationship between the quality of case management delivery and project spending decisions.

Case Study : The Cost of Case Management in an Orphans and Vulnerable Children Project in Rwanda

Background

Interventions for orphans and vulnerable children (OVC) are socioeconomically driven, community-based services for children under age 18 who have lost one or both parents to AIDS (United States President's Emergency Plan for AIDS Relief [PEPFAR], 2012). OVC programs aim to improve children's resilience to meet their basic needs of health, safety, stability, and schooling, through the provision of such services as case management, psychosocial support, early childhood development, and household economic strengthening. The end goal of OVC programming is to reduce vulnerability to HIV and AIDS (PEPFAR, 2015).

Case management is a cornerstone of orphans and vulnerable children (OVC) programming and the platform on which services are delivered. Few cost analyses have attempted to disaggregate the cost of case management from other OVC program service areas (Santa-Ana-Tellez et al, 2011). This study engaged with OVC projects in six countries to gain insight on current approaches to OVC case management implementation, map how costs can be linked to case management activities and determine the cost of case management per beneficiary reached in each project. This case study outlines findings from the Turengere Abana program in Rwanda.

Methods

Data collection took place from June 25, 2017 to July 10, 2017 in Kigali, Burera, and Huye districts in Rwanda. Retrospective financial costs and beneficiary data were collected simultaneously with the implementation of in-depth qualitative interviews with project staff and PSWs. The interviews explored a wide range of experiences related to case management delivery, capacity, and quality. Staff self-reported their level of effort (LOE) spent on case management. Qualitative data were analyzed through content analysis of researcher interview notes and were used to support the categorization and assignment of costs. Quantitative data were analyzed through a combination of activity-based costing and step-down cost accounting. Project beneficiary data were used to calculate the cost per beneficiary, and the qualitative results were used to explore and interpret the quantitative findings and identify seven main cost centers.

Allocations were used to proportionally attribute shared costs to case management. Current values were applied to recurrent inputs and an annualized value to capital inputs. Administrative costs were assessed and allocated based on their relevance to the activity, often using average estimates of personnel's LOE. Indirect costs were excluded. The total amount spent on case management was calculated and compared with the total expenditures for each project. Project beneficiary data were used to calculate the cost per beneficiary.

Findings

Case management in Turengere Abana is carried out by volunteer CWs called para social workers (PSWs), who are also project beneficiaries. Their role is to conduct monthly home visits to households identified and enrolled in the project, to support household economic strengthening and internal savings and lending groups (ISLGs) for caregivers; provide educational support; provide healthcare insurance support for families; make referrals and linkages to healthcare; and offer HIV prevention, treatment, and support services. The PSWs also provide education to their assigned households, on such topics as home gardens, nutrition, sanitation in the home, and gender-based violence. The PSWs are directly supervised by FXB staff, called field facilitators.

Turengere Abana PSWs manage an average of 21.7 households (ranging from 10 to 30 households), with an estimated caseload of approximately 50 beneficiaries. The PSWs are selected by their peers in the ISLG to lead their group and provide case management for those eligible in their community. The PSWs must be able to read and write. Most PSWs have a primary level of education, with about six years of schooling and 1.6 years of experience. They receive training from their supervisors on a wide range of topics, including how to visit homes, complete forms, provide referrals, and other case management-related activities.

The total cost of case management for the 2.7 years of the current project phase was \$360,120, which averages to \$134,489 annually. This figure constitutes 22 percent of Turengere Abana's total OVC project costs. With 16,401 beneficiaries served by the project during this period, the cost per beneficiary comes to \$21.96. This cost does not account for the opportunity cost of the PSWs' time and labor, donated village office spaces, and any out-of-pocket expenses the PSWs incur for cell phone use or travel.

The cost of case management in the Turengere Abana project is heavily driven by the supervision cascade and staffing to oversee, support, and provide on-the-job training to the PSWs (44.7%). The other major cost driver is transportation-related costs (22.4%), because the field facilitators are provided motorbikes, and regional managers use vehicles to get to the field offices and local communities. Turengere Abana has conducted a lot of in-service training for paid staff over the life of the project. The training costs related to case management concerned beneficiary selection criteria and tailoring services to needs, FXB's data collection tools, and conducting meetings and managing the PSWs (8.9%). CW support—a category that captures the costs of CW stipends, printing of household visit and other case management forms, monthly meetings, CW transportation, and communication—constitutes a small proportion of case management-related expenses (1.2%). The other cost categories consist largely of allocated costs for general personnel (21.2%), office support (1.1%), and M&E (0.5%).

Conclusions

Case management was viewed as an integral part of programming, and the high estimates from staff on the proportion of their time spent on case management reflect this finding. Apart from staff time, the project was challenged to define the costs and activities that were related to case management. FXB Rwanda's Turengere Abana case management approach relies heavily on the time and supervisory capacity of project staff. The PSWs, who are also project beneficiaries, do not receive extensive financial support for their role, but they reported having capacity and motivation to give back to their communities.

Assessing the cost of a single component of OVC programming is easier to do with a mixed-methods approach, bolstering the quantitative cost data with qualitative research methods. The parallel approach of collecting both quantitative and qualitative data allowed the researchers to better understand the costs as they related to case management, increasing both the validity of the results and the level of detail when looking at the data. Further research should consider the quality of case management as it relates to cost, to better understand the benefits and drawbacks of supervision-driven case management, such as the approach found in the Turengere Abana project.

Source: Gobin & Foley (2019b)

Case Study: The Costs of HIV Treatment, Care, and Support Services in Uganda

Background

This study assessed the cost of HIV care services in public and non-public facilities that provide HIV care and support services in Uganda. The research gathered data from 12 sites and included adult and pediatric patients. The objectives were:

- To determine the **average annual unit cost per patient** (adult and child) for specific HIV treatment, care, and support services.
- To establish the key **cost components or “drivers”** of the HIV treatment, care, and support services.
- To determine the **costs borne by patients** (“out-of-pocket costs”) that are not incurred in a clinical facility.
- To **compare cost variations** by level of service delivery.

Methods

The data collected were used to estimate the per-patient costs for both ART and non-ART patients. Patient-level information was gathered on services received, physical functionality, socioeconomic background characteristics, and costs incurred by patients when receiving care.

A purposive sample of 12 of Uganda's accredited ART centers was chosen to reflect key characteristics thought to influence unit cost, including: level of service delivery; major implementing partner; type of ownership; and geographic location. The main sampling criterion was level of service delivery, with the final sample proportionally selected to represent the five service delivery levels in the country.

The study aimed to conduct 600 patient interviews (an average of 50 per site) with adults, and 200 (an average of approximately 16 per site) with children (or their caregivers.) The numbers constituting each sample were determined to ensure a reasonable representation of the site populations and to facilitate any required subgroup analysis. Key informant interviews were conducted with clinicians, nurses, social workers, pharmacists/dispensers, hospital administrators or managers, and accountants to obtain information on service costs data.

Data were collected at the facility level on major cost elements, such as staffing, patient load, and services provided. The study captured both financial and economic costs.

Direct costs collected were:

- Staff time in caring for clients
- Commodities, including drugs to prevent and treat opportunistic infections; antiretrovirals (ARVs); and medical consumables and supplies used for clinic visits and for laboratory testing
- Capital expenditures for medical equipment; vehicles used directly for client care; and physical infrastructure used for client care

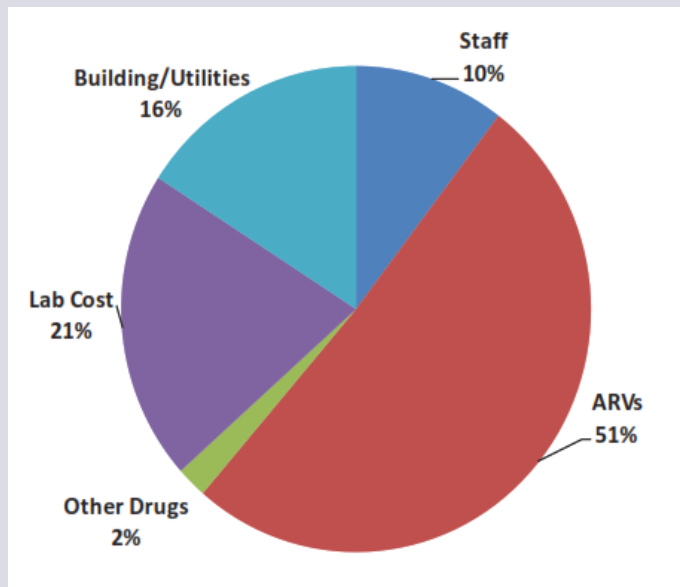
Indirect cost data collected at the facility were:

- Labor cost of administrative staff
- Overhead expenses (e.g., office supplies, travel expenses, communication), and equipment and building use generically at a facility
- Depreciation of equipment and assets (e.g., equipment and furniture in the clinical and diagnostic units)
- Costs to program clients (e.g., client time, transport, meals, out-of-pocket payments, user fees for services or drugs) and waiting time at the facility (to assess lost work time)

Findings

Results showed that the annual facility-level cost of providing HIV treatment, care, and support to adult HIV patients ranged from Ugandan shillings (UGX) 254,000 to UGX 824,000 (\$116.28 to \$376.20) across the 12 sites, with a median cost of UGX 567,000 (\$258.78). When restricted to adults on ART, annualized costs ranged from UGX 403,000 to UGX 1,330,000 (\$183.54 to \$606.48), with a median of UGX 734,000 (\$335.16). For child HIV patients, the study found that costs ranged from UGX 190,000 to UGX 1,869,000 (\$86.64 to \$852.72), with a median cost of UGX 630,000 (\$287.28).

Figure 9. Distribution of cost elements, adults, 12 sites



Source: Moreland, et al., 2013

The main cost drivers for both adults and children were ARVs and laboratory tests (Figure 9). Among all adults, ARVs accounted for 51 percent of costs, and for children, ARVs accounted for 69 percent. For laboratory services, the percentages were 21 percent and 16 percent, respectively. First-line drugs accounted for the largest share of ARVs. The study also found that among the facilities contacted, public hospitals had a higher cost per adult patient than did non-public, not-for-profit hospitals, but only by approximately 12 percent (UGX 655,018 versus UGX 582,894.) Public hospitals had lower staff costs per patient whereas laboratory costs were higher. By contrast, PHCs had much lower per-patient costs than non-PHCs for adults (UGX 335,625 versus UGX 512,073.) The major factor here was much cheaper staff costs in PHCs (UGX 28,850 versus UGX 168,845.) Similar patterns were found for children.

Source: Moreland, Namisango, Paxton, & Powell, 2013.

Further Reading

Because many aspects of costing can be specific to the condition being studied, a variety of costing guides and tools have been developed for specific public health services and diseases. Here are a few examples:

- **HIV/AIDS and HIV prevention:** Paxton, A., & Carvalho, N. (2013). *HIV/AIDS program costing tools: Concepts and methods used under the USAID Health Policy Initiative, Costing Task Order*. Washington, DC: Futures Group, Health Policy Initiative, Costing Task Order. Available from http://www.healthpolicyplus.com/archive/ns/pubs/hpi/1551_1_Costing_Tool_Guide_FINAL.pdf.
- **Family planning:** Moreland, S. (2011). *Methods for estimating the costs of family planning: Report of the expert group meeting on family planning costing* (WS-11-19.pdf). Chapel Hill, NC, USA: MEASURE

Evaluation, University of North Carolina. Available from
<https://www.measureevaluation.org/resources/publications/ws-11-29>

- **OVC:** Larson, B.A. & Wambua, N. (2011). How to calculate the annual costs of NGO-implemented programmes to support orphans and vulnerable children: A six-step approach. *Journal of the International AIDS Society*, 14, 59. Available from <https://www.ncbi.nlm.nih.gov/pubmed/22182588>.
- **mHealth:** LeFevre, A.E., Shillcutt, S.D., Broomhead, S., Labrique, A.B., & Jones, T. (2017). Defining a staged-based process for economic and financial evaluations of mHealth programs. *Cost Effectiveness and Resource Allocation*, 15, 5. Available from <https://resource-allocation.biomedcentral.com/articles/10.1186/s12962-017-0067-6>.

3. MEASURING IMPACT ON LIVES

Chapter 2 presented ways to estimate the costs of a public health intervention. Perhaps by definition, public health programs are designed to prevent loss of life, extend it, or prevent and reduce cases of disease. Economic analysis requires a way to measure the value of the number of human lives saved by an intervention or the cases of a particular disease or condition that can be avoided. In this chapter, summary measures of the burden of disease on human lives are defined. They can be used to compare the magnitude of impact of different diseases and the effects of interventions. The measures discussed in this chapter are mainly used in cost-benefit and cost-utility analyses (which are reviewed in Chapter 6). Because costs are measured in monetary units, health benefits need to be monetized to perform a comparison. This chapter discusses measures of the value of healthy life that economic evaluations use to make such comparisons. How these measures are constructed, their limitations, and how they can be used to estimate the value of the benefits of public health programs are discussed.

Learning Objectives

Understanding disability-adjusted life years
Understanding quality-adjusted life years
What is the value of a statistical life?
How is willingness-to-pay used?

Disability-Adjusted Life Years

Disability-adjusted life years (DALYs) represent years of healthy life lost due to mortality and morbidity from a disease. DALYs measure the burden of disease by aggregating years of life lost due to early death and a measure of years of disability in a single measure of mortality and morbidity. The measure of disability is a weighted fraction of the number of years lived with the disability, based on a disability weight (DW) that represents the severity of the illness (Mathers, Vos, Lopez, Salomon, & Ezzati, 2001).

DALYs are calculated for specific diseases and causes of death or disability. They allow different health challenges that are usually measured with different metrics to be compared. The DALY metric was designed to facilitate these comparisons so that cost-effectiveness studies can use a common measure that encompasses deaths across age groups and non-fatal outcomes.

Calculating DALYs

In simple terms, total DALYs are the sum of **years of life lost (YLL)** and **years lived with disability (YLD)**:

$$DALY = YLL + YLD$$

In turn, YLL estimates the number of additional years a person would have lived in the absence of the disease. It is a function of the number of deaths (N) and the standard life expectancy at the age of death (L):

$$YLL = N \times L$$

YLD quantifies the impact of disability on a person's life, both in terms of how long the person lives in less than ideal health and also the severity of poor health. It is a function of the number of cases per year

(incidence I), the DW , and the average duration of the disease (L). The DW reflects the severity of the disease on a scale from zero (perfect health) to one (dead).

The stylized formula for YLD is:

$$YLD = I \times DW \times L$$

DALY weights summarize the severity of the impact of a condition on the lives of patients. The impact can include pain, loss of mobility, mental health conditions, or other disabilities. In its Global Burden of Disease Study, WHO provides estimates of DALY weights for numerous conditions, for treated and untreated cases, and disaggregated by the ages of patients. Some of these estimates are broken down by sequelae, or consequence, for conditions that have a variety of presentations. For example, malaria DALY weights are estimated for standard cases and for complications, such as anemia and neurological symptoms.

DALY weights are generally estimated using stated preference methods, where people are surveyed about their preferences between disability scenarios (Salomon, Haagsma, Davis, de Noordhout, Polinder, Havelaar, ... Vos, 2015). A weight of 0.5 does not mean that a year of life with a certain condition is halfway between life and death, rather, the weight orders the level of disability for this condition relative to others.

Example:

Brucellosis is a bacterial infection that can be contracted by eating infected meat or milk or encountering infected animals. Symptoms of brucellosis include fever, headache, and fatigue and, if untreated, may lead to arthritis, chronic fatigue, and swelling of internal organs. The DW for brucellosis has been estimated at 0.11. In Ethiopia, the average brucellosis infection lasts for 21 days. (This estimate may differ across countries based on the likelihood that an infection is correctly diagnosed and treated.)

Ethiopia had a population of 99,169,144 in 2015. In that year, about 126,000 cases of brucellosis were reported, with approximately 2,300 deaths associated with the infections. Assuming that cases and deaths are uniformly distributed across the population, the average YLL equals the standard life expectancy (about 86 years) minus the average age of the population (19 years).

$$YLL = N \times L = 2,300 \times (86 - 19) = 154,000$$

$$YLD = I \times DW \times L = 126,000 \times 0.11 \times (21 / 365) = 797$$

$$DALY = YLL + YLD = 154,797$$

In this case, the duration of the disease is short, and the symptoms are, on average, not very severe, so the YLL far outnumber the YLD.

DALY calculations can be disaggregated by age and sex to reflect differences in the impact of disease across the population. For example, some diseases may disproportionately affect the elderly, whereas occupational diseases may affect males and females (who often have different life expectancies) differently. Diseases that affect younger populations result in a greater number of years lost per affected person, so the age distribution assumption can greatly affect DALY results.

WHO provides a calculation template² for disaggregated DALY calculations. The template also allows for age-weighting, where years of life for young adults are more valuable than those of the very young or the very old. Most comparisons across countries and diseases do not use age-weighting, but it may be of interest in certain cases. Comparison studies generally use standard expected life expectancy for all regions. Although actual life expectancies differ, some applications may call for actual life expectancy data. In the numerical example above of brucellosis in Ethiopia, this would mean using a life expectancy of 65.5³ rather than 86.

The disaggregated DALY calculation requires population data by sex and age group⁴ and the distribution of cases of the disease under consideration. In the absence of disaggregated incidence data, it can be assumed that the disease affects all populations equally. DWs can also vary by age range, where estimates by age are available and appropriate.

Considerations for Estimating DALYs

DALY measurements were designed to combine the impact of mortality and morbidity of disease in a single metric, allowing for comparisons across diseases. Comparing across diseases requires the use of DWs that rank conditions in terms of severity. Constructing meaningful estimates of DALYs depends on careful application of the correct DWs to the reported cases (in terms of the sequelae reported, age group, and treated versus untreated cases). When possible, differences in incidence by age and sex should be considered. If the data are not available to perform this disaggregation, qualitative discussion can be used to address any potential bias in the results.

Age-weighting involves a judgement about the relative value of years of life at different ages. Similarly, using a local estimate of life expectancy instead of the standard value discounts the lives of people who live in regions with shorter average lifespans. Assumptions about these model parameters should be stated transparently, and decisions to use age-weighting or non-standard life expectancies should be explained in the analysis. Sensitivity analysis can be performed around these assumptions and the DWs.

Quality-Adjusted Life-Years

Quality adjusted life years (QALYs) are another measure that combines mortality with quality of life that can be used in economic evaluation (Whitehead & Shehzad, 2010). As opposed to DALYs, QALYs do not use age-weighting to value life at different ages differently. For DALYs, a year of healthy life is represented as zero years lost, a year of life lost is represented as one year, and a year of life lived with disability is a fraction of a year determined by the DW (and potentially age-weighting).

The WHO considers health to be “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” (Grad, 2002). Therefore, QALYs measure the overall quality of life instead of the magnitude of disabilities. QALYs represent a year of healthy life as one year, a year of life lost as zero years, and a year of life lived with disability as a fraction of a year determined by the quality of life weighted for the year and condition, Q .

² http://www.who.int/healthinfo/global_burden_disease/tools_national/en/

³ <https://www.worldlifeexpectancy.com/ethiopia-life-expectancy>

⁴ Available from the Spectrum DemProj tool at <https://www.avenirhealth.org/software-spectrummodels.php#demproj>

$$QALY_t = 1 * Q, Q \leq 1$$

The quality-adjusted life expectancy (QALE) is then defined as the sum over the life expectancy of QALYs per year:

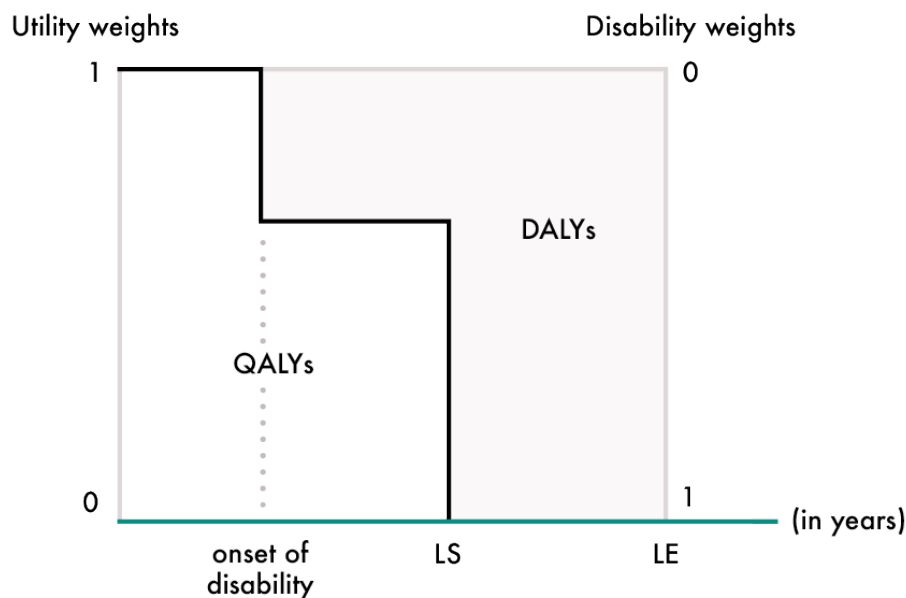
$$QALE = \sum_t QALY_t$$

Figure 10 compares DALYs and QALYs graphically. LE represents life expectancy at birth, and LS represents the actual lifespan that has been shortened due to a health condition. The difference between the two lifespans represents the YLL. The difference between the onset of disability and the actual lifespan represents the amount of time lived with the disability. In the DALY calculation, these years are multiplied by the DW to get YLD. The shaded area represents the DALYs.

On the other hand, QALYs measure the quality of life instead of the amount of disability. From birth until the onset of disability, each year of perfect health is valued at one. The years lived with the disability are multiplied by the utility weight to represent the reduction in quality of life. The QALYs for the individual's life are represented by the white area in Figure 10.

Comparing the two areas in the figure, where $Q = 1 - D$, we can convert between QALYs and DALYs, assuming there is no age-weighting and the treatment is effective over the entire remaining lifespan (Sassi, 2006).

Figure 10. Comparison of DALY and QALY calculations



Adding up QALYs across the population and comparing the outcomes under different scenarios of lifespan and disability provide an estimate of the quality of life gained as a result of an intervention. QALYs can be gained by applying treatments that extend lifespans and/or improve the quality of life in each year of life. As with DALYs, the QALYs should be discounted over time. The formula for QALYs gained is:

$$QALYs\ gained\ due\ to\ intervention = \sum_{t=a}^{a+L} Q_i / (1+r)^{t-a} - \sum_{t=a}^{a+L} Q_t / (1+r)^{t-a}$$

where a is the patient's current age, L is the duration of illness, L_i is the duration of the benefit of treatment, and r is the discount rate.

As with other measures, QALYs cannot capture all the benefits of health interventions. Improving quality of life can allow people to become more productive and generate economic benefits. Improving the quality of life of an individual may also improve the quality of life of their family members or other members of the community.

As opposed to DALYs, QALYs do not allow for age weighting. In some cases, policymakers may be interested in investing in the health of young, productive people, in which case the DALY calculation allows for young adults to receive higher age weighting. QALYs consider improvements in the health of any member of the population to be of the same value.

Health-Related Quality of Life

QALY utility weights are estimated using survey methods. The **time trade-off** method asks people to choose whether they would prefer to live a longer life with a specific disability or a shorter life in perfect health. The trade-off between the years of life informs the utility weights, known as **health-related quality of life** (HRQOL) measures. An alternative estimation technique, known as the **standard gamble**, asks people whether they would risk a given lifespan with disability for a gamble resulting in either perfect health or death. The probability of death that people will accept then informs the HRQOL score for the condition.

HRQOL may also refer to a public health indicator that measures health disparities and population trends in health using self-assessments (Centers for Disease Control and Prevention [CDC], 2018). Respondents are asked to report their general health, mental health, physical health, and ability to perform normal activities over the past 30 days, known as Healthy Days measures. These measures can be used to examine disparities in health status between different populations.

Note that the Healthy Days measures and the utility weights used in QALY calculations are different methods that are referred to by the same name. Karimi and Brazier (2016) differentiate the different uses of the terms "quality of life" and "HRQOL" in the literature.

Value of a Statistical Life

Background

The benefit to society of preventing a fatality can be measured using the **value of a statistical life** (VSL). This measure has been constructed to represent the amount people are willing to pay for improved safety to save one life. In high-income countries, such as the United States (US), VSL measures are often considered part of the benefits of public health, environmental health, or public safety interventions.

Measures of the value of a human life were first created in the 1960s and included only the income lost due to premature death. These measures understated the value of human lives, and so the willingness to pay for safety approach was developed (US Department of Transportation, 2016). Most modern VSL measures are estimated by examining preferences for safety using the additional compensation required by people to accept riskier jobs (hedonic wage studies). Other types of estimates are based on observed purchasing data, such as preferences

for vehicles with different safety features (revealed preference studies) or data collected from surveys (stated preference studies).

An application of the VSL is the estimation of the benefits of reduced mortality due to a policy intervention. Multiplying the number of lives saved by the VSL gives an estimate of the monetary value of these benefits, which can be used in a CBA, for example.

Using the VSL

The current VSL estimate used by the US government to conduct CBA is \$9.1 million in 2012 US dollars. When using this value for policy analysis, the VSL number should be converted to dollars for the base year of the analysis, using the following formula:

$$VSL_t = VSL_{t0} * (P_t / P_{t0}) * (I_t / I_{t0})^\epsilon \quad (1)$$

Where t is the base year of the analysis, t_0 is the original year of the VSL estimate, P_t is the price levels in each year, and I_t is real incomes in each year. The adjustment in price levels can be based on the consumer price index. Because VSL measures are based on willingness to pay for safety, people with different levels of income have different preferences; therefore, the VSL estimate should change annually with real income. The US Department of Transportation uses an income elasticity of 1.0 to update its annual estimates.

Transferring the VSL from one population to another—such as US VSL estimates to other countries—requires information about their relative incomes and an income elasticity that measures how responsive the VSL is to changes in income. Households with lower incomes spend a larger proportion of their budgets on basic goods, such as food and energy, and are not willing to spend the same relative amount of income for improvements in safety.

Studies have shown that VSL increases with income and that higher income populations have proportionally higher VSL (Hammit & Robinson, 2011). This implies an income elasticity greater than one, although empirical studies to estimate the elasticity are sparse (US Environmental Protection Agency [USEPA], 2016). When using VSL estimates for lower-income countries, sensitivity analysis around values of the income elasticity greater than one can be helpful. To calculate the VSL for country c , use the ratio of gross domestic product (GDP) per capita at purchasing power parity (PPP) for the US and country c and the selected income elasticity (Viscusi & Masterman, 2017).

$$VSL_c = VSL_{US} * (GDP\ PPP_c / GDP\ PPP_{US})^\epsilon \quad (2)$$

Example:

The US VSL for 2015 is \$9.6 million. To transfer this VSL to Ethiopia, the relative incomes for the US and Ethiopia are needed.

US PPP GDP per capita in 2015 = \$56,444

Ethiopia PPP GDP per capita in 2015 = \$1,633

Using this formula, the VSL for Ethiopia is calculated for three different income elasticity values (Table 2).

Table 2. VSL calculated using different income elasticities

	$\epsilon=1$	$\epsilon=1.5$	$\epsilon=2$
VSL Ethiopia 2015	\$277,741	\$47,242	\$8,035

When using the VSL as part of a CBA, for example, the value of lives saved can be calculated for multiple years. As discussed above, the VSL in each year should be adjusted to account for changing income and price levels. For each year of the analysis, the base year VSL should be adjusted by the assumed rate of real income growth, g .

$$VSL_t = VSL_{t0} * (1 + g)^{t-t0}$$

Limitations of the VSL

Because the true value of the income elasticity that governs preferences for safety is unknown, the analyst should choose a range of values, typically greater than one. Sensitivity analysis can be useful to demonstrate the magnitude of the results under different income elasticity assumptions, such as central, high, and low cases. Once an elasticity value is chosen, the calculated VSL is applied in aggregate to all people regardless of their individual incomes, risks, and preferences. This measure is useful for CBA but ignores issues of how benefits are distributed across the population.

VSL does not measure the morbidity impact of injury or disease; this should be estimated separately. By their nature, VSL measurements may make readers uncomfortable. When using VSL estimates, be clear about the interpretation of the measurement and what it does and does not include.

Value of a Year of Healthy Life

As compared to the VSL measure of the economic impact of mortality, DALYs combine both mortality and morbidity in one measurement. However, DALYs are expressed in units of time instead of monetary units. For CBA, measurement of the benefits of reduced mortality and morbidity in monetary units that can be compared with program costs is needed. The simplest assumption for the value of a DALY is the value of GDP or gross national income (GNI) per capita. This measure assumes that the value of a DALY is equivalent to the lost income due to a year of lost working time. As with the VSL, this measure should be considered a lower-bound on the value of a year of life because people value healthy life for more reasons than earning potential.

Ideally, estimates of the willingness to pay (WTP) to avoid health conditions is available for a variety of conditions and income levels. In practice, estimates of WTP values for non-fatal health risks are often unavailable. Lost income can be calculated from the amount of missed work time associated with a particular health condition and per capita income, but again, this measure excludes the discomfort experienced by those who suffer the disability or the potential missed work of family members due to caretaking activities. Averted medical expenditures can also be included as an economic impact of a public health intervention.

Value of a Statistical Life Year

Another proxy for the value of human health is an annualized VSL measure, calculated as an estimate of the willingness to pay (WTP) for a year of healthy life. The value of a statistical life per year (VSLY) annualizes the VSL based on average life expectancy. In this formula, the VSL is discounted during the remaining years of life:

$$VSLY_{US} = VSL_{US} / \sum_t^{US \text{ remaining life expectancy}} [(1 + r)^{-t}]$$

Using the 2015 US VSL of \$9.6 million, an average age of 40 years and life expectancy of 78 years (for a remaining life expectancy of 38 years), and a three percent discount rate, the US VSLY is estimated to be about \$409,000.

Because we previously converted VSL from the US value to a value for low-income countries, we can also convert VLSY values using an income elasticity. The US VSL is consistent with US life expectancy, so we convert the annual value, rather than converting the VSL, then discounting over life expectancy. The VLSY for country c is calculated using the ratio of income (at PPP) and an income elasticity.

$$VSLY_c = VSLY_{US} * (GDP\ PPP_c / GDP\ PPP_{US})^\epsilon$$

Example:

We have previously calculated a 2015 US value for VSLY of \$409,000.

US PPP GDP per capita in 2015 = \$56,444

Ethiopia PPP GDP per capita in 2015 = \$1,633

Using formula above, the VSLY for Ethiopia is calculated for three different income elasticity values (Table 3).

Table 3. VSLY calculated using different income elasticities

	$\epsilon=1$	$\epsilon=1.5$	$\epsilon=2$
VSLY Ethiopia 2015	\$11,833	\$2,013	\$342

In this example, the VSLY value is 23 percent higher than GDP per capita.

WTP for a DALY or QALY

We can now combine a measurement of the total impact of mortality and morbidity with the value of a year of healthy life to calculate the total impact of a specific disease or condition in monetary terms.

$$Total\ impact\ of\ disease_t = VSLY_t * DALY_{s_t}$$

Example:

We previously computed the annual DALYs due to brucellosis in Ethiopia to be 154,797 in 2015. We also computed the 2015 VSLY for Ethiopia to be \$2,103 (assuming an income elasticity of 1.5). The total impact of brucellosis in 2015 is then:

$$154,797 * \$2,103 = \$325,538,091$$

GDP in Ethiopia in 2015 was \$64 billion, so the impact of brucellosis is approximately 0.5 percent of GDP.

For a CBA, we need to compute the value of the health benefits of a policy intervention over a period of several years. We can calculate DALYs in each year in the absence of an intervention, and then with an intervention, using different assumptions about parameters, such as incidence (intervention results in fewer cases), deaths (intervention reduces mortality), or the number of treated and untreated cases (intervention results in more treated cases that have better outcomes).

In each year, we should adjust the VSLY to account for growth in real income, g :

$$VSLY_t = VSLY_{t_0} * (1 + g)^{t-t_0}$$

The total monetary impact of a disease is then equal to the sum over the time horizon, T , of the VSLY times the DALYs in each year. The total impact for each year should be discounted appropriately based on the assumed rate for the analysis, r :

$$Total\ impact\ of\ disease = \sum_t^T [VSLY_t * DALYs_t * (1 + r)^{-t}]$$

Alternately, the value of the QALYs gained can be calculated instead of the value of DALYs avoided. As discussed above, DALYs and QALYs can be converted under certain assumptions and should produce results in similar orders of magnitude.

Limitations of the VSLY

As with the estimates of VSL, the value of VSLY depends on the choice of income elasticity. Estimates of this parameter are difficult to find. Countries should differ in their VSLY due to differences in income, preferences, and life expectancy. Ideally, WTP should be used to avoid each specific condition for each country, but these estimates are also generally unavailable. As with the other benefit measures described above, sensitivity analysis can be used to demonstrate a possible range of impact.

VSLY estimates are helpful for comparing the social costs of different conditions and policy options to mitigate them. When presenting these costs, care should be taken to explain the meaning and interpretation of the measurements.

Case Study: Cleft Palette Repair in sub-Saharan Africa

Treatment for cleft lip and palette is not widely available in sub-Saharan Africa (SSA) due to a lack of surgical facilities. Although the condition is not fatal, children born with it are at risk of "problems with feeding and speaking clearly and can have ear infections. They also might have hearing problems and problems with their teeth" (CDC, 2017). Organizations that treat cleft lip and palette in the developing world have estimated that repair costs are approximately \$250 per case.

The 2004 Global Burden of Disease estimates of DWs are shown in Table 4. Note that the treatment does not eliminate the disability but reduces its magnitude.

Table 4. Disability weights for cleft lip and palette

	Untreated	Treated	Incidence rate (US per 1,000 births)
Cleft lip	0.098	0.016	0.6
Cleft palette	0.231	0.015	0.5

The number of cases in SSA was estimated from US incidence rates applied to the SSA population, resulting in an estimate of 34,683 new cases in 2008. The disability is assumed to persist throughout the expected life span. The authors calculated the DALYs due to cleft lip and palette in SSA with and without discounting (at a rate of 3%) and with and without age-weighting. They calculated DALYs using the same age weights for all countries, and again, using country-specific age weights. This second set of age weights was also applied to the VSLY calculation and was constructed so that the VSLY peaks at two-thirds of life expectancy for each country.

The DALY estimates are shown in Table 5.

Table 5. Estimates of potential DALYs averted through cleft lip and palette repair in SSA

Condition	Total cases	Total DALYs averted		
		3% discount rate 0.04 age weight	No discounting No age weighting	3% discount rate country-specific age weights
Cleft lip	18,918	45,896	80,493	36,013
Cleft palette	15,765	100,747	176,692	79,251
Total	34,683	146,643	257,185	115,354

The authors calculated the VSL for each country using an income elasticity of 0.55 and the VSLY using the country-specific age weights. They computed two values of VSLY for each country—one based on adult VSL and one based on child VSL (1.8 times higher). These VSLY values were multiplied by the DALYs to estimate the magnitude of the impact of cleft lip and palette. For comparison, the authors also used GNI per capita as an estimate of annual income to calculate the loss of earnings due to these conditions.

Using GNI per capita, the total income loss due to cleft lip and palette ranged from \$251.7 million to \$441.1 million. Using VSLY, the estimates ranged from \$5.4 billion to \$9.7 billion. The authors argued that additional surgical capacity in the region needed to realize these benefits could extend the ability to treat numerous other conditions and improve healthcare while delivering large returns on investment.

Source: Alkire, Huges, Nash, Vincent, & Meara, 2011

Further Reading

Measuring Impact on Lives

The WHO has studied the global burden of disease (GBD) since 1990, using DALYs as a metric, to compare the importance of diseases across populations. More information is available from <http://www.healthdata.org/gbd>.

For an overview on the use and history of QALYs and DALYs, see: Gold, M.R., Stevenson, D., & Fryback, D.G. (2002). HALYs and QALYs and DALYs, oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health*, 23(1), 115–134. Available from <https://www.annualreviews.org/doi/full/10.1146/annurev.publhealth.23.100901.140513>.

For a discussion of the different ways the terms “quality of life” and “HRQOL” are used in the literature: Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: What is the difference? *Pharmacoeconomics*, 34(7), 645–649. Available from <https://doi.org/10.1007/s40273-016-0389-9>.

Valuing Health Impact

For a technical discussion of how to value QALYs for economic analysis, see Adler, M.D. (2006). QALYs and policy evaluation: A new perspective. *Yale Journal of Health Policy, Law, and Ethics*, 6(1), Article 1. Available from <http://digitalcommons.law.yale.edu/yjhple/vol6/iss1/1>.

4. EVALUATING IMPACT

In a world of finite resources, donors—whether national governments, foundations, or bilateral and multilateral aid agencies, increasingly want to know whether the programs they fund are successful and are having measurable and verifiable impact. This is especially true for public health. Evaluation is the main way these questions are answered. *Economic evaluation* includes the cost dimension of the program or intervention in the evaluation.

Therefore, all forms of economic evaluation—cost-effectiveness, cost-benefit, and cost-utility analyses—require an estimate of the impact of the evaluated intervention or program. For example, to estimate the cost per HIV infection averted by a condom distribution program, the results of an evaluation are needed to estimate the number of HIV infections averted. In short, evaluations provide data for the denominators used in economic evaluations.

It is useful to specify some of the common reasons to do an evaluation because they may influence the kinds of information an evaluation provides, who will use the information, and how it will be used, including what decisions it can inform. Table 6 lists some of the reasons to do an evaluation and why it can be useful.

Learning Objectives

What are the reasons to evaluate a health project or program?

What are the uses of evaluations?

What are the main types of evaluations?

What is a “counterfactual” and how is it used?

Understanding the role of mixed methods in evaluations

How are complex interventions evaluated?

Table 6. Reasons to conduct an evaluation

Why do impact evaluations?	When are impact evaluations most useful?
To decide whether to continue or expand an intervention	Interventions where there is not a good understanding of their impact, and better evidence is needed to inform decisions about whether to continue funding them or to redirect funding to other interventions.
To learn how to replicate or scale up a pilot	Innovative interventions and pilot programs that, if proven successful, can be scaled up or replicated.
To learn how to successfully adapt a successful intervention to suit another context	Periodic evaluations of the impact of a portfolio of interventions in a sector or a region to guide policy, future intervention design, and funding decisions.
To reassure funders, including donors and taxpayers (upward accountability) that money is being wisely invested	Interventions with a higher risk profile, such as a large investment (currently or in the future), high potential for significant negative impact, or sensitive policy issues.
To inform intended beneficiaries and communities (downward accountability) about whether, and in what ways, a program is benefiting the community	Interventions where there is a need for stakeholders to better understand each other's contributions and perspectives.

Source: Bonbright, 2012

Using the Results of an Evaluation

Before providing an overview of some of the key approaches and issues around public health evaluation, it is important to address how the evaluation results may (or may not) be used and why. There may be resistance to an evaluation by the implementing agency for several reasons, which can impede or even prevent a successful evaluation. Based on Bonbright (2012), four challenges can be identified. If these challenges are recognized at the outset and taken into account in the design and strategy for the evaluation, the chances of success and usefulness are increased.

- Many programs do not understand the value and uses of impact evaluation. The value to the program is not clear, whereas the financial costs are significant.
- Conducting impact evaluations requires a level of technical skill and knowledge that many organizations do not have. An external evaluator may be viewed with suspicion.
- There may be fear that the evaluation will be judgmental and cast a negative light on the organization; for example, it could threaten the organization's reputation, the main concern being reduced future funding.
- According to Bonbright, "donors typically are, if anything, at a more primitive stage in their own understanding and use of evaluation than their grantees. Surveys consistently show that donors neither provide sufficient funding for nor understand how to support and use evaluation well. It is important to clarify that this characterization is of donors as a whole, encompassing individuals and institutions."

Table 7 provides some useful tips to overcome these challenges and increase the use of evaluation findings.

Table 7. Tips for ensuring the use of evaluations

An operational checklist for using evaluation results
Planning & management
1. Develop a value proposition for each potential user.
2. Estimate what evidence will be useful for which user at what time.
3. Recruit a team from across all organizational units that will be required to ensure that identified user needs are met.
4. When possible, make full use of existing research data to reduce the burden of original data collection for staff.
5. Build a communications strategy for evaluations that differentiates internal and external communications and includes user-appropriate reporting formats.
6. Map existing systems and activities for opportunities to include small steps for evaluation use.
7. Routinely include discussions of evaluation findings in staff meetings.
8. Track awareness and use through an evaluation scorecard that aggregates up to higher-level organizational scorecards. Indicators should include timeliness and quality of staff inputs and resulting actions.
User engagement & measuring evaluation use and impact
9. Before the evaluation begins, engage users to test the evaluation hypotheses and proposed indicators, and to determine when and how to best report the findings.
10. Validate tentative findings and deepen interpretations through consultations with users.
11. Conduct assessments one month after the evaluation has been reported to learn where and how the findings of the evaluation are known (awareness) and used.
12. Conduct assessments six months after the evaluation has been reported to learn how the evaluation may have changed users' beliefs and behaviors.

Source: Bonbright, 2012

Types of Evaluations

There are several kinds of evaluations that can be conducted; however, the most appropriate one for economic evaluation is *impact evaluation*. According to USAID:

“**Impact evaluations** measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a treatment or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.” (USAID, 2011)

The purpose of a **process evaluation** (sometimes called a performance or program evaluation) is to assess and document how the program or project is being implemented. Key questions a process evaluation helps address are: What difficulties or challenges were encountered when implementing the program? How well was the program implemented? Did the program reach its intended beneficiaries? Are expected results being achieved?

An **outcome evaluation** assesses the program or project's effectiveness in achieving the intended changes or outcomes. Examples of questions an outcome evaluation addresses are: Did the program succeed in helping households increase their dietary intake? Was knowledge of sustainable farming increased as a result of the workshops or training?

Implementation science looks at the extent to which effective health interventions are effectively integrated in actual, real-world public health and clinical service systems. "It compares multiple evidence-based interventions, identifies strategies to encourage the provision and use of effective health services, promotes the integration of evidence into policy and program decisions with the goal of adapting interventions to a range of populations and settings, and identifies approaches for scaling up effective interventions to improve health care delivery" (Spiegelman, 2016).

Comparative effectiveness research (CER) is an evaluation approach that compares existing health care interventions to determine which interventions provide the greatest benefits and possible harm to patients. The main questions of CER are which treatment works best, for whom, and under what circumstances. According to the US Institute of Medicine, "The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels" (Institute of Medicine, 2009). CER uses many of the other evaluation approaches discussed in this guide, including CEA, QALYs, and randomized controlled trials (RCTs).

Impact Evaluations

Of these approaches, impact evaluation is the most useful for economic evaluation. Two key concepts to understand in impact evaluation are (1) what is meant by the **impact** (or "observed change" in the 2011 USAID definition above);, and (2) the **counterfactual**. The **impact** of public health interventions normally occurs over a long term, in comparison with outcomes that are more intermediate in nature. Changes in the uptake of an improved health service is an example of an intermediate output, whereas an improvement in health status, perhaps measured by a lower level of morbidity for a disease, constitutes an impact. Knowing when to draw the line between intermediate and longer-term effects of an intervention can be guided by a well-thought-out logic model or theory of change. Some projects focus their objectives on outputs and not on impact. In these cases, an evaluation that focuses on impact is difficult, because impact may not be measured.

The other important element is the **counterfactual**. The reason a counterfactual is used is because the evaluation seeks to attribute the impact to the intervention. Ideally, the counterfactual measures outcomes that would have occurred in the absence of the program or intervention. An evaluation is interested in measuring impact due *only* to the intervention. But as Lance, et al. (Lance, Guilkey, Hattori, & Angeles, 2014) observe, for an individual, this is not possible because the individual "cannot be observed in two states at the same time," and the time when the individual is observed is varied (for example, before and after an intervention), and other factors that are affected by time may creep in and could affect the outcome. The solution that is often used, therefore, is to observe *different* people, some in the program and some outside the program. The

challenge here is that there may be fundamental differences in the characteristics or behaviors between those in the program and those who are not; that is, it may not be possible to choose the non-participating people in such a way as to say that the only differences in observed outcomes between them and those in the program are due to the program.

Impact evaluation focuses on high-level questions, essentially around whether and to what extent an intervention or program achieved the intended results in terms of its ultimate impact on health.

The questions that an impact evaluation typically answer are:

- Did the project meet its objectives?
- What was the project's impact?
- Who benefited?
- Is the impact sustainable?
- What other factors were at work that may have helped or hindered the project in achieving impact?

There are several possible methods that can be used for impact evaluation. The following guidance from Rogers (2012) is useful in making the choice:

Clarification of the Value of the Evaluation

It is important to define what is considered a success or a failure. For example, in a public health campaign to reduce malaria by distributing bed nets, what level of reduction in reported malaria cases in the project area would be judged a success? The distribution of benefits is also important: who benefits and who does not? In some cases, these criteria have been defined in the project or intervention design. However, it is still important to articulate them when designing the evaluation.

Framework Development

When conducting an impact evaluation, it is very useful to have a framework in mind for how the intervention is intended to lead to the expected impact or outcomes. Such a framework is often referred to as a logic model, theory of change, or results chain, although there are slight differences among these three frameworks, which are discussed here.

Logic Models

Logic models link inputs to outcomes and impact in a direct and linear fashion, as shown in Figure 11.

Figure 11. Logic model linking inputs to impacts



Inputs refer to human, financial, social, political, or other resources needed to begin and complete the program. Examples of inputs are hiring trained staff, developing and printing behavior change communication materials, and providing technical assistance.

Processes refer to the activities that programs carry out to achieve their objectives. Examples are the distribution of family planning commodities, training, or educational events designed to increase the use of contraceptives.

Outputs refer to the results of the efforts at the program level. In population, health, and environment projects, outputs refer to training, behavior change communication activities, delivery of selected health services, and completion of community-based natural resource management plans.

Outcomes refer to changes measured at the population and habitat levels. Examples are changes in the target population's knowledge and behaviors, and increased tree and wildlife species in the target habitat. Long-term outcomes also refer to coverage and disease prevalence.

Impact is outcomes, but at a higher or longer-term level. Sometimes impact and outcomes are merged. A program can have an outcome that is measured by a reduction in a specific disease morbidity whereas its impact can be measured in DALYs or in increases in life expectancy.

Theory of Change

A ***theory of change*** is a depiction of a “big picture” scenario of how a program intends to lead to a desired outcome or impact. It is often described using a diagram with nonlinear pathways from one step to the next, is often unstructured, and may vary from one program to the next. Theories of change often include external factors that are outside the control of the program but have the potential to affect one or more program outcome. The basic characteristics of theories of change, as detailed by Bullen (2014) are:

- Gives the big picture, including issues related to the environment or context that you cannot control.
- Shows all the different pathways that can lead to change, even if those pathways are not related to your program.
- Describes *how* and *why* you think change happens.
- Can be used to complete the sentence “if we do X then Y will change *because...*”.
- Is presented as a diagram with narrative text.
- The diagram is flexible and does not have a particular format; it can include cyclical processes, feedback loops, one box could lead to multiple other boxes, different shapes can be used, etc.
- Describes why you think one box will lead to another box (e.g., if you think increased knowledge will lead to behavior change, is that an assumption or do you have evidence to show it is the case?).
- Is mainly used as a tool for program design and evaluation.

Logical Frameworks

Logical frameworks are often used as an alternative organizing framework for linking interventions to impact and can be very similar to logic models. In a way, logical frameworks fall somewhere between logic models and theories of change. Again, to quote Bullen (2014), a logical framework:

- Gives a detailed description of the program, showing how the program activities will lead to the immediate outputs, and how these will lead to the outcomes and goal (the terminology used varies by organization).
- Can be used to complete the sentence “we plan to do X, which will give Y result.”
- Is normally shown as a matrix, called a log frame. It can also be shown as a flow chart, which is sometimes called a logic model.
- Is linear, which means that all activities lead to outputs, which lead to outcomes and the goal; there are no cyclical processes or feedback loops.
- Includes space for risks and assumptions, although they are usually only basic. It does not include evidence for why you think one thing will lead to another.
- Is mainly used as a tool for monitoring.

Understanding Impact

As we have discussed, one of the central questions that an impact evaluation tries to answer is the extent to which an intervention or program can be seen as responsible for a result. However, rarely is an intervention the sole source for a change in health outcomes. Public health interventions often work in concert with other programs or factors that are favorable to the desired outcome.

For example, a program to increase the uptake of family planning services by improving the quality of service delivery can also be implemented with a program of behavior change communication. In such a case, the researcher may only be interested in knowing if the combined interventions have made a difference, in which case the logic model or theory of change would include both interventions, and the main challenge could be controlling for non-program factors, such as demographic or education changes that could affect the demand for family planning services and affect uptake.

In cases where it is desired to understand the specific role of a single intervention, one then relies on the theory of change and focuses on three components (Rogers, 2012). The first is understanding what actual change took place and comparing that with what was expected in the theory of change. This can involve looking closely at the M&E data, interviewing key stakeholders, and confirming that the program inputs/interventions took place, when, and at what level. The second component is estimating what would have happened without the intervention, the counterfactual. There are several strategies to estimate the counterfactual, some of which involve observing or estimating effects among non-intervention populations. The “gold standard” here is the RCT, in which communities or households are randomly assigned to participate in the intervention or be in a control group not receiving the intervention, and the results are compared. The third component is identifying and ruling out alternative explanations that could explain the change. This could be the case if another intervention took place simultaneously to the one in question.

Evaluation Data and Mixed Methods

Real world programs are complex; it is therefore challenging for evaluations to capture the necessary data. Thus, evaluators should use and combine different evaluation frameworks, tools and techniques. Many evaluations use a “mixed methods” approach that combines quantitative data and qualitative data.

Several data collection approaches are used in impact evaluations. The choice of which quantitative or qualitative method to use varies with each evaluation. Examples of issues commonly studied using quantitative and qualitative data are given in Table 8, and examples of common data collection techniques used in evaluations are listed in Table 9. When a mixed methods approach is chosen, the researcher uses both.

Table 8. Common issues studied using quantitative and qualitative data

Quantitative	Qualitative
Household demographics	Processes in households, communities, and organizations
Targeting accuracy	Beliefs, norms, attitudes, and social relationships
Participation rates	Gender relations; women's status
Impact	Experiences with institutions
Intrahousehold decision making	Institutional and political dynamics
Service quality	Service delivery practices and staff attitudes
Test scores	Local satisfaction with program design, targeting, and administration

Source: Adato, 2011

The advantages of a mixed methods approach to data collection have been summarized by Adato (2011):

While surveys provide generalizable findings on what outcomes or impacts have or have not occurred, qualitative methods are better able to identify the underlying explanations for these outcomes and impacts, and therefore enable more effective responses. Qualitative methods also inform survey design, identify social and institutional drivers and impacts that are hard to quantify, uncover unanticipated issues, and trace impact pathways. When used together, quantitative and qualitative approaches provide more coherent, reliable, and useful conclusions than do each on their own.

Table 9. Data collection techniques for evaluations

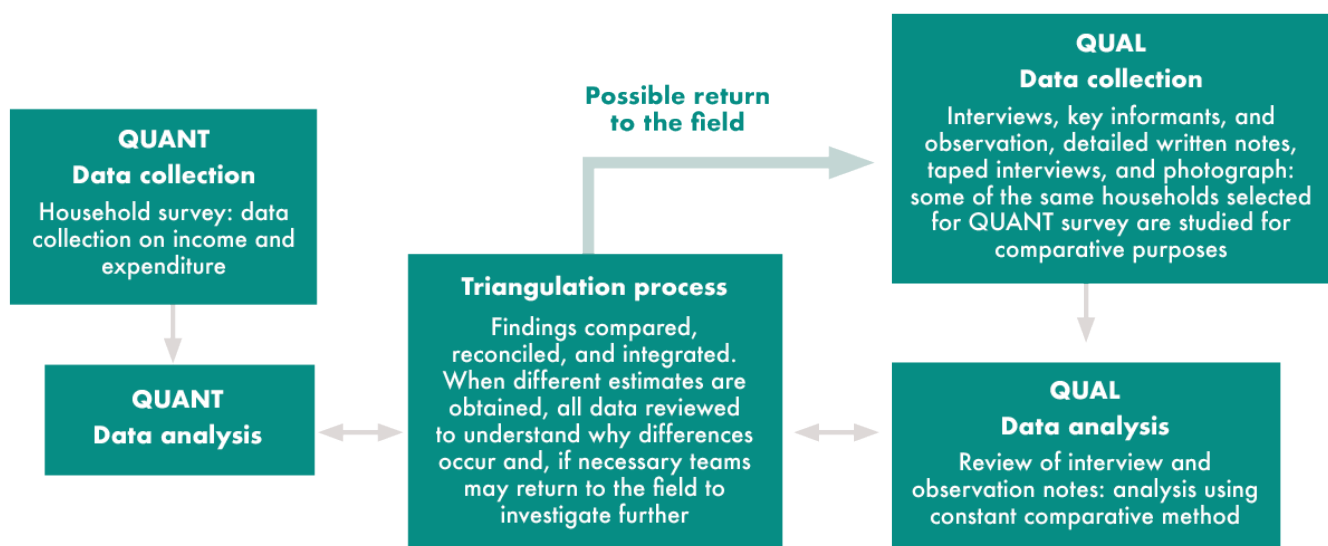
Quantitative	Qualitative
<ul style="list-style-type: none"> • Structured surveys of households, farms, users of public services, etc. • Structured observation guides • Anthropometric measures of height and weight • Anemia and HIV tests using blood sample collection and tests • Automatic counters (e.g., people entering a building) • Sociometric analysis ± ** • Geographic information system (generation and analysis of Global Positioning System maps)** • Program MIS on inputs and outputs data • Review of institution data—clinic records, school records, etc.** 	<ul style="list-style-type: none"> • In-depth interviews • Key informants • Participant observation • Non-participant observation** • Case studies • Client exit interviews** • Simulated patient studies • Video or audio recording** • Photography • Document analysis** • Artefacts • Group interviews (e.g., focus groups, community meetings)** • Participatory group techniques (e.g., Most Significant Change technique) • Internet surveys
<p>± Survey techniques to study group formation, how information spreads, identification of opinion leaders, and other patterns of social organization in a community or group.</p> <p>** Indicates that these techniques can be used both quantitatively and qualitatively. They are placed in the column where they are most commonly used.</p>	

Source: Adato (2011)

Triangulation

When using a mixed methods approach that delivers both quantitative and qualitative information, triangulation is key. Triangulation involves comparing results between the two data sources, cross-checking findings, and filling in missing information. More importantly, qualitative information can be used to “explain” or to cast light on quantitative findings. For example, a quantitative survey may find a sudden increase in pregnancies despite no change in family planning service quality or changes in demand. But the qualitative evidence from key informant interviews may reveal that there were significant stockouts of contraceptives during the period in question. Figure 12 provides a framework for triangulation.

Figure 12. Triangulation framework



Source: Bamberger, 2012

Evaluating Complex Interventions

Many public health programs often have several types of stakeholders and cover several contextual areas. They are often developed and executed in dynamic settings using integrated program models. Programs aimed at improving the lives of OVC, for example, can involve multiple interventions, multiple targeted outcomes, and even different target populations (e.g., children in different age groups.) Using only traditional methods for evaluating a program is often time-consuming, expensive, and insufficient to describe all the complex interactions between actors and outcomes. These kinds of programs can also have nonlinear or unknown outcomes or involve multidirectional pathways toward intended or unintended outcomes. These programs are often referred to as “complex” and require innovative ways of evaluating their outcomes and impact.

There is an emerging interest and discipline in methods of evaluating the outcomes and impact of complex programs. Four evaluation methods are briefly outlined.

Contribution Analysis

The reasons for conducting an impact evaluation or outcome assessment are usually to provide answers about the efficacy and efficiency of a program: Does it work? To what extent are the outcomes seen related to the intervention or program? **Contribution analysis**, a theory-based evaluation method, provides a different option for exploring these questions when a traditional experimental or quasi-experimental design is not possible, feasible, or desired (Mayne, 2001). National or population-level impact or long-term outcomes, such as behavior change, often cannot be understood or quantified for many years after the intervention, yet many donors and beneficiaries are eager to understand how the program is or is not meeting its intended objectives. Contribution analysis attempts to provide clear pathways from inputs to outcomes and yields information on whether a program is likely to produce the intended impact.

Mayne theorized that this method could be used to “address attribution through performance measurement” (Mayne, 1999). In contribution analysis, using a well-developed theory of change along with an assessment of all alternative or counterfactual theories for the outcome, one can create a “performance story” that can relatively assess a plausible attribution (Kotvojs & Shrimpton, 2007).

What makes contribution analysis unique is that it focuses on attribution of the intervention, and more actively seeks out counterfactuals during the process to help strengthen and validate the attribution. Although many donors require information on outcomes or impact, they most often also require data on inputs, processes, and outputs, which this method does not readily capture. Contribution analysis should be used in conjunction with monitoring data to ensure a complete picture of program performance.

Most Significant Change

The **Most Significant Change (MSC)** method has more recently gained traction as a valid and rigorous qualitative evaluation technique. It was developed by Davies in the 1990s to help evaluate a complex rural program in Bangladesh (Davies, 1998). Rather than focusing on measuring precise inputs, processes, and outputs, this method focuses on outcomes and impact. The MSC method is highly participatory in nature because the stakeholders themselves are involved in data collection, analysis, and sometimes dissemination. It is also a purely qualitative technique that does not employ any quantitative data or methods; rather, it relies on stories gleaned from stakeholders. Although the MSC approach is primarily of a qualitative nature, the qualitative data can be triangulated with quantitative monitoring data of inputs, processes, and outputs to provide a more robust picture of a program’s impact (Serrat, 2009). Although the MSC stories can include stakeholder estimates of quantitative impact, MSC cannot produce the kinds of data needed for economic evaluations.

Outcome Harvesting

Outcome harvesting is defined as the identification, formulation, analysis, and interpretation of outcomes to answer useful questions (Wilson-Grau, 2015). This method first collects evidence of what has been achieved and then works backward to determine whether and how the project contributed to the change, rather than measuring progress toward predetermined outcomes or objectives (as other evaluation methods do). Outcome harvesting is a fairly new approach to understanding complex programs, projects, and policies that aim to capture outcome-level indicators and the theories of change that contributed to the outcomes. This method involves six iterative steps:

- Stakeholders identify useful questions that help guide the process of the harvest.
- Through various primary and secondary sources, the “harvester” gathers data (e.g., through interviews, data sources, observations) to understand what changes have occurred due to the intervention or program and why beneficiaries feel those changes occurred.
- The “harvester” creates outcomes descriptions, based on the information gathered during steps 1 and 2, with the program stakeholders.
- The information is validated.
- The data are analyzed and interpreted.

- Information is disseminated and used for evidence-based programming (Wilson-Grau, 2015).

The information collected during this process is validated by other independent stakeholders at the individual and group levels to understand such questions about the program outcomes and impact as “what happened in this program?” and “why is it important?” This method is especially useful to understand how individual-level outcomes affect broader systemwide changes and impact. Outcome harvesting is well suited for understanding complex relationships and undefined or unknown outcomes and causal effects. Similar to the other emergent evaluation methods described here, this method is more useful for understanding outcomes than it is for understanding inputs, processes, and outputs. For this reason, outcome harvesting should be used in conjunction with quantitative or mixed methods approaches of monitoring to gather information. Outcome harvesting can also be used as a monitoring tool, an evaluation tool, or both, depending on how frequently it is used.

Participatory Evaluation

In addition to including field-based teams and other internal stakeholders in participatory planning and data collection methods, there are several methods for integrating external stakeholders in the M&E data collection and analysis processes. Participatory evaluation is a broader term and method than those described above. It involves many qualitative participatory methods, such as social and community mapping, scoring and ranking, storytelling, social network analysis, and diagramming. Many of these innovative and emerging qualitative approaches have been borrowed from other disciplines and have been found to be useful for monitoring and evaluating complex, integrated programs, such as population, health, and environment programs. However, participatory evaluation can also be used to collect quantitative data through the participatory practice of designing surveys, collecting data, and consensus building around results and analysis. Program beneficiaries and the communities in which they live are increasingly included in the entire programming process, beginning with program inception and ending with data dissemination and use.

Participatory evaluation is also useful for analysis and dissemination to a program/intervention’s beneficiaries. Community members and other stakeholders can assist in validating key findings and in identifying culturally appropriate ways to disseminate and present key findings relevant to and appropriate for different audiences.

Challenges in Impact Evaluation

This review has already alluded to many of the challenges in implementing an impact evaluation. In this section, a few of the more common challenges in carrying out an impact evaluation are highlighted.

Establishing a baseline is critical so that the evaluation can show change in the target indicators **and** attribute the change in some way to the intervention. However, establishing a baseline at the beginning of a project may be a challenge if the target beneficiaries have not been well identified, which is sometimes the case. Also, if the evaluation is being done externally, there may be issues of coordination between the implementing organization and the evaluators in terms of the start-up of activities. It is not uncommon for a baseline survey to be carried out after a program has started.

Timeframe of impact. The impact of an intervention may take a long time, often years, to be observed and yet donors and decision makers want to have information in a shorter period. The theory of change may need to be amended to include intermediate “benchmark” results that logically point to longer term impact.

Statistical modelling can also be used to project expected impact based on inputs that are known to lead to changes in the desired impact. For example, if the desired impact is a reduction in the fertility rate (which may require a household survey), one could look at the uptake of contraceptive use from service delivery statistics and use a model to estimate the impact on fertility.

Other factors and program influences. As we have noted, most development interventions in public health take place in an environment in which other programs and contextual factors come into play. One way to deal with this is to take into account such factors in the theory of change and control for non-program changes in the impact analysis. However, this requires gathering data and information about such factors.

Heterogeneity of impact. A given intervention has a different impact for different people (Rogers, 2012). What works for one subpopulation may not work for another subpopulation. In addition, even when we could expect similar impact across populations, the evaluation needs to consider the quality and “quantity” (dosage) of an intervention, especially across geographic areas. A program may have been well executed in one district but not in another. Similarly, there may be cultural factors at play between regions that may influence the success of a program, even when the program inputs and quality are the same.

Budget and resource constraints. Impact evaluations that are done “correctly” can be expensive and take time. Depending on the size of the program and its timeframe, an evaluation can cost millions of dollars and take years to complete. The data needed to carry out the evaluation may be judged to be too expensive to collect or staff resources may not exist to collect the data. When possible, program M&E data may be a good source of data when those data are deemed appropriate and of good quality. Alternatively, qualitative information on an intervention’s impact can be gathered through key informant interviews with stakeholders familiar with the situation. Last, program donors need to budget for evaluations along with the program itself, and the evaluation should be planned and initially included in the terms of reference for the intervention.

Case Study: Messaging and Malaria Treatment in Uganda

A common public health issue is that patients often do not finish their medication. The authors conducted an evaluation of alternative approaches to increasing adherence to malaria medication in Uganda.

The authors conducted a RCT in a high malaria prevalence area of Uganda, in which 2,641 households were given access to subsidized artemisinin-based combination therapies (ACTs) at local drug shops. In the study area, 35 percent of patients did not complete the full ACT treatment course. The study experimented with several ACT packages designed to increase adherence. The first approach to boosting ACT adherence rates was the one used by ministries of health and social marketing organizations in several African countries. It involved specialized packaging (“CAPSS” package) that has “pictorial instructions for illiterate patients, and a colorful, glossy design” that aimed to increase adherence by improving comprehension of dosing and by indicating the high quality of the drugs. Results showed that this approach had no significant effect on adherence. They also tested the impact of two inexpensive stickers for the ACT package with “short, targeted messages about adherence.” Both stickers highlighted the importance of adherence and encouraged patients to finish all pills. Results of the RCT found that both stickers increased treatment completion by roughly six percentage points to nine percentage points and reduced the number of remaining pills by 29 percent.

Source: Cohen & Saran, 2018

Case Study: Malaria and Child Mortality in Liberia

Malaria is the leading cause of morbidity and mortality in Liberia, with the entire population at risk. An evaluation was co-commissioned by Liberia's Ministry of Health and the National Malaria Control Program, and the United States President's Malaria Initiative, to report on the impact of these investments on morbidity and mortality among children under five years of age during the period 2005 to 2013.

Evaluation Design

The evaluation was based on a before-and-after assessment, which used a plausibility evaluation design that measured changes in malaria intervention coverage, malaria-related morbidity, and all-cause child mortality (ACCM) in those under five years of age, while accounting for other contextual determinants of child survival during the evaluation period. ACCM was used as the primary measure of impact. Further analyses investigating the relationship between household insecticide-treated bed net (ITN) ownership and malaria parasitemia were conducted using multiple logistic regression to support the plausibility design.

Data Sources

Data came from the following five large population-based household surveys: the 2005 Malaria Indicator Survey (MIS), 2007 Demographic and Health Survey (DHS), 2009 MIS, 2011 MIS, and the 2013 DHS. These national survey data were supplemented by data from the National Malaria Control Program and the Liberia health management information system (HMIS), the 2009 and 2013 Liberia Health Facility Surveys, World Bank data, and country project reports.

Interventions

The proportion of households with at least one ITN and/or indoor residual spraying in the last 12 months reached 59 percent by the end of the evaluation period; however, most of this was attributed to increased household ITN ownership. Overall, household ownership of one or more ITNs grew steadily during the evaluation period; however, household access to an ITN (defined as one ITN for every two people in the household) only reached 22 percent by 2013.

Implementation of intermittent preventive treatment in pregnancy (IPTp) began in Liberia in 2005. Coverage of IPTp increased substantially from four percent in 2005 to just under 50 percent by 2013, although coverage reached 45 percent by 2009 and then remained relatively unchanged between 2009 and 2013.

Morbidity

Malaria parasitemia prevalence among children ages 6 to 59 months measured through rapid diagnostic tests saw a substantial decline, from 66 percent in 2005 to 37 percent in 2009, but then increased to 45 percent in 2011. However, malaria parasitemia measured through microscopy showed a slight decline, from 32 percent in 2009 to 28 percent in 2011 among children ages 6 to 59 months. (No data were available from the beginning of the evaluation period.)

Data from the HMIS showed that the number of confirmed malaria cases among children under five years of age and people five years of age and above gradually increased between 2009 and 2012, before declining in 2013. Overall, trends in malaria parasitemia showed an overall decline during the evaluation period, whereas confirmed malaria cases suggested a decline toward the end of the evaluation period.

Mortality

ACCM declined by 14 percent during the evaluation period, from 109 to 94 deaths per 1,000 live births between 2002 and 2006 and between 2009 and 2013. When assessing trends in ACCM by age group, the greatest relative decline between the two survey periods was among infants (25%). Declines were observed in all age groups, except child mortality (mortality between ages 12 and 59 months); however, the decline was only significant among infants. ACCM was greater in rural areas in the 2009 to 2013 period, and greater relative declines were observed in urban areas compared with rural areas between the two survey periods. Altogether, the data suggested a small decline in ACCM during the evaluation period that was mainly due to a decline in infant mortality.

Contextual Factors

The evaluation included a comprehensive review of contextual determinants of child survival that could have contributed to the observed changes in mortality during the evaluation period. Among the social and economic determinants, improvements were seen in access to health facilities and the overall health system infrastructure, GDP per capita, total health expenditure per capita, women's education and literacy, and household asset ownership (telephones). Moreover, improvements in several maternal and child health interventions were observed during the evaluation period, including antenatal care (ANC) attendance, tetanus toxoid vaccination, delivery at a health facility and with a skilled attendant, immunization coverage, and vitamin A supplementation for children ages 6 to 59 months.

Statistical Analysis

Multiple logistic regression analyses assessing the association between household ITN ownership with parasitemia prevalence (via rapid diagnostic tests) among children ages 6 to 59 months showed a protective effect that fell just short of statistical significance. However, a second similar model examining the association between household ITN ownership by age of the net and parasitemia prevalence showed a significant protective effect on parasitemia prevalence for children living in households that owned an ITN for 0 to 6 months. Both models demonstrated that other variables, including age of the child, region of the country, place of residence, malaria risk, and household wealth, to be significantly associated with parasitemia prevalence.

Source: Malaria Impact Evaluation Group, Liberia & Herrera, 2018.

Further Reading

General economic evaluation education online courses:

- Global Health eLearning Center. (2011). Economic evaluation basics course. Available from <https://www.globalhealthlearning.org/course/economic-evaluation-basics>
- United States National Library of Medicine. (2002). Health economics information resources: A self-study course. Available from <https://www.nlm.nih.gov/nichsr/edu/healthecon/>
- Bamberger, M., Rugh, J., & Mabry, L. (2012). Real world evaluation: Working under budget, time, data and political constraints. 2nd edition. Los Angeles, CA, USA: Sage Publications.

5. COST-EFFECTIVENESS ANALYSIS

In Chapter 2, we noted that **CEA** is the measurement of “the cost to achieve one unit of the desired health effect.” In this chapter we delve in more detail about CEA, what it means, and its limitations.

CEA compares an intervention’s costs with its outcomes. CEA expresses outcomes by a measure of some health outcome unit, such as the number of malaria cases prevented or the number of lives saved. Therefore, the CEA metric is the cost per health outcome unit achieved, i.e., the cost per malaria case prevented or cost per life saved.

Although there are several possible uses of CEA, at the risk of oversimplification, we list two uses. First, CEA can assist with the achievement of *technical efficiency* by helping with the choice of an intervention or interventions based on the lowest cost per unit of health benefit achieved. Alternatively, CEA can be used for *allocative efficiency*: choosing an intervention or interventions to achieve the maximum benefits at the population or societal level.

CEA involves two kinds of estimates. First, we need the numerator (cost) of an intervention or program, and second, we need to measure effectiveness in the denominator (health outcome.) Measurement of the effectiveness most often comes from an impact evaluation of the intervention or program, and the choice of the outcome indicator depends on what the evaluation could measure and/or the objectives of the program. Similar projects may measure impact differently. For example, a family planning program could measure impact in terms of the number of new users of modern family planning methods. Alternatively, the family planning program could measure impact by the number of births averted. Because impact is always measured in terms of a change (number of new users, births averted), a baseline should be established to measure the change from the initiation of the program to some point in the program when the impact is measured. This is represented conceptually in Figure 13.

Figure 13. Conceptual representation of measuring impact of an intervention from a baseline



Alternatively, one can have two programs that focus on the same outcome but with different approaches. For example, two programs could focus on reducing obesity, one through an exercise program and the other through diet alteration. We can compare the effectiveness of the programs if we choose the same outcome

Learning Objectives

What are the main measures used in cost-effectiveness analysis?

How are event trees or decision trees used in cost-effectiveness analysis?

How do you conduct cost-effectiveness analysis of multiple interventions?

What decision rules can one use?

How can you measure the efficiency of health units?

measure (i.e., change in the mean body mass index of the target population). A more complete discussion of approaches to measuring effectiveness is in Chapter 4.

Outcomes in CEA can be defined narrowly or broadly, although broad definitions are more appropriate for public health policy decisions. Narrowly defined effects are those that are intermediate in nature and that can be easier to capture, such as immediate increases in the use of bed nets for malaria or decreases in the number of malaria cases. More broadly defined effects, for example, are reductions in malaria deaths or DALYs.

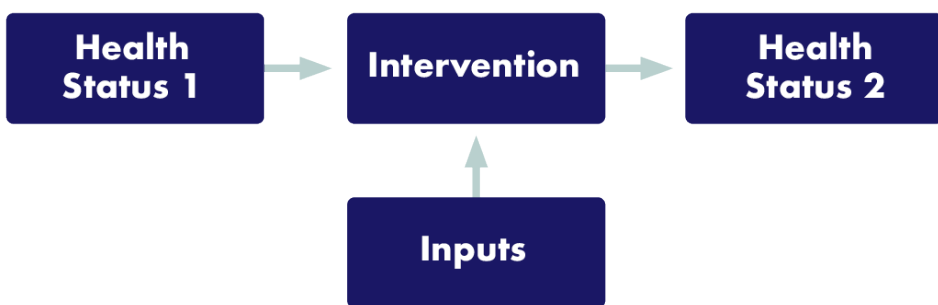
CEA can be used to inform a variety of perspectives. For example, a CEA study of a family planning program may evaluate the cost per user added across different clinics that have different cost structures. This example could represent the perspective of a program that needs to decide how to provide services most effectively (a question of technical efficiency). A different CEA study could look at the cost per client added using different interventions to promote the use of family planning. The second example represents the perspective of a government agency that needs to set budget priorities based on the impact of interventions across the health system (a question of allocative efficiency).

Although the computation of CEA metrics requires a single measure of effectiveness, many public health interventions may result in multiple outcomes. For example, family planning programs not only reduce unintended pregnancies, they also benefit the health of contracepting women by reducing the incidence of the negative health consequences of a pregnancy and/or delivery. Although the focus of a CEA could be on reducing the number of unintended pregnancies, the possibility of other effects means that the intervention could also be compared with other interventions that target these other effects. Of course, this would require a separate CEA for the other effects. The additional analysis could show that the intervention is the most cost-effective alternative for the secondary benefit and thus provide additional decision support for the intervention. If the intervention dominates for one outcome but not for the other, the data could be presented to decision makers to decide on the trade-off.

Alternatively, the analysis could be altered to be a **CUA**. In CUAs, the multiple indicators of effectiveness are combined in a single metric by weighting the various benefits. The weights can be established by key stakeholders.

The other measurement we need for CEA is costs, which we treated in more detail in Chapter 2. A common misconception pertaining to costs is that we can directly measure the cost of an outcome. Although one objective of CEA is to know what it costs to achieve a unit of some health outcome, it is important to remember that we can only measure the costs of *inputs* involved in the intervention or program. Therefore, we measure the cost of staff, medications, facilities, training, etc. How the inputs influence the outcomes is hypothesized through the intervention's theory of change or log frame. A simple theory of change is shown in Figure 14.

Figure 14. A simple theory of change linking cost inputs to outcomes



- Newly trained staff
- Newly available medications
- Refurbished facilities
- Additional administration

In this sense, then, the costs of the outcomes are estimated indirectly.

Summary Measures

CEA uses two summary measures. One summary measure is the ratio of net programmatic costs divided by net program effects. Programmatic costs are program costs minus the cost of illness averted by the program in the case of a disease control project. If there are two interventions with the same outcome measure, the CEA ratios can be compared to see which program yields the lower cost per unit of outcome.

Alternatively, two interventions affecting the same health outcome can be compared in terms of incremental costs of one program compared with the other, divided by the incremental effects of one program compared with the other. This approach uses a measure called the **incremental cost-effectiveness ratio (ICER)**, which is defined as

$$ICER = \frac{\Delta c}{\Delta e} = (c_2 - c_1) / (e_2 - e_1)$$

in which c_2 is the cost of the intervention, e_2 is the effect of the intervention, c_1 is the cost of the status quo (or an alternative intervention), and e_1 is the effect of the status quo (or an alternative intervention).

An example of the two CEA measures is presented in Table 10. In this example, two interventions are being assessed for their cost-effectiveness in averting HIV infections.

We see that intervention 2 has a lower cost per infection averted than does intervention 1: \$100 versus \$200.

Alternatively, the ICER for intervention 1 versus intervention 2 is calculated as: $(\$100,000 - \$50,000) / (1000 - 250) = \67 . This means that the cost for every additional HIV infection averted by intervention 2 is \$67.

Table 10. Example of cost-effectiveness assessment

	Interventions	
	1	2
Cost	\$50,000	\$100,000
HIV infections averted	250	1000
CEA ratio (cost per infection averted)	\$200	\$100
ICER	\$67	

Decision Rules

The use of the CEA ratio to inform decisions on which intervention to choose is straightforward: choose the intervention with the lowest cost per unit of impact. However, the use of the ICER to make decisions is more esoteric. Its interpretation falls on “WTP.” This amounts to a judgment about whether society would be willing to pay for the value of a given effect. Essentially, it means putting a value on an incremental outcome. For example, what is the value of a woman not having an unwanted pregnancy? Or for a child to be immunized? Or for a person not to be infected with TB or HIV?

Another approach to using the results of a CEA is to apply a decision rule, such as in Table 11.

Table 11. Decision rules in cost-effectiveness analysis

Net effects	Net costs positive	Net costs zero or negative
Positive	CASE 1. Cost-effectiveness = net costs/net health effects. Select most efficient programs for improving health (lowest ratios).	CASE 2. Program economically valuable. Should generally be implemented.
Zero or negative	CASE 3. Program benefits offset by morbidity and inconvenience. Program generally should not be implemented.	CASE 4. Cost effectiveness = net costs/net health effects. Select most efficient programs for containing costs (highest ratios).

Source: Shephard & Thompson, 1979

Many stakeholders in public health can have different views about cost-effectiveness values. For example, the donor, a department head in the MOH, a citizen in that country, the WHO, and the researcher or evaluator could have diverse notions about WTP. The process of interpretation of the ICER requires the researcher or evaluator to be aware of the perspective and to be transparent about it in the CEA.

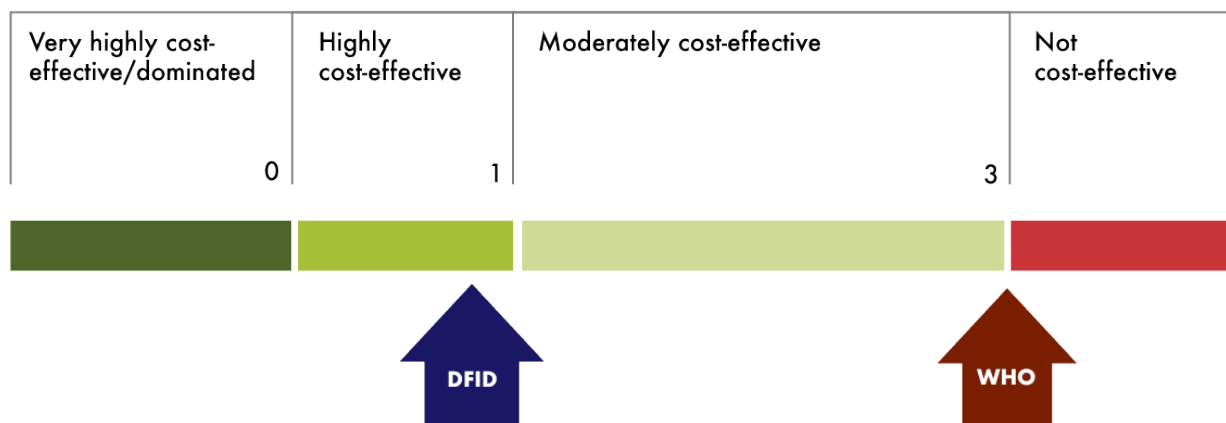
No consensus exists on the value of an ICER being cost-effective (Robinson, Hammitt, Chang, & Resch, 2017). The general practice in public health is to define WTP as some factor of per capita income for every additional DALY in a program’s country. This metric can generally be used for infections averted or other natural events.

The Department for International Development (DFID), the United Kingdom’s development agency, has a standard that the cost of the ICER of all interventions that save or avert one DALY below one unit of per capita income in the recipient country is cost-effective (e.g., US\$10,000 in South Africa, US\$600 in Mozambique). WHO has a guideline that the ICER threshold for an intervention should be below three times per capita income to be cost-effective (e.g., US\$30,000 in South Africa, US\$1,800 in Mozambique) (WHO, 2001). USAID does not have a standard in its guidelines.

The number lines in Figure 15 provide guidance on interpreting the ICER. If the ICER is negative, it is called “dominated”; a negative ICER suggests that the intervention would save resources and should be done. One important qualification is that if e_1 is greater than e_2 , which means that the reference has more of a positive effect than the intervention—the denominator in the ICER ratio is negative. If the net change in costs is positive and the net change in effects is negative, the ICER would be negative, but this is a false signal. If the net effect were negative, the researcher should report that the intervention did not have the intended effect and should not continue with a CEA study.

If the ICER is positive and it is less than one unit of per capita income, this could be called highly cost-effective. The debate begins if the ICER to avert one DALY is above one unit of per capita income. If the ICER is between one and three times per capita income, the intervention could be considered moderately or weakly cost-effective. WHO’s more liberal standard of cost-effectiveness is not surprising, because WHO promotes health and argues that seemingly more expensive interventions are appropriate (Baltussen, Adam, Tan-Torres Edejer, Hutubessy, Acharya, Evans, ... World Health Organization, 2003).

Figure 15. Number line to represent ICER threshold to save one DALY by unit of per capita income



Source: Baltussen, Adam, Tan-Torres Edejer, Hutubessy, Acharya, Evans, ... World Health Organization, 2003

We noted above that a program’s outcomes can be measured by two different indicators. However, an important caveat in conducting CEA is that outcomes in natural units cannot be combined and should be considered separately. As noted by the CDC:

For example, a physical activity program may have two intended effects: lowering blood pressure and decreasing body mass index. Because these two effects can’t be combined in a cost-effectiveness analysis, the summary measure for the analysis would

be cost per 1 percent reduction in blood pressure and cost per 1 percent decrease in body mass index. However, the cost in these two summary measures is the same, so the ratios are somewhat misleading. This makes cost-effectiveness ratios using natural units difficult for policy-makers to translate. (Centers for Disease Control and Prevention, n.d.)

A useful construct in thinking about CEA, especially the ICER, is the decision tree. A decision tree or events tree consists of mapping all the possible states that a beneficiary or group of beneficiaries could achieve under various interventions (or lack thereof). For example, if a man were to be medically circumcised, he could either become HIV-infected or stay healthy. Likewise, if he were not circumcised (the status quo branch of the tree), he could either become HIV-infected or stay healthy. This creates four branches of the tree, as shown in Figure 16.

Consider an intervention to provide VMMC as a strategy to reduce the transmission of HIV. As we move through the decision tree in this example, we need the probabilities of someone on the VMMC side who is either HIV-infected or staying healthy. These are the only two states on the VMMC branch, so the sum of their probabilities needs to equal one. We should also have similar calculations on the bottom branch. These probabilities are essential to the calculation of the incremental costs and effects and, eventually, the ICER. The probabilities express the degree of certainty that an event or the health state will happen (being HIV positive or negative in this example). Chapter 4 reviews evaluation approaches that can provide these probabilities.

In the example in Figure 16, the cost of the VMMC intervention is \$100 per client and the probability of becoming infected with HIV is 0.06. The cost of treatment for HIV is \$1,000 in the event of infection.

Figure 16. Decision or events tree

Decision tree	VMMC	Stay healthy	Probability: 0.94	Cost: \$100	Effect 0.94 * \$100=\$94
		HIV infection	Probability: 0.06	Cost \$100+\$1000	Effect 0.06 * (\$1100)=\$66
	Status quo	Stay healthy	Probability: 0.85	Cost \$0	Effect 0.85 * \$0=\$0
		HIV infection	Probability: 0.15	Cost \$1000	Effect 0.15 * \$1000=\$150

Conversely, for the status quo branch, the probability of infection without circumcision is higher, at 0.15, and there is no intervention cost; however, the cost of treatment for HIV is \$1,000, as before.

For each of the four ultimate states (or effects), we calculate the cost by multiplying the probability of being in the state by the cost. For example, the expected cost for a man circumcised who becomes infected is the sum of the intervention cost and treatment cost (\$100+\$1,000) times the probability of infection (0.06), which equals \$66. The effect or benefit of the intervention is measured as the probability of staying healthy after circumcision: 0.94. Therefore, in this example, the net cost of VMMC is the sum of the expected costs of both states, which is \$94 plus \$66 = \$160.

Similarly, the net cost of the status quo/do nothing is the sum of \$0 and \$150 = \$150. The effect or benefit of the intervention is measured as the probability of staying healthy without circumcision: 0.85.

If we substitute these results in the ICER equation, we calculate:

$$ICER = \frac{\Delta c}{\Delta e} = (\$160 - \$150) / (0.94 - 0.85) = (\$10) / (0.09) = \$111.11$$

The interpretation of the ICER in this example is that every HIV infection averted costs of \$111.11.

Choice of Comparisons

In almost all cases, CEA involves comparing two or more interventions, whether using the CEA ratio or the ICER. An important design component of a CEA is what interventions to compare. Suppose we want to expand a program to distribute insecticide-treated bed nets to households to reduce cases of malaria. The program aims to increase the percentage of households with bed nets to 30 percent. But suppose that before the program, 25 percent of households used bed nets. The ICER would be different if we compared a program that compares the increase with 30 percent to the current 25 percent level than if we compared the 30 percent program with a situation where no one is using bed nets. In one case, we are calculating the cost-

benefit of a small increment in bed net use over the status quo, and in the other comparison, with essentially no program at all.

The choice of the comparison depends on the use of the CEA. If the CEA is being conducted merely to decide if an intervention should be undertaken relative to a current practice, then the current practice can serve as the comparison. In this case, the decision is a marginal one, essentially aimed at improving *technical efficiency*. In this case, the CEA does not evaluate the efficiency of the current mix of interventions; it only looks at the marginal changes of introducing a new intervention. Therefore, although it can tell you whether a new intervention is worth doing, it is silent on whether the current situation was worth doing.

If the CEA aims to choose an intervention that maximizes a population's health benefits, then a counterfactual or "null scenario" is more appropriate. This more general question is one of achieving *allocative efficiency*. As Baltussen, Adam, Tan-Torres Edejer, Hutubessy, Acharya, Evans, ... & World Health Organization (2003) state, "From the starting point of the situation that would exist in the absence of the interventions being analyzed, the costs and effects on population health of adding interventions singly (and in combination) can be estimated, to give the complete set of information required to evaluate the health maximizing combination of interventions for any given level of resource constraints."

Other Considerations

Sensitivity Analysis

Despite the best efforts of researchers to provide good estimates of costs and efficiencies, there are often areas where there may be uncertainty about a parameter's value or where the value could be subject to change. For example, the efficacy of an intervention in treating a disease or in preventing an infection may be based on a clinical trial; however, the trial could produce a range of efficacy or we could expect the efficacy to differ between one population and another, for example. Or, the costs of drugs could be expected to change over time.

In these cases, we can perform different forms of sensitivity analysis. In a **one-way sensitivity analysis**, a single parameter is changed in value over a range and it is fed into the calculations. In the VMMC example above, suppose the cost of the intervention changed. In Table 12, we show how sensitive the ICER is to this parameter. Changing the cost of the intervention by a mere five percent or \$5 cuts the ICER in half, from \$111 to \$55.

Table 12. One-way sensitivity analysis of VMMC

Cost of intervention	ICER
\$95	\$55
\$100	\$111

In a two-way sensitivity analysis, two parameters are varied at the same time. Using the VMMC ICER example, suppose we change not only the cost of the intervention, but also its efficacy as measured by the probability of

remaining HIV-free if circumcised. Table 13 illustrates the impact on the ICER for two values of the cost and two values of the efficacy of the intervention. Again, we see not only how sensitive the ICER is to each parameter, but also the large range of estimates, from \$55 to \$428.

Table 13. Two-way sensitivity analysis of VMMC

Cost	Efficacy	
	0.94	0.92
\$95	\$55	\$357
\$100	\$111	\$428

Sensitivity analysis can be useful in two ways. First, it can guide decisions based on the value of the parameter. In the VMMC one-way example above, suppose that the government is considering the VMMC program, but it does not have the funding to justify an investment when the ICER is \$111 and when the intervention cost is \$100 per client compared with another intervention. Table 12 shows us that if the intervention cost can be reduced by five percent, the ICER would be much lower and the intervention could be judged cost-effective compared with other interventions.

Second, sensitivity analysis helps identify which beneficiary populations have the lowest cost-effectiveness. This could be reflected in the cost per client (\$95 versus \$100) or in the efficacy (0.92 versus 0.94).

Linearity

Most CEA assumes, at least implicitly, that health programs are divisible and exhibit what economists call “constant returns to scale.” This means that if the size of the program is doubled, the benefits would double. This is referred to as linearity. This assumption could derive from practical data limitations in conducting CEA that may not permit calculating CEA ratios at different levels of scale, in which case the CEA ratio is implicitly an average. Another reason could be that the intervention or program could be indivisible. Because a CEA ratio is the cost per unit of benefit (i.e., \$100/case averted), the linearity assumption implies that the cost per benefit for a small program is the same as for a large program.

Most economic models of production do not assume linearity. There can be “economies of scale,” especially if fixed capital costs of the intervention are high and these costs can be spread over more and more beneficiaries as a program scales up. Or, there could be “diminishing returns to scale,” such as when a program expands to reach clients who are harder and harder to reach.

Linearity can pose a problem for the decision maker in a couple ways. First, if the CEA is being used to cost a scale-up, it could overestimate the costs if they are non-linear due to economies of scale, or it could underestimate them if unit costs escalate with larger programs. Second, when choosing between multiple programs at different levels of scale, if the costs are non-linear, using a CEA ratio that is not influenced by scale could lead to an incorrect decision.

Multiple Interventions

A fundamental issue in CEA is defining the intervention (Murray, Evans, Acharya, & Baltussen, 2000). This is not as straightforward as it may appear; it leads us to a discussion of CEA of multiple interventions. In studies

of multiple interventions, the comparison is almost always the counterfactual/do-nothing scenario. For example, consider alternative interventions to reduce malaria cases. They could include environmental management, behavior change communication, indoor residual spraying, ITNs, and malaria prophylaxis. They could also include combinations of these interventions, which would introduce interactions between them, i.e., ITN uptake reduces the need for prophylaxis. In the case of multiple interventions, the costs and benefits can be represented in a league table, such as Table 14. In this fictitious example, the lowest CEA ratio is 0.4 for the combined prophylaxis and ITN program. Of course, decision makers can choose an intervention with a higher CEA ratio if other factors preclude the selection of the intervention with the lowest ratio.

Table 14. League table for malaria interventions

	Costs	Benefit	CEA ratio
Prophylaxis and ITNs	200	450	0.4
Prophylaxis only	100	200	0.5
ITNs	150	250	0.6
Behavior change communication	200	300	0.7
Indoor residual spraying	600	500	1.2
Environmental management	500	400	1.3

Comparability Across Populations

Decision makers may want to use a CEA conducted for one population to guide decisions about adopting the intervention for another population. It is tempting for decision makers in resource poor settings to use results of a CEA in one country to guide decisions in another country because the cost and time required to conduct a good CEA, especially the evaluation (effectiveness) component, may be prohibitive. For example, results from a CEA of ART for HIV in South Africa could be used to initiate a similar program in Uganda.

But using data from one country in another one should be done with caution. First, at the very least, the CEA should use a counterfactual null case for comparison because using a status quo for comparison in one country may not be relevant to the country under consideration. The current situation could be very different. Second, resource costs for such inputs as staff time, medical supplies, and pharmaceuticals will be different from one country to another. It is possible to adjust such costs to take differences in unit costs into account between countries by using a price index, per capita income, or other approaches. Third, the effectiveness of the same combination of inputs across contexts will differ, although this can be attenuated by choosing the counterfactual null case, as noted above. However, as noted by Murray, et al., “Nevertheless there are clear limits to the comparability across populations of the counterfactual null set” (Murray, Evans, Acharya, & Baltussen, 2000).

Limitations of CEA

Although CEA is a very powerful tool to assess the value for money of discrete health interventions, other tools and methods are sometimes required, depending on the question the analyst is asking. Beyond the challenges of identification and measurement of the cost and effect of a health intervention, there are limitations to CEA, which follow:

Scale, feasibility, and affordability. Most CEA studies are incremental and, as such, do not address scale (Marseille, Larson, Kazi, Kahn, & Rosen, 2015; Boardman, Greenberg, Vining, & Weimer, 2011). If a pilot program has an ICER that is deemed to be low and cost-effective, the assumption that the costs will remain the same if the program were scaled up may not hold unless the unit costs are linear, as we discussed earlier. Also, a new larger program could have issues with administration, targeting, and supply chain, among others, which increase the cost per client/patient served. When scaling up a successful pilot, the recommendation is to calculate a new ICER for the new program, incorporating all the potential costs of the intervention's expansion and using sensitivity analysis to calculate from a domain of rising cost inputs.

Moreover, CEA does not consider feasibility. If the analysis shows a high degree of cost-effectiveness (even to the point of cost savings), the MOH may not have the administrative or technical capacity to implement the program. In many settings, the lack of health workers is a significant impediment to the expansion of clinical services. For example, in Malawi and Zambia, the initial rollout of HAART in the mid-2000s was delayed because neither health system had sufficient frontline health workers (medical doctors, nurses, and midwives).

CEA also does not deal with the issue of affordability. If a developing country has a consultant advising that an intervention is cost-effective, the country may not have the foreign exchange or financial resources to implement it. In global health, many countries have used global mechanisms, such as the Global Alliance for Vaccines and Immunizations, the Global Fund to Fight AIDS, Tuberculosis and Malaria, the Gates Foundation, and bilateral donors, to purchase vaccines, commodities, and medications. As foreign assistance in global health declines, the sustainability of successful programs becomes increasingly important. Apart from CEA, the examination of fiscal space, budget maximization, and the broader tax effort become paramount.

Ethical Considerations

There have been longstanding concerns that CEA would be used with an “audit mentality”: if an intervention was deemed too expensive, it would not be implemented (Morrow & Bryant, 1995). In the early 1990s, the state of Oregon in the United States organized a list or league table of ICERs for clinical procedures that it wanted to use to help its budgeting of Medicaid expenditures, but the state needed the Department of Health and Human Services (HHS) to approve the document for its use. Although the list was only meant to be an additional tool to aid decision making, the social uproar that the document sparked prevented its submission to HHS by Oregon, and the list was never used (Allhoff, 2005). The Patient Protection and Affordable Care Act of 2010 (also known as “Obamacare”) in the United States prohibits the use of thresholds by the Patient-Centered Outcomes Research Institute in HHS to determine whether a procedure is cost-effective.

No one in health economics would suggest that a league table or threshold should be the final or deciding metric to determine whether one intervention should be funded versus another. These examples highlight how politically delicate any discussion of cost-effectiveness can be. Economic evaluation of programs provides

evidence of value for money, but there may be other reasons—reasonable or not—to implement an expensive intervention. The main point is not to fall in the trap of using CEA results with an audit mentality.

Complicated Logic Models

One of the key assumptions of CEA is that one program has one effect and other programs (and their effects) are mutually exclusive from each other. In some program areas (e.g., OVC in HIV), a family unit could receive a set of interventions at the same time, which renders attribution of an effect to one intervention nearly impossible. In this situation, qualitative comparative analysis is an approach that creates pathways, or a set of interventions linked to one effect. A pathway is a configuration of two or more conditions (interventions or programs) that are necessary but insufficient to achieve the outcome. This process allows each pathway to differ on one intervention that would become subject to comparison in the ICER calculation. Although qualitative comparative analysis was originally designed for research with a small number of cases, it can be used to help the researcher or evaluator understand the causal arrows and theory of change.

Equity and Gender

CEA can be a very powerful tool to assess the value for money of discrete health interventions and to compare them against one another. Yet important concepts, such as equity or gender, do not fit well in CEA. Several attempts in advanced industrialized democracies have been made to incorporate equity or gender by using different metrics and methods. Efforts to weigh the effects toward the poor or females often cloud the analysis instead of clarifying the intervention's effects on equity and gender.

If there is interest in the effect of an intervention on health or social equity, one approach is to perform a distributional or extended cost-effectiveness analysis (DCEA or ECEA) (Asaria, Griffin, & Cookson, 2016). DCEA or ECEA adds only one additional step to the CEA five-step process: if the data are available, the evaluator needs to establish the distribution of whatever is being studied by population quintile and compare the data before (baseline) and after (the program's effect) by quintile to isolate the effect of the intervention on the lower quintiles. Gender analysis follows a similar approach. Perhaps the emphasis on special steps to address equity or gender in CEA is not salient in global health. Many programs in global health directly target the lower quintiles of the income distribution and females in their design, so the additional analysis may add little. One potential pitfall of DCEA or ECEA is that both require a great deal of data that the program may not have collected.

Data Envelope Analysis

DEA is another cost-effectiveness approach that is designed to help measure and improve the performance of an organization. The measurement of efficiency of an organization or organizational unit (such as a service delivery point) can be complex, especially in situations where there are multiple inputs and multiple outputs (Bhat, Verma, & Reuben, 2001). Originally developed by Charnes, et al. (1985), DEA uses linear programming to measure efficiency. It focuses on technical efficiency and compares the relative efficiency of organizational "units" in a health organization where the units perform similar tasks. DEA can also be used in other sectors, such as banking or manufacturing. Unlike CEA, DEA uses physical inputs and outputs to compare technical efficiency across units rather than costs. There may be considerable differences in the way in which individual

units combine inputs to produce outputs caused by the technology they have used, their geographic location, or other factors, such as the profile of the catchment population.

DEA uses efficiency scores for all units being analyzed to estimate an efficiency “frontier.” The efficiency of each unit is compared with that of the most efficient unit, showing how much inefficient units need to reduce their inputs or increase their outputs to become more efficient.

For example, consider the case of a health service organization that has multiple clinics. The clinics could offer multiple services (family planning, birth delivery, immunization, well baby care, ANC, etc.) and use multiple inputs (clinical staff time, clinic facilities, drugs, etc.). How then can we rate the efficiency of the clinics and rank them?

Start with a simple example. If the clinics in our example have a single input and a single output, the efficiency of converting the input into an output for each clinic is measured by:

$$\text{Efficiency} = \frac{\text{output}}{\text{input}}$$

Suppose we measure efficiency by the number of patients seen in the ANC unit per month per staff member. Table 15 gives this fictional example.

Table 15. Simple DEA example

Clinics	Number of staff	ANC patients served per month	ANC patients/staff	Relative efficiency
Clinic 1	10	150	15	0.5
Clinic 2	20	250	13	0.4
Clinic 3	16	100	6	0.2
Clinic 4	25	500	20	0.6
Clinic 5	30	1000	33	1.0

In this example, we see that Clinic 5 is the most efficient and Clinic 3 is the least efficient: 33 patients per staff versus six patients per staff, respectively. The column marked “relative efficiency” is calculated by dividing each clinic’s efficiency score (patients/staff) by the score of Clinic 5, the most efficient. For example, this means that Clinic 3’s efficiency is 20 percent that of Clinic 5. If Clinic 3 were to increase its efficiency to the level of Clinic 5, it could serve 533 patients per month (100/0.2) rather than the 100 it currently serves. Alternatively, Clinic 3 could serve the same 100 patients with only about three staff if it were as efficient as Clinic 5.

Let’s turn to a more complex example in which the clinics offer two services: ANC and family planning. Table 16 shows the number of ANC patients seen per month and the number of family planning patients seen per month. For simplicity, in this example, the same staff serve both types of patients.

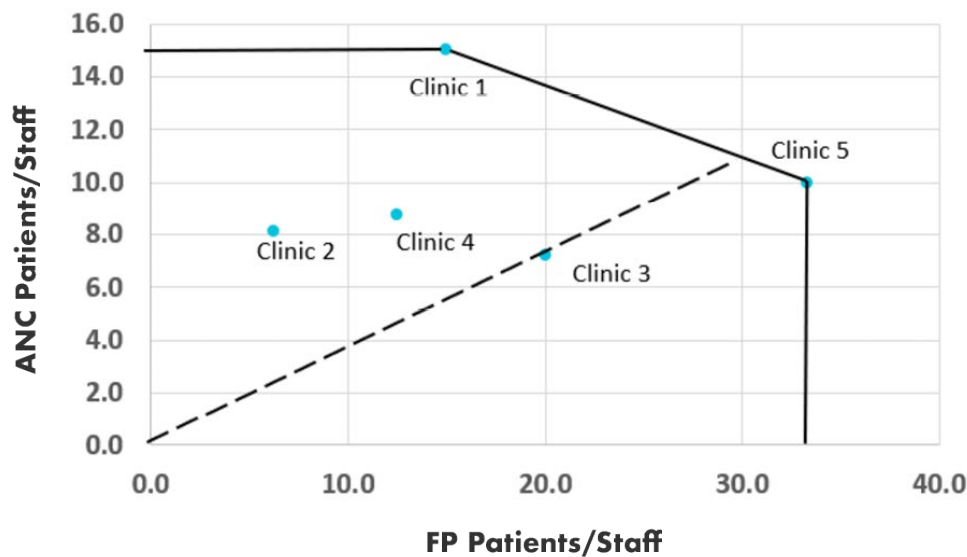
Table 16. Two dimensional DEA

Clinics	Number of staff	ANC patients served per month	ANC patients/staff	Family planning patients per month	Family planning patients/staff
Clinic 1	10	150	15.0	150	15.0
Clinic 2	20	250	12.5	175	8.8
Clinic 3	16	100	6.3	130	8.1
Clinic 4	25	500	20.0	180	7.2
Clinic 5	30	1000	33.3	300	10.0

Now the efficiency of each clinic is calculated with the same indicators (patients/staff) for each of the two services. We see that Clinic 5 is the most efficient for ANC services, but Clinic 1 is the most efficient for family planning services.

These data can be graphed as shown in Figure 17, which is called a “frontier graph.” The frontier graph represents the positions of each clinic in terms of patients per staff. The line drawn between Clinics 1 and 5 shows the “efficiency frontier.” It is derived from the most efficient clinics and represents the best achieved performance. It can be used as a standard with which to measure the performance of the other clinics. We see that clinics 2, 3, and 4 are inside the frontier. A dotted line from the origin through Clinic 3 to the intersection of the frontier shows how the potential efficiency of Clinic 3, for example, could be increased if it were to proportionally increase its patients served while keeping its inputs the same. Alternatively, Clinic 3 could reduce its inputs while keeping its patients served the same.

Figure 17. Frontier graph



There are two versions of DEA. The first version assumes “constant returns to scale” technology (CRS). This is appropriate when all units are operating at an optimal scale. CRS means that a unit’s outputs increases in the same proportion as its inputs. For example, if all inputs were to double, the outputs would double. Although this is rarely the case, CRS is often used as a simplifying assumption, similar to the assumption that unit costs stay constant if a program is scaled up. However, note that this is quite an ambitious assumption. The CRS model calculates an efficiency score, called CRS technical efficiency (Huguenin, 2012). The second DEA approach assumes “variable returns to scale” technology (VRS model). In a VRS model, an increase in output(s) is less than proportional to increases in input(s). The VRS model calculates an efficiency score called variable returns to scale technical efficiency.

Case Study: Family Planning Vouchers in Rural Pakistan

Introduction

This study reported on the effectiveness and efficiency of the Suraj Social Franchise (SSF) voucher program from the donor’s perspective. Under the program, private healthcare providers in remote rural areas were identified, trained, upgraded, and certified to deliver family planning services to underserved married women of reproductive age (MWRA) in 29 districts of Sindh and three districts of Punjab province, Pakistan between October 2013 and June 2016.

Methods

A decision tree compared the cost of implementing the SSF and the results of its provision of additional couple years of protection (CYPs) to targeted women compared with business-as-usual. Costs included vouchers given to women to receive a free contraceptive method of their choice from the SSF provider. The vouchers were then reimbursed to the SSF provider by the program.

Results

A total of 168,206 MWRA received SSF vouchers between October 2013 and June 2016, which cost \$3,278,000 (\$19.50/recipient). The average effectiveness of the program per voucher recipient was an additional 1.66 CYPs, giving an incremental cost-effectiveness of \$4.28 per CYP compared with not having the program (95% confidence interval [CI]: \$3.62–\$5.31).

Conclusion

The results compared favorably with other interventions having similar objectives and appeared affordable for the Pakistan national healthcare system. The approach was recommended to help address unmet need for contraception among MWRA in these areas of Pakistan and was worthy of trial implementation in the country more widely.

Source: Broughton, Hameed, Gul, Sarfraz, Baig, & Villanueva, 2017

Further Reading

Baltussen, R. M., Adam, T., Tan-Torres Edejer, T., Hutubessy, R. C., Acharya, A., Evans, D. B., ... & World Health Organization (WHO). (2003). *Making choices in health: WHO guide to cost-effectiveness analysis*. Geneva, Switzerland: WHO. Available from http://www.who.int/choice/publications/p_2003_generalised_cea.pdf

Drummond, M., Sculper, M. J., Claxton, K., Stoddart, G., & Torrance, G. (2015). *Methods for economic evaluation of health care Programmes*, Fourth edition. Oxford, UK: Oxford University Press.

Hutubessy, R., Chisholm, D., Tan-Torres Edejer, T. & WHO-CHOICE. (2003). Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost Effectiveness and Resource Allocation*, 1, 8. Available from <https://resource-allocation.biomedcentral.com/articles/10.1186/1478-7547-1-8>

Jamison, D. T. (2009). Cost effectiveness analysis: concepts and applications. In Detels, R., McEwen, J., Beaglehole, R., & Tanaka, H. (Eds.), *Oxford textbook of public health: Volume 2, The methods of public health*, Fifth edition (pp. 767–782). Oxford, UK: Oxford University Press. Available from <http://depts.washington.edu/cfar/sites/default/files/uploads/core-program/user164/Jamison%20CEA%20Concepts%20and%20Applications.pdf>

Muennig, P. & Bounthavong, M. (2016). *Cost-effectiveness analysis in health: A practical guide*, Third edition. San Francisco, CA, USA: Jossey-Bass.

Shillcutt, S.D., Walker, D.G., Goodman, C.A., & Mills, A.J. (2009). Cost-effectiveness in low- and middle-income countries: A review of the debates surrounding decision rules. *PharmacoEconomics*, 27(11), 903–917. Available from <https://link.springer.com/article/10.2165%2F10899580-000000000-00000>

6. COST-BENEFIT ANALYSIS

Introduction

CBA is another tool for making decisions about investing in public health interventions. It helps policymakers make informed decisions by comparing the overall costs of an intervention with societal benefits over time. By comparing costs and benefits in monetary terms, the costs of different program options can be directly compared. CBA can be used to inform the following types of decisions:

- Should a public health intervention be implemented?
- Which public health intervention should be selected from a list of choices?
- How should public health interventions be prioritized in the context of limited budget?

CBA is often used to evaluate investments in large-scale programs, such as national strategies for vaccination or HIV treatment, recommendations to increase screening for disease, or investments in health insurance or infrastructure. Although costing studies are carried out from a range of perspectives, they often involve the perspective of individual patients or healthcare providers. CBA often takes the broader perspective of a government or society because it can include a broader range of impact of an intervention. A government and society may be most interested in understanding the broader economic impact of a policy or intervention compared with the costs.

In conducting CBA, the preferred outcome is the one that maximizes **net benefits**, or the amount by which benefits exceed costs. The benefits can include health outcomes, such as improved quality of life for patients, and non-health outcomes, such as improved worker productivity. Any costs and benefits that can be monetized can be captured in the analysis. When conducting CBA, stakeholders should be consulted early and often to ensure that all costs and benefits are included in the analysis.

The costs and benefits of an intervention can occur at different points in time. An intervention may require investment in the early years of a program, such as purchasing equipment and training staff, whereas the benefits of the intervention to patients could extend during their entire lifetimes. Comparing the costs and benefits using a common unit of measurement allows for the capture of the intervention's long-term benefits. At the end of the analysis, if the projected benefits of an intervention exceed the estimated cost, we can expect that the intervention will improve overall well-being.

Because of the broader range of impact included in a CBA compared with a costing study, CBA can be time-consuming and expensive to undertake. This is another reason why CBA is generally used to inform high-level decision making, although it can be used at any level if the scope of the study is clearly defined and the data are

Learning Objectives

What questions can cost-benefit analysis answer?

What benefits and what costs are in a cost-benefit analysis in public health?

What are the main steps for conducting a cost-benefit analysis?

What are the main criticisms of cost-benefit analysis?

available to support the study. The study's perspective will determine what types of costs and benefits should be involved in the CBA, but all costs and benefits relevant to that perspective should be included. For example, from a patient perspective, costs should include not just out-of-pocket expenses, but also travel costs, missed work time to attend appointments, and any other opportunity costs of receiving treatment. From a societal perspective, per patient costs may include these same types of costs for all patients, along with the costs incurred by caretakers who accompany patients to appointments. These patient and caretaker costs should be included in societal costs along with other types of costs of providing the services.

Summary Measures

Cost-benefit comparisons can be expressed as a **cost-benefit ratio** (benefits/costs) or as **net benefits** (benefits–costs):

- Cost-benefit ratios demonstrate the savings per dollar of program expenditure.
- Net benefits express the total savings due to the intervention. The intervention should be implemented if net benefits are greater than zero.

In this chapter, we discuss the principles of CBA and how they are applied in public health. First, we discuss how the costs estimated in Chapter 2 should be applied in CBA. Next, we describe the types of benefits that should be included in an analysis and how they are measured. We then introduce the concept of discounting to allow for a comparison of benefits and costs that occur at different points in time in a consistent way. After the cost and benefit concepts are defined, sample applications are provided, followed by a discussion of cautions to keep in mind when using the cost-benefit approach. Finally, we summarize the steps for conducting a CBA.

Measuring Costs and Benefits

Costs

The cost measure used in a CBA should summarize the total costs to society of implementing an intervention. They include **direct costs** and **indirect costs**. They also include costs that are **tangible**, that is, they can be counted and measured, and those that are **intangible**. Costs can represent one-time expenditures on physical equipment or recurring expenditures that should be considered into the future. Refer to Chapter 2 for more information about estimating the total cost of an intervention.

Cost information on its own can be useful for evaluating an intervention. Treatment costs can be compared across diseases. Estimates of costs can be used to advocate for resources to reduce the burden of disease. However, by combining costs and benefits, the interventions that maximize the benefits for a level of expenditure can be prioritized.

Benefits

As with costs, all benefits to society should be considered and measured (if possible) for inclusion in a CBA. In this section, we discuss the types of benefits to consider and how to monetize some of these types of benefits.

Direct benefits include avoided medical expenditures resulting from an intervention due to treatment of a condition or by preventing disease and illness. **Indirect benefits** are less tied to health outcomes and generally include increases in worker productivity due to better health of workers. Improved productivity increases

earnings for workers and employers and should be measured as part of a CBA. **Intangible benefits** include the other quality of life benefits that come from better health. There are several ways to quantify intangible benefits.

The **human capital approach** to measuring the benefits of an intervention considers lost earnings due to mortality and morbidity. In this approach, the earnings that are lost when a worker loses productive time because of premature mortality or morbidity are computed as a measure of the cost of illness to society. The value of morbidity is estimated as a loss of earnings from lost years of working life. The value of mortality is estimated as the earnings lost during time lost from work due to illness, and lost productivity because of disability. However, measurements using the human capital approach are not complete because they exclude the following:

- Workers who experience illness also lose the ability to enjoy leisure time and experience pain and suffering, which are not valued.
- This approach does not include the lost time of people outside the labor force, including retirees, children, and disabled people. It also assigns lower values to the time of people who work in lower wage jobs.
- The time of people who are involuntarily unemployed is likewise not valued, although they have the potential to produce value and also suffer similar effects of illness.

Two other tools for valuing benefits attempt to understand peoples' preferences: revealed and stated preference methods. **Revealed preference methods** use market prices to infer the value of non-market goods, such as improved health. **Stated preference methods** use survey questions to capture preferences for improved health. The measures generated from these surveys are **willingness to pay** (how much a person would give up to enjoy a better health outcome) and **willingness to accept** (how much compensation a person requires to accept a worse health outcome). These measures are more commonly used in CBA for environmental regulation. For public health, other non-monetary measures have been developed to consider the impact on quality of life.

Non-health benefits can be considerable for some health interventions. In his review, Weatherley (2009) lists benefits in criminal justice, education, law enforcement, housing, employment, the environment, and transportation, among others. Many of these benefits stem from improvements in health (or reductions in unhealthy conditions), which have downstream effects on the affected population. This is especially true for preventive health services, where the benefits are not only reductions in morbidity or averted mortality but can also be averted health treatment costs. For example, CBAs of family planning programs typically have very high cost-benefit ratios, mainly because the benefits are measured in terms of reduced expenditures on education, housing, water and sanitation, and immunization costs. Moreover, family planning investments can produce benefit streams over a long period of time. In a cost-benefit study of family planning in 16 SSA countries that included multi-sector benefits, Moreland and Talbird (2006) found benefit-cost ratios ranging from 1 to 13.

Discounting

Costs and benefits of an intervention can occur at different points in time. To conduct CBA, a consistent measure of value over time should be used. Therefore, **discounting** is used to compare the **present value** of the stream of costs and benefits of an intervention using consistent monetary units over time.

Investing in an intervention now has an **opportunity cost**. Every dollar that is spent on the intervention could be invested to produce a rate of return, allowing for increased spending in the future. Discounting accounts for the lost value of future consumption. Moreover, in general, people prefer consumption in the present over the future. Discounting reduces the value of consumption in the future to account for this preference. Last, because most economies grow over time, expenditures now are more valuable in the present compared with a future with more resources. Discounting accounts for the reduced value of future consumption.

Discounting benefits is also appropriate. Just as people prefer consumption in the present over the future, so do they prefer health benefits now compared with health benefits in the future. Discounting only the costs is not appropriate because delaying the intervention makes it appear cheaper compared with the same level of health benefits. The accepted method for CBA includes discounted benefits and costs.

The formula for the **discount factor** is: $1 / (1+r)^t$

Where r is the **discount rate** and t is number of years in the future when the cost or benefit is expected to occur. As mentioned above, the costs of an intervention could be greater in the earlier years of a program, whereas the benefits could begin to accrue later or even beyond the time when the program ends. The benefits should still be considered for each year in the future that they are expected to occur. Real costs, that is, those costs without the effects of inflation, should be considered as all costs and benefits are valued in present year dollars. Discount rates generally range from three percent to seven percent, considering projected economic growth and citizens' preferences (Office of Management and Budget [OMB], 1992).

Uncertainty

When conducting CBA, care should be taken to clearly define the scope of the study. Because costs and benefits can continue into the future, a time horizon for the analysis that is long enough to capture as many costs and benefits as possible but that is still feasible to estimate should be defined. In some cases, only the private sector costs and savings of a particular intervention can be considered. In other cases, government expenditures and savings that would result from the intervention may need to be included. Defining the scope of the analysis helps identify which items should and should not be counted in the analysis.

Cost-benefit exercises usually project the costs and benefits of interventions over time, introducing uncertainty about the future. As we project further out in time, assumptions about the population, economic values, and technology become more uncertain. There are two types of uncertainty that may need to be considered (USEPA, 2010):

- **Statistical uncertainty:** Many benefits are based on statistical estimates of the effects of an intervention on a given population. Our uncertainty about these estimates can be measured by using statistical techniques. If the probability of different outcomes is known, we can assign them values and determine the most likely magnitude of the benefits.

- **Incomplete knowledge:** Even using our best estimates about the costs and benefits of a public health intervention leaves uncertainty. The sources of this uncertainty include, but are not limited to, incomplete understanding of health processes and imperfect forecasts of future circumstances. Although this uncertainty cannot be eliminated, we can attempt to disclose all the sources of uncertainty that can be identified. By being transparent about our assumptions about the future, and providing sensitivity analysis when possible, we can better conceptualize the possible future outcomes of an intervention.

When presenting uncertain information, estimates that represent the best knowledge available should be used, not “worst-case” scenarios. One should maintain consistency of assumptions across the estimation of costs and benefits.

Conducting Cost-Benefit Analysis

The following is a summary of the steps involved in conducting a CBA (Lave & Joshi, 1996; CDC, 2016; OMB, 2003).

1. Clearly define the scope of the intervention and its goals. Correctly identifying the costs and benefits of an intervention depends on having a clear picture of the desired outcome of a project. Define a time horizon for the analysis that is long enough to capture all the costs and benefits that could result from the intervention.
2. Consult with stakeholders and experts about the range of possible costs and benefits of the project. This step helps identify potential risks and rewards of the project that may not be expected, especially because public health interventions may require the work of multiple agencies. The more comprehensive the analysis, the stronger it will be.
3. List all costs of the intervention and identify when they will occur in time. Costs can include:
 - a. Direct costs, such as the costs of medical supplies and personnel. Where market prices do not reflect the true cost of an item, use the appropriate cost.
 - b. Indirect costs, such as support staff that may be needed to administer the program
 - c. Intangible costs, such as the frustration of patients who must wait longer for services under the new program. Valuing these costs may not be possible, but they should still be identified and considered in the analysis.
 - d. Potential risks of undertaking the project should also be considered. Like intangible costs, risks may not be quantified, but should be discussed and considered as part of the analysis.
4. List all benefits of the intervention and identify when they will occur in time. Benefits can include:
 - a. Direct benefits, such as savings of medical costs in the future by preventing disease in the present.
 - b. Indirect benefits, such as earnings that result from improved worker productivity.
 - c. Intangible benefits, such as improved quality of life of people who receive the intervention and their families. Even if these benefits cannot be fully quantified or monetized, they

should still be discussed and considered as part of the analysis. Items that cannot be monetized can sometimes be counted, but be careful not to double count any costs or benefits, even those that are difficult to value.

5. Monetize costs and benefits in common units:
 - a. Quantify as many costs and benefits as possible. For those that can be expressed in monetary terms, convert to currency units.
 - b. Convert into common monetary measures by discounting future costs and benefits so that they can be directly compared. Multiply the appropriate discount factor for each year by the total costs and benefits in each year.
6. Compare costs and benefits by calculating the net benefits: $NB = B - C$. Do not use a cost-benefit ratio because it can be manipulated by moving items between cost and benefit categories.
7. Report the findings of your CBA.
8. Document all your assumptions about the time horizon you have considered, the discount rate used, and the data and models that were used to calculate the costs and benefits.
9. Discuss the additional costs and benefits that you have identified that cannot be quantified or monetized.
10. Discuss sources of uncertainty in your analysis and perform sensitivity analysis if large uncertainties are possible.
11. Use the analysis to inform decisions about:
 - a. Whether to implement a specific intervention
 - b. Which project from a set of alternatives should be selected
 - c. What are the best uses of limited funds

The CBA provides important information about the overall benefit to society of an intervention. However, officials should also consider the distributional effects of who pays for the costs of the intervention and who enjoys the benefits. Officials should also consider the costs and benefits that cannot be monetized and that were identified in the analysis, and the key sources of uncertainty identified in the CBA process.

Criticisms of Cost-Benefit Analysis

A common criticism of CBA is that costs are easier to quantify than health benefits, making it harder to demonstrate the value of the intervention. As discussed above, the commonly used human capital approach underestimates the value of mortality and morbidity of people who do not work, such as the young, elderly, disabled, or unemployed. For those in the labor force, quality of life outside of their ability to work is likewise not captured.

Discounting of benefits suggests that the lives of future generations are less valuable than the lives of the current generation. However, failing to discount future health benefits may lead to decisions that delay

investing in a public health intervention. Just as people prefer consumption now over consumption in the future, they also prefer to enjoy health benefits in the present.

When projecting the future costs and benefits of an intervention, the health risks associated with inaction and the health benefits of the program may be very uncertain. Sensitivity analysis can be used to examine the relative costs and benefits of interventions using different discount rates or using different assumptions about the costs and benefits over time. Creating a range of possible outcomes under different scenarios can help guide the policy discussion about an intervention, provided that the scenarios are plausible.

Because of the need to consider all costs and benefits to society, CBAs can take months or sometimes years to complete. This may be appropriate for expensive, large-scale interventions, but could risk delaying investment in life-saving interventions.

Related Concepts

Return on investment (ROI) is another way to express costs and benefits. ROI expresses benefits as a percentage of the initial investment.

$$ROI = 100 * Net\ benefits / Total\ costs$$

In most cases, ROI of a successful project should be greater than zero. However, if the project results in mainly intangible benefits that cannot be quantified in monetary terms, ROI may not be an appropriate measure. ROI can be calculated for a smaller set of benefits and costs than a full CBA. Care should still be taken to clearly identify the set of costs and benefits included in the analysis. As in CBA, costs and benefits over time should be discounted when calculating ROI (Applied Geographics, 2009).

ROI is increasingly used as an advocacy tool for public health interventions. According to Brousselle, et al. (2016), “ROI is a timely tool with which to advocate for public health interventions that have long term implications and require substantial investments. Assessing the economic value of public health programs and interventions using ROIs, cost-offsets, and profitability threshold analysis could provide robust arguments in their defense” (Brousselle, Benmarhania, & Benhadj, 2016). However, ROI as a decision rule to choose between alternative interventions can change funding priorities, especially when there are significant differences between the interventions. As reported by Brousselle, et al. (2016), water fluoridation has an ROI of 3700 percent whereas vaccination has an ROI of 1500 percent. Does this mean that fluoridation programs should be funded in priority over vaccination programs? As discussed in Chapter 1, economic evaluation approaches like ROI, CBA, or CEA can provide metrics to guide decisions but may not be the only factors in making these decisions.

Some organizations use the concept of **value for money (VFM)** to evaluate investment decisions alongside, or instead of, CBA. The United Kingdom’s DFID defines VFM as “the optimal use of resources to achieve intended outcomes” to “maximize the impact of each pound spent to improve poor people’s lives” (DFID, 2011). The VFM approach looks not just at the total costs and benefits of an intervention, but includes the “three E’s” of economy, efficiency, and effectiveness.

Economy refers to the prices of inputs that meet quality standards, such as drugs, equipment, or staff.

Efficiency refers to the ability of the project to turn inputs into measurable outputs. Effectiveness refers to the

ability of a project to achieve expected outcomes. The VFM approach can be used at smaller project scales than CBA and emphasizes ongoing evaluation of cost efficiency during the project lifetime (Jackson, 2012).

Case Study: Childhood Immunization

A 2016 study estimated the costs and benefits, and ROI, of vaccination against ten diseases in LMICs during the period 2011 to 2020. The costs of providing vaccination included the costs of the vaccines, supply chain costs (including transportation, storage, and labor), and costs of providing vaccination services. The authors used vaccine demand forecasts to determine the number of doses given during the period of analysis, along with price forecasts, to calculate the total cost of vaccination.

To calculate the benefits of vaccination, the authors used health impact models to estimate the reduction in cases of disease and associated reductions in mortality and morbidity. They estimated avoided cost-of-illness based on treatment costs (including transportation), lost productivity due to death or disability, and lost productivity of caretakers. For lost productivity, the authors represented annual income by GDP per capita, and translated the lost work time into lost earnings.

In addition to lost earnings, the authors estimated the benefits of additional healthy years of life. They valued each year of life saved at 1.6 times GDP per capita to represent the value people place on good health (based on studies of the value of a statistical life). They valued years of disability avoided at this rate times a DW for each condition. The authors noted that many potential benefits were missing from the analysis, including the benefits of herd immunity for those who cannot be vaccinated, additional macroeconomic benefits of a healthier population, and demographic changes due to fewer deaths from these illnesses.

The authors summarized the costs and benefits through the measure of ROI, or net benefits per dollar invested. Using only the cost-of-illness approach, the benefits were 16 times higher than the costs. Including the broader economic benefits of healthy life-years gave the benefits as 44 times higher than the costs. The ROI varied across the different vaccines and regions considered, but all vaccines had a positive ROI. The authors also conducted sensitivity analysis to create an uncertainty range around the ROI estimates.

Source: Ozawa, Clark, Portnoy, Grewal, Brenzel, & Walker, 2016.

Case Study: Treating Noncitizens with HIV in Botswana

In 2016, Botswana introduced a “Treat All” strategy with the aim of HIV epidemic control. Under this policy, citizens of Botswana living with HIV receive ART for free. However, noncitizen residents, representing seven percent of the population, were not eligible for treatment under this program. Surveys conducted in the country estimated that only 27 percent of noncitizens living with HIV in Botswana were accessing ART.

Compared with the other countries in the region, Botswana had similar treatment guidelines and levels of ART coverage, but was the only one of its neighbors that did not provide treatment to noncitizens. Surveys conducted in Botswana showed that most noncitizens were between the ages of 15 and 49 years, which is a demographic that is both sexually active and economically productive. In addition, one-third of people in Botswana ages 15 to 34 engaged in multiple concurrent sexual partnerships. Because ART reduces the risk of transmission of HIV, having a low percentage of sexually active noncitizens on treatment increased the risk of infection among citizens and noncitizens alike.

In 2018, an interagency working group prepared a CBA of adding noncitizens to the ART program. They estimated the costs of treating noncitizens from 2018 to 2030 to be \$18 million, using the Spectrum modeling system (Stover, 2003). This model also projected the HIV and TB infections averted. The net cost of the program included the ART costs for the added noncitizens, minus the cost savings for treatment of the HIV and TB infections avoided. The model estimated cost savings of \$116 million between 2018 and 2030.

In addition to the treatment cost savings, the team estimated the monetary gains in productivity that resulted from treated noncitizens returning to the workforce at \$30 million by 2030. Additional non-monetary benefits included the number of lives saved (about 6,700) and infections averted about (23,000), which benefitted both the citizen and noncitizen populations. The program was projected to improve indicators of epidemic control and accelerate progress toward the government's treatment targets.

Source: MEASURE Evaluation. (2018). Saving lives, transforming the economy: Making treat all real and saving 23,000 more from HIV by 2030. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina.

Further Reading

Harvard University's T.H. Chan School of Public Health has received funding from the Bill and Melinda Gates Foundation to develop Guidelines for Benefit-Cost Analysis. Follow its progress here:

<https://sites.sph.harvard.edu/bcaguidelines/>.

The US Office of Management and Budget has guidelines for regulatory analysis, including CBA found here:

<https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A4/a-4.pdf>.

CBA has been used in environmental regulation more frequently than in public health. Read about the EPA

Guidelines for Preparing Economic Analyses here: <https://www.epa.gov/environmental-economics/guidelines-preparing-economic-analyses>.

The UK Independent Commission for Aid Impact has produced recommendations for using the VFM in the international aid context: <https://icai.independent.gov.uk/report/value-for-money/>

READERS' DISCUSSION GUIDE

Chapter 1: Introduction

Economic evaluation helps decision makers look at options for health projects, programs, and interventions from an economic perspective. What are some of the limitations and challenges of taking an economic perspective?

Economic evaluation involves analyses that can assist with public health program and policy decisions based on economic principles. What are some of the other factors that decision makers might take into account? How might they weight these criteria relative to economic factors?

Chapter 2: Cost Analysis

Below are a few examples of questions addressed by specific costing studies. As you read them, think about the decision problems you would want to answer in designing each study.

- What is the cost per service for HIV testing at a government PHC in Kano, Nigeria? For those found HIV positive, what is the cost of immediate ART consultation and follow-up adherence care?
- Is there variation in the cost per service at different service points delivering HIV tests at PHCs in Kano? What drives variation (staff time, demand, etc)?
- What is the difference in cost per beneficiary for three OVC implementing partners in Kenya that provide nutritional support?

Here are some additional examples of questions about program changes that cost data help inform. Again, think about the decision problems that need to be answered in designing each study.

- What are the savings due to a shift in delivery of family planning counseling from nurses to community health workers?
- What is the additional cost of adding social quality early childhood development services to existing pediatric treatment and PMTCT programming?
- How much would it cost for the government of South Africa to scale up a VMMC program from one district to 25 districts?
- How much would it cost to transition HIV testing of OVC from the health facility to the community?

Choose an example of a costing study from the list above. What is the *perspective* that would find this study most useful? What types of costs would you want to collect to inform this study? What are some potential sources of the cost data?

Chapter 3: Measuring Impact on Lives

DALYs can be used to compare the burden of disease across populations—including by geography, age group, or income level—or to compare the burden of different diseases. What are some applications where this information could help make the case for public health interventions? What types of information would researchers need to collect to conduct a study that presents DALYs as an output?

What are some reasons why researchers may want to convert DALYs or QALYs to monetary units?

Chapter 4: Evaluating Impact

A major challenge in conducting an impact evaluation of health programs is trying to attribute any changes in the outcome measures to the intervention or program. What are some ways that researchers can deal with this challenge?

Many impact evaluations often assume that projects are implemented as planned and/or in the same way in different locations. How can evaluation researchers account for the realities of the diversity of projects and their deviation from original designs?

Chapter 5: Cost-Effectiveness Analysis

The perspective of analysis is an important determinant of the technical approach one can use in computing cost-effectiveness or efficiency. Which perspective is the most appropriate and which analytical approach would you use in the following examples? What would be the best measures of impact?

1. Calculating the cost-effectiveness of a mobile clinic to provide family planning services to a community.
2. Calculating the cost-effectiveness of a new self-administered injectable contraceptive.
3. Calculating the cost-effectiveness of a project to increase the use of condoms by female sex workers.

Chapter 6: Cost-Benefit Analysis

What are some benefits of public health interventions that can be expressed in monetary units? What are some benefits that are harder to quantify? How can these additional benefits be described to complement a CBA?

Think about the perspective of a costing study, as described in Chapter 2. What perspectives would find CBA most useful?

REFERENCES

- Adato, M. (2011). *Combining quantitative and qualitative methods for program monitoring and evaluation: Why are mixed-method designs best?* Washington, DC, USA: World Bank. Retrieved from <https://openknowledge.worldbank.org/handle/10986/11063>.
- Akazili, J., Adjuik, M., Jehu-Appiah, C., & Zere, E. (2008). Using data envelopment analysis to measure the extent of technical efficiency of public health centres in Ghana. *BMC International Health and Human Rights*, 8(1), 11. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/19021906>.
- Alkire, B., Hughes, C.D., Nash, K., Vincent, J.R., & Meara, J.G. (2011). Potential economic benefit of cleft lip and palate repair in sub-Saharan Africa. *World Journal of Surgery*, 35(6), 1194–1201. Retrieved from <https://doi.org/10.1007/s00268-011-1055-1>.
- Allhoff, F. (2005). The Oregon plan and QALYs. *Virtual Mentor*, 7(2). Retrieved from <https://journalofethics.ama-assn.org/article/oregon-plan-and-qalys/2005-02>.
- Applied Geographics. (2009). Economic justification: Measuring return on investment (ROI) and cost benefit analysis (CBA). USGS Contract #: 08HQC�0024. Retrieved from <https://www.fgdc.gov/initiatives/50states/newsbp/EconomicJustification ROI-CBA-Tutorial v2 052809 FinalVersa.pdf>.
- Asaria, M., Griffin, S., & Cookson, R. (2016). Distributional cost-effectiveness analysis: a tutorial. *Medical Decision Making*, 36(1), 8–19. Retrieved from <https://journals.sagepub.com/doi/pdf/10.1177/0272989X15583266>.
- Baltussen, R. M., Adam, T., Tan-Torres Edejer, T., Hutubessy, R. C., Acharya, A., Evans, D. B., ... & World Health Organization (WHO). (2003). *Making choices in health: WHO guide to cost-effectiveness analysis*. Geneva, Switzerland: WHO. Retrieved from <https://apps.who.int/iris/bitstream/handle/10665/42699/9241546018.pdf?sequence=1&isAllowed=y>.
- Bamberger, M. InterAction. (2012). *Introduction to mixed methods in impact evaluation* (No. 3). Retrieved from: <http://www.interaction.org/sites/default/files/Mixed%20Methods%20in%20Impact%20Evaluation%20%28English%29.pdf>.
- Bamberger, Michael, (2012) *Introduction to Mixed Methods in Impact Evaluation*, Impact Evaluation Notes #3, August, 2012. <https://www.interaction.org/resources/training/guidance-note-3-introduction-mixed-methods-impact-evaluation>
- Bhat, R., Verma, B.B., & Reuben, E. (2001). Data envelope analysis. *Journal of Health Management*, 3(2). Retrieved from <https://journals.sagepub.com/doi/10.1177/097206340100300207>.
- Bamberger, M., Rugh, J. and Mabry, L., (2012) *Real world evaluation: Working under budget, time, data and political constraints*. 2nd edition. Los Angeles, CA, USA: Sage Publications.
- Boardman, A., Greenberg, D., Vining, A., & Weimer, D. (2011). *Cost-benefit analysis: Concepts and practice*, Fourth edition. Upper Saddle River, NJ, USA: Pearson/Prentice-Hall.

- Bonbright, D. (2012). *Use of impact evaluation results*, Impact Evaluation Notes #4. Retrieved from http://www.pointk.org/resources/files/Use_of_Impact_Evaluation_Results.pdf.
- Broughton, E.I., Hameed, W., Gul, X., Sarfraz, S., Baig, I.Y., & Villanueva, M. (2017). Cost-effectiveness of a family planning voucher program in rural Pakistan. *Frontiers in Public Health*, 5, 227. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5614961/>.
- Brousselle, A., Benmarhania, T., & Benhadj, L. (2016) What are the benefits and risks of using return on investment to defend public health programs? *Preventive Medicine Reports*, 3:135-8. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/27419005>.
- Bullen, P.B. (2014). Theory of change vs logical framework: What's the difference? Retrieved from <http://www.tools4dev.org/resources/theory-of-change-vs-logical-framework-whats-the-difference-in-practice/>.
- Byford, S., & Sefton, T. (2003). Economic evaluation of complex health and social care interventions. *National Institute Economic Review*, 186, 98–108. Retrieved from https://www.jstor.org/stable/23876789?seq=1#page_scan_tab_contents.
- Centers for Disease Control and Prevention, Division for Heart Disease and Stroke Prevention. (2016). Five Part Webcast on Economic Evaluation. Retrieved from https://www.cdc.gov/dhdsp/evaluation_resources/economic_evaluation/index.htm
- Centers for Disease Control and Prevention. (2017). Facts about cleft lip and palette. Retrieved from <https://www.cdc.gov/ncbddd/birthdefects/cleftlip.html>.
- Centers for Disease Control and Prevention. (2018). Health-related quality of life (HRQOL): Methods and measures. Retrieved from <https://www.cdc.gov/hrqol/methods.htm>.
- Chan, Y.L. (1993). Improving hospital cost accounting with activity-based costing. *Health Care Management Review*, 18(1):71–77. Retrieved from <http://sit.tums.ac.ir/UserFiles/mali/Ghavanin2/1-activity%20based%20cost%20accounting.pdf>.
- Charnes, A., Cooper, W.W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelope analysis for Pareto-Koopmans efficient production functions. *Journal of Econometrics*, 30(1–2), 91–107. Retrieved from https://econpapers.repec.org/article/eeeeconom/v_3a30_3ay_3a1985_3ai_3a1-2_3ap_3a91-107.htm.
- Cohen, J., & Saran, I. (2018). The impact of packaging and messaging on adherence to malaria treatment: Evidence from a randomized controlled trial in Uganda. *Journal of Development Economics*, 134, 68–95. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0304387818305066>.
- Cokins, G. (1996). *Activity-based cost management: Making it work*. New York, NY, USA: McGraw-Hill.
- Conteh, L., & Walker, D. (2004). Cost and unit cost calculations using step-down accounting. *Health Policy and Planning*, 19(2), 127–135. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/14982891>.
- Cunnam, L., Sinanovic, E., Ramma, L., Foster, N., Berrie, L., Stevens, W., ... Vassall, A. (2016). Using top-down and bottom-up costing approaches in LMICs: The case for using both to assess the incremental costs of

new technologies at scale. *Health Economics*, 25(Suppl. 1):53–66. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/26763594>.

Davies, R. (1998). *Order and diversity: Representing and assisting organisational learning in non-government aid organisations*. PhD Thesis, University of Wales. Retrieved from https://www.academia.edu/1058592/Order_and_Diversity_Representing_and_Assisting_Organisational_Learning_in_Non-Government_Aid_Organisations.

Department for International Development. (2011). DFID's approach to value for money (VfM). Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49551/DFID-approach-value-money.pdf.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*, Fourth edition. Oxford, UK: Oxford University Press.

Dutta, A., N. Perales, R. Silva, & L. Crivillé. (2014). *Estimated resource needs and impact of Mozambique's Plano Estratégico do Sector Saúde, 2014–2019*. Washington, DC, USA: Futures Group, Health Policy Project. Retrieved from https://www.healthpolicyproject.com/pubs/242_MozambiqueRelatórioCustosparaPESSEenglish.pdf.

Elliott, R., & Payne, K. (2005). *Essentials of economic evaluation in healthcare*. London, UK: Pharmaceutical Press.

Gobin, S., & Foley, S. (2019a). Rapid costing assessment of USAID-funded structural & behavioral HIV prevention activities, Part II: Results of quantitative cost analysis. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from <https://www.measureevaluation.org/resources/publications/tr-19-323>.

Gobin, S., & Foley, S. (2019b). The cost of case management in orphans and vulnerable children programs: Results from a mixed-methods, six-country study. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from: <https://www.measureevaluation.org/resources/publications/tr-19-327>.

Gold, M. R., Stevenson, D., & Fryback, D. G. (2002). HALYS and QALYS and DALYS, Oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health*, 23(1), 115–134. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/11910057>.

Grad, F. P. (2002). The preamble of the constitution of the World Health Organization. *Bulletin of the World Health Organization*, 80 (12), 981–981. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2567708/>.

Hammitt, J. K., & Robinson, L. A. (2011). The income elasticity of the value per statistical life: Transferring estimates between high and low-income populations. *Journal of Benefit-Cost Analysis*, 2(1), 1–29. Retrieved from <https://www.cambridge.org/core/journals/journal-of-benefit-cost-analysis/article/income-elasticity-of-the-value-per-statistical-life-transferring-estimates-between-high-and-low-income-populations/3AE5B0BB1034B7D898E2C9B3C1507794>.

Hendriks, M.E., Kunda, P., Boers, A.C., Bolarinwa, O.A., Te Pas, M.J., Akande, T.M., ... Swan Tan S. (2014). Step-by-step guideline for disease-specific costing studies in low- and middle-income countries: A mixed methodology. *Global Health Action*, 7, 23573. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/24685170>.

Huguenin, J-M. (2012). *Data envelopment analysis (DEA): A pedagogical guide for decision makers in the public sector*. Lausanne, Switzerland: Institut de hautes études en administration publique. Retrieved from https://serval.unil.ch/resource/serval:BIB_0FC432348A97.P001/REF.

Hutubessy, R., Chisholm, D., Tan-Torres Edejer, T. & WHO-CHOICE. (2003). Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost Effectiveness and Resource Allocation*, 1, 8. Retrieved from <https://resource-allocation.biomedcentral.com/articles/10.1186/1478-7547-1-8>.

Institute of Medicine. (2009). Chapter 2, What is comparative effectiveness research? In *Initial National Priorities for Comparative Effectiveness Research*. Washington, DC, USA: The National Academies Press.

Jackson, P. (2012). Value for money and international development: Deconstructing myths to promote a more constructive discussion. Retrieved from <https://www.oecd.org/development/effectiveness/49652541.pdf>.

Jamison, D. T. (2009). Cost effectiveness analysis: Concepts and applications. In Detels, R., McEwen, J., Beaglehole, R., & Tanaka, H. (Eds.), *Oxford textbook of public health: Volume 2, The methods of public health*, Fifth edition (pp. 767–782). Oxford, UK: Oxford University Press. Retrieved from http://depts.washington.edu/cfar/sites/default/files/uploads/core-program/user164/Jamison_CEA_Concepts_and_Applications.pdf

Janowitz, B., & Bratt, J. (1994). *Methods for costing family planning services*. New York, NY: United Nations Population Fund. Retrieved from <https://www.k4health.org/sites/default/files/unpf0050.pdf>.

Javid, M., Hadian, M., Ghaderi, H., Ghaffari, S., & Salehi, M. (2016). Application of the activity-based costing method for unit-cost calculation in a hospital. *Global Journal of Health Science*, 8(1): 165–172. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4803948/>.

Joint United Nations Programme on HIV/AIDS (UNAIDS). (2000). *Costing guidelines for HIV prevention strategies*. Geneva, Switzerland: UNAIDS. Retrieved from http://data.unaids.org/publications/irc-pub05/jc412-costguidel_en.pdf.

Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: What is the difference?. *Pharmacoeconomics*, 34(7), 645–649. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/26892973>.

Kotvojs, F., & Shrimpton, B. (2007). Contribution analysis: A new approach to evaluation in international development. *Evaluation Journal of Australasia*, 7(1), 27–35. Retrieved from <https://www.aes.asn.au/images/stories/files/conferences/2006/papers/022%20Fiona%20Kotvojs.pdf>.

Lance, P., Guilkey, D., Hattori, A. & Angeles, G. (2014). *How do we know if a program made a difference? A guide to statistical methods for program impact evaluation* (ms-14-87-en). Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from <https://www.measureevaluation.org/resources/publications/ms-14-87-en>.

Larson, B. A. & Wambua, N. (2011). How to calculate the annual costs of NGO-implemented programmes to support orphans and vulnerable children: A six-step approach. *Journal of the International AIDS Society*, 14, 59. Retrieved from <https://jiasociety.biomedcentral.com/articles/10.1186/1758-2652-14-59>.

Lave, L. B. & Joshi, S. V. (1996). Benefit-cost analysis in public health. *Annual Review of Public Health*, 17, 203–219. Retrieved from <https://www.annualreviews.org/doi/abs/10.1146/annurev.pu.17.050196.001223>.

LeFevre, A. E., Shillcutt, S. D., Broomhead, S., Labrique, A. B., & Jones, T. (2017). Defining a staged-based process for economic and financial evaluations of mHealth programs. *Cost Effectiveness and Resource Allocation*, 15, 5. Available from <https://resource-allocation.biomedcentral.com/articles/10.1186/s12962-017-0067-6>. Levin, C. (n.d.). Methods for data collection and analysis: Costing methods in global health. Retrieved from <http://dcp-3.org/sites/default/files/resources/Levin%20CFAR%20CE%20research%20methods%20workshop%20Session%203%20.pdf>

Luce, B., Manning, W., Siegel, J., & Lipscomb, J. (1996). Estimating costs in cost effectiveness analysis. In Gold, M.R., Siegel, J.E., Russell, L.B., & Weinstein, M.C. (Eds.) *Cost effectiveness in health and medicine*. New York, NY, USA: Oxford University Press.

Malaria Impact Evaluation Group, Liberia & Herrera, S. (2018). *Evaluation of the impact of malaria control interventions on all-cause mortality in children under five years of age in Liberia, 2005–2013*. Washington, DC, USA: USAID. Retrieved from <https://www.pmi.gov/docs/default-source/default-document-library/pmi-reports/liberia-malaria-impact-evaluation-group-full-report.pdf?sfvrsn=6>.

Marseille, E., Larson, B., Kazi, D. S., Kahn, J. G., & Rosen, S. (2015). Thresholds for the cost–effectiveness of interventions: Alternative approaches. *Bulletin of the World Health Organization*, 93(2), 118–124. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4339959/>.

Mathers, C.D., Vos, T., Lopez, A.D., Salomon, J., Ezzati, M. (Eds.) (2001). *National burden of disease studies: A practical guide*. Edition 2.0. Global Program on Evidence for Health Policy. Geneva, Switzerland: WHO. Retrieved from <https://www.who.int/healthinfo/nationalburdenofdiseasemanual.pdf>.

Mayne, John (1999), Addressing attribution through contribution analysis: Using performance measures sensibly, Discussion paper, Office of the Auditor General Canada. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.456.9890&rep=rep1&type=pdf>.

Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *The Canadian Journal of Program Evaluation*, 16, 1–24. Retrieved from <https://www.betterevaluation.org/en/resources/guide/addressing-attribution-through-contribution-analysis>.

McBain, R.K., Jerome, G., Leandre, F., Browning, M. Warsh, J., Shah, M., ... Kaplan, R. (2018). Activity-based costing of health-care delivery, Haiti. *Bulletin of the World Health Organization*, 96, 10–17. Retrieved from <https://www.who.int/bulletin/volumes/96/1/17-198663/en/>.

MEASURE Evaluation. (2018). Saving lives, transforming the economy: Making “treat all” real and saving 23,000 more Batswana from HIV by 2030. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina.

- Mogyorosy, Z., & Smith, P. (2005). *The main methodological issues in costing health care services: A literature review*. York, UK: The Centre for Health Economics, The University of York. Retrieved from https://www.york.ac.uk/media/che/documents/papers/researchpapers/rp7_Methodological_issues_in_costing_health_care_services.pdf
- Moreland, S., and Talbird, S. (2006). *Achieving the Millennium Development Goals: The contribution of fulfilling unmet need for family planning*. Washington, DC, USA: The Futures Group/ POLICY Project. Retrieved from <http://www.healthpolicyplus.com/archive/ns/pubs/hpi/MDGMaster%209%2012%2006%20FINAL.pdf>.
- Moreland, S. (2011). *Methods for estimating the costs of family planning: Report of the Expert Group Meeting on Family Planning Costing*. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from <https://www.measureevaluation.org/resources/publications/ws-11-29>.
- Moreland, S., Foley, S., & Gobin, S. (2018). *Reproductive health cost reporting system: A user guide*. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from <https://www.measureevaluation.org/resources/publications/ms-18-126>.
- Moreland, S., Namisango, E., Paxton, A., & Powell, RA. (2013). *The costs of HIV treatment, care, and support services in Uganda* (tr-13-94.pdf). Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from <https://www.measureevaluation.org/resources/publications/tr-13-94>.
- Morrow, R. H., & Bryant, J. H. (1995). Health policy approaches to measuring and valuing human life: conceptual and ethical issues. *American Journal of Public Health*, 85(10), 1356–1360. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1615602/>.
- Muennig, P., & Bounthavong, M. (2016). *Cost-effectiveness analysis in health: A practical approach*, Third edition. San Francisco, CA, USA: Jossey-Bass.
- Murray, C. J., Evans, D. B., Acharya, A., & Baltussen, R. M. (2000). Development of WHO guidelines on generalized cost-effectiveness analysis. *Health Economics*, 9(3), 235–251. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/10790702>.
- Office of Management and Budget. (1992). Circular No. A-94: Guidelines and discount rates for benefit-cost analysis of federal programs. Retrieved from <https://www.wbdg.org/FFC/FED/OMB/OMB-Circular-A94.pdf>.
- Office of Management and Budget. (2003). OMB Circular A-4: Regulatory analysis. Retrieved from <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A4/a-4.pdf>.
- Orlewska, E., & Mierzejewski, P. (2003). Guidelines for cost calculation in economic evaluations of healthcare programs (PROJECT). Retrieved from: <http://www.czytelniamedyczna.pl/424,guidelines-for-cost-calculation-in-economic-evaluations-of-healthcare-programs-p.html>.
- Ozawa, S., Clark, S., Portnoy, A., Grewal, S., Brenzel, L. & Walker, D.G. (2016). Return on investment from childhood immunization in low- and middle-income countries, 2011–20. *Health Affairs*, 35(2), 199–207. Retrieved from <https://www.healthaffairs.org/doi/10.1377/hlthaff.2015.1086>.

Paxton, A., & Carvalho, N. (2013). *HIV/AIDS program costing tools: Concepts and methods used under the USAID Health Policy Initiative, Costing Task Order*. Washington, DC: Futures Group, Health Policy Initiative, Costing Task Order. Retrieved from http://www.healthpolicyplus.com/archive/ns/pubs/hpi/1551_1_Costing_Tool_Guide_FINAL.pdf.

Perkins, C., Steinbach, R., Tompson, L., et al. What is the effect of reduced street lighting on crime and road traffic injuries at night? A mixed-methods study. Southampton (UK): NIHR Journals Library; 2015 Sep. (Public Health Research, No. 3.11.) Chapter 4, Cost–benefit analysis: Methodological challenges of evaluating large-scale public health interventions and a worked example of the costs and benefits of part-night lighting. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK316508/>

Robinson, L. A., Hammitt, J. K., Chang, A. Y., & Resch, S. (2017). Understanding and improving the one and three times GDP per capita cost-effectiveness thresholds. *Health Policy and Planning*, 32(1), 141–145. Retrieved from <https://academic.oup.com/heapol/article/32/1/141/2555408>.

Rogers, P. (2012). InterAction Guidance Note 1: Introduction to impact evaluation. Retrieved from <http://www.managingforimpact.org/resource/interaction-guidance-note-1-introduction-impact-evaluation>.

Salomon, J. A., Haagsma, J. A., Davis, A., de Noordhout, C. M., Polinder, S., Havelaar, A. H., ... Vos, T.. (2015). Disability weights for the Global Burden of Disease 2013 study. *The Lancet Global Health*, 3(11), e712–723. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/26475018>.

Sassi, F. (2006). Calculating QALYs, comparing QALY and DALY calculations. *Health Policy and Planning*, 21(5), 402–408. Retrieved from <https://academic.oup.com/heapol/article/21/5/402/578296>.

Serrat, O. (2009). *The most significant change technique*. Washington, DC, USA: Asian Development Bank. Retrieved from <https://digitalcommons.ilr.cornell.edu/intl/196/>.

Shepard, D.S., & Thompson, M.S. (1979). First principles of cost-effectiveness analysis in health. *Public Health Reports*, 94(6), 535–543. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1431742/>.

Shiell, A., Donaldson, C., Mitton, C., & Currie, G. (2002). Health economic evaluation. *Journal of Epidemiology & Community Health*, 56(2), 85–88. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1732075/>.

Shillcutt, S. D., Walker, D. G., Goodman, C. A., & Mills, A. J. (2009). Cost-effectiveness in low- and middle-income countries: A review of the debates surrounding decision rules. *Pharmacoeconomics*, 27(11), 903–917. Retrieved from <https://link.springer.com/article/10.2165%2F10899580-000000000-00000>.

Slothuus, U. (2000). *An evaluation of selected literature on the measurement of costs in health economic evaluation*. Health Economics Papers 3. Odense, DK: University of Southern Denmark. Retrieved from https://www.sdu.dk/-/media/files/om_sdu/centre/cohere/working+papers/20003.pdf.

SoCha LLC. (2016). Saving Mothers, Giving Life Phase 1 in Uganda: A cost-effectiveness analysis.

Spiegelman, D. (2016). Evaluating public health interventions: 1. Examples, definitions, and a personal note. *American Journal of Public Health*, 106(1), 70–73. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4695951/>.

- Stover, J. (2003). AIM. Version 4. A computer program for making HIV/AIDS projections and examining the social and economic impacts of AIDS. Retrieved from <https://www.avenirhealth.org/software-spectrum.php>
- United States Agency for International Development Global Health eLearning Center. (2011). *Economic evaluation basics* [Online course]. Retrieved from <https://www.globalhealthlearning.org/course/economic-evaluation-basics>.
- United States Department of Transportation. (2016). Revised departmental guidance 2016: Treatment of the value of preventing fatalities and injuries in preparing economic analysis. Retrieved from <https://cms.dot.gov/sites/dot.gov/files/docs/2016%20Revised%20Value%20of%20a%20Statistical%20Life%20Guidance.pdf>.
- United States Environmental Protection Agency (USEPA). (2010). Guidelines for preparing economic analyses. Retrieved from <https://www.epa.gov/environmental-economics/guidelines-preparing-economic-analyses>.
- USEPA. (2016). Recommended income elasticity and income growth estimates: Technical memorandum. Retrieved from [https://yosemite.epa.gov/sab/SABPRODUCT.NSF/81e39f4c09954fcb85256ead006be86e/0CA9E925C9A702F285257F380050C842/\\$File/Income+Elasticity+Technical+Memorandum_final_2_5_16_docx.pdf](https://yosemite.epa.gov/sab/SABPRODUCT.NSF/81e39f4c09954fcb85256ead006be86e/0CA9E925C9A702F285257F380050C842/$File/Income+Elasticity+Technical+Memorandum_final_2_5_16_docx.pdf).
- Vassall, A., Sweeney, S., Kahn, J., Gomez, G. B., Bollinger, L., Marseille, E., ...Levin, C. (2017). Reference case for estimating the costs of global health services and interventions. Retrieved from https://ghcosting.org/pages/standards/reference_case.
- Viscusi, W. K., & Masterman, C.J. (2017). Income elasticities and global values of a statistical life. *Journal of Benefit-Cost Analysis*, 8(2), 226–250. Retrieved from <https://www.cambridge.org/core/journals/journal-of-benefit-cost-analysis/article/income-elasticities-and-global-values-of-a-statistical-life/5AE299883F668DCC265C41A377E1E063>.
- Waters, H., Abdallah, H., & Santillán, D. (2001). Application of activity-based costing (ABC) for a Peruvian NGO healthcare provider. *The International Journal of Health Planning and Management*, 16(1), 3–18. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/hpm.606>.
- Weatherly H, Drummond M, Claxton K, Cookson R, Ferguson B, Godfrey C, Rice N, Sculpher M, Sowden A. (2009), Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. *Health Policy*, 93(2-3):85–92 . Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/19709773>
- Whitehead, S. J., & Ali, S. (2010). Health outcomes in economic evaluation: The QALY and utilities. *British Medical Bulletin*, 96(1), 5–21. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/21037243>.
- Wilson-Grau, R. (2015). Outcome harvesting. Better Evaluation. Retrieved from https://www.betterevaluation.org/en/plan/approach/outcome_harvesting.
- World Health Organization (WHO), Commission on Macroeconomics and Health. (2001). *Macroeconomics and health: Investing in health for economic development*. Geneva, Switzerland: WHO. Retrieved from <http://www1.worldbank.org/publicsector/pe/PEAMMarch2005/CMHReport.pdf>.

Zimmerman, J. (2003). *Accounting for decision-making and control. International Edition.* (Fourth Edition). Boston, MA: McGraw-Hill Irwin.

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. MS-19-162
ISBN: 978-1-64232-135-7

