

Evaluation FAQ: How Much Will an Impact Evaluation Cost?

The cost of an impact evaluation is dependent on many factors, each particular to the program being evaluated, the outcomes of interest, and the local context for the work. Other factors are particular to the evaluation process—for example, the front-end work required for planning an evaluation is often substantial and should be considered in budgeting for an evaluation. Some considerations for cost include the following.

- **Evaluation questions:** The number and type of evaluation questions will affect costs. A large number of evaluation questions will typically increase sample size and/or the number of different data collection instruments needed. Evaluation questions that require comparing different combinations of activities imply multiple evaluation arms, which increases cost. Questions on differential impact of interventions on different target populations typically require a larger sample size.
- **Implementing program features:** The primary outcome(s) to be measured and how fast the outcome is likely to change will have implications for sampling. The scale of the program and the timeline for implementation will drive choices about design, as will additional program features relevant to (a) identifying a suitable comparison group (see below), (b) understanding the impact pathway, and (c) controlling for potential confounding factors.
- **Evaluation design:** The critical design feature for an impact evaluation is the identification of a comparison group that will provide a

counterfactual situation. The sample size, geographic coverage, and the types of data collection needed for program and non-program populations (e.g., population-based survey, facility-based survey, in-depth participant interviews) all influence costs. Keep in mind that you typically need a larger sample size to detect statistically significant change in an outcome than would be needed for a point estimate at the same level of precision, particularly if the outcome is relatively rare or is likely to change slowly. Estimating the difference in change in an outcome between program and non-program areas typically requires even larger sample sizes. Including biomarkers also adds to costs, often substantially.

- **Country context:** Local data collection costs vary across countries, depending on local capacity for data collection, transport costs, ethics review costs, etc. It is important to include an explicit emphasis on local capacity building and data use, but note that this step adds further to the cost, depending on the extent of such activities. It is also important to highlight the commitment to disseminate and act on the evaluation findings, and so one should factor in costs for engaging stakeholders in design and intended data use.

Keeping in mind the breadth of options available, many impact evaluations are focused on change in a particular health behavior or outcome at the population level. A typical design for such evaluations includes conducting baseline and endline household surveys in program and non-program areas. The sample size for such surveys is typically large (1000s).

Table 1 provides examples of costs from recent evaluation studies conducted by MEASURE Evaluation. This table is meant to illustrate the range of likely costs, but each impact evaluation must be budgeted based on its own specifics. The examples below represent one round of data collection. However, typically one would need at least two rounds a few years apart for an impact evaluation. These costs do include both direct and indirect costs.

For more information

MEASURE Evaluation supports the international development goal of maximizing program results through the systematic collection and analysis of information and evidence about health program performance and impact. Evaluation findings provide accountability to stakeholders and donors, and inform decisions—supporting organizational learning, strengthening health systems, and improving program effectiveness so that people live healthier lives. See <http://www.cpc.unc.edu/measure/our-work/evaluation> or <http://www.cpc.unc.edu/measure/about/services>.

Table 1: Examples of Potential Impact Evaluation Costs

Study	Budget in USD (approx.)	Comments
Bangladesh Smiling Sun Franchise Endline Survey	\$850K	Endline household survey in program and non-program areas (national) focused on key maternal and child health (MCH) indicators.
Guatemala Feed the Future (FTF) baseline survey	\$3.5mil	Baseline household survey in program and non-program areas (subnational) focused on MCH and nutrition outcomes. Difficult field terrain and long instrument. \$2.4 million for in-country data collection costs.
Nigeria COMPASS program endline survey	\$400K	Endline household, facility, and school surveys in five program local government areas. No comparison areas. Limited technical assistance.
Ukraine tuberculosis impact evaluation baseline data collection	\$400K	Facility-based sampling for retrospective medical record chart abstraction plus facility survey in program and non-program areas (subnational). Focused on two research questions requiring different sampling and follow-up plans.
Jamaica randomized controlled trial for HIV prevention interventions	\$552K	Cluster randomized controlled trial with baseline and follow-up surveys. Site-based sampling of patrons.

measure@unc.edu

www.measureevaluation.org

