METHODS BRIEF

PCE Impact Evaluation Methodological Summary

David E Phillips PhD, On behalf of the IHME/PATH PCE Consortium and Starley Shade MPH, PhD, On behalf of the EGH/UCSF PCE Consortium

> This is a working document and may be updated as specific details in each country emerge.
> February 25, 2019

1 Purpose of this Document

The purpose of this document is to explain the analytical approach which we have selected for impact evaluation, with the goals of methods improvement through feedback, transparency, and demonstration of the robustness (both strengths and limitations) of the methods. A complete summary of the final format for PCE reporting, qualitative methods and work plans and timelines may be better suited for a separate document.

The intended audience for this document was originally the PCE consortium of IHME, PATH, CIESAR, IDRC and PATH DRC, drafted in April 2018. It has been updated to include all eight PCE countries, with the intended audience being both GEPs, all CEPs, the TERG and TERG Secretariat.

2 Overview

The basic approach to this impact evaluation is to take a differentiated approach by disease and country. This can be summarized as three separate approaches, brought together by a common analytical framework (the PCE Theory of Change and Results Chains, described elsewhere):

1. Mixed methods analysis of results chains (alone)
2. Targeted impact analysis within results chains
3. Comprehensive impact modeling along results pathways

Fundamentally, all three approaches are to measure many separate indicators along the results chains, then either discuss or measure their relationships. The added value of the PCE approach is to do so with:

1. A high level of detail (both in terms of number of indicators in each section of the results chain, and subnational resolution)
2. Complementary mixed methods (i.e. qualitative information that explains and adds depth to quantitative findings)
3. Attention to, *and corrections for*, data quality that go beyond simply taking reported data at face value

Control variables are especially critical for impact analysis. The basic approach of "measuring many indicators" also includes measuring covariates and controls that are not necessarily the indicators of interest, but are essential to understanding the relationship between expenditure, outputs and outcomes. The PCE will rely on internationally-vetted measures of controls as much as possible.

3 Subject of the Impact Evaluation

The PCE impact evaluation is to provide answers to four of the 13 PCE synthesis evaluation questions. These are:

- What are the trends and distribution of HIV, TB and malaria-related health outputs and outcomes?
- To what extent do Global Fund resources contribute to improvement in health outputs and outcomes for HIV, TB and malaria?
- How well are key and vulnerable populations defined and effectively addressed through Global Fund investments?
- What are the trends and distribution of Global Fund resources and how do they compare with need?

The core hypothesis of this impact evaluation is that changes in Global Fund investments result in observable changes in health systems outputs. Additionally, the hypothesis is that those changes in outputs result in observable changes in intervention coverage, which subsequently result in improvements in burden of disease. The PCE will take an approach of analyzing all available data (historically and along the results chain), with an emphasis on the most recent data and subnational data wherever possible.

Changes in Global Fund investments cannot be analyzed without also considering investments from other development partners, as well as government and private/out-of-pocket health expenditure however. Furthermore, individual outputs are often impossible to attribute specifically to one source of funding, given the role of partnerships in investment decisions and the pooled nature of resources that channel through national programs.

For these reasons, the subject of the impact evaluation is twofold:
1 The Global Fund's contribution to national program activities and outputs
2 The national program (and wider effort to fight the three diseases)'s impact on burden of disease

4 Differentiated Approaches by Country and Disease

Figure 1 displays the anticipated approach that will be taken in each country for each disease. These selections were primarily based on CEP knowledge of country context and data (see Section 5). To be clear, the approach for every disease is to analyze indicators along the results chains. The differentiation of approaches pertains more to what the PCE will do with the results of those individual indicators.

Figure 1 Anticipated Approaches to Impact Evaluation

	HIV	TB	Malaria
DRC	*		
GTM	+		
KHM			
MOZ			
MMR			
SEN	+		
SDN			
UGA	*	*	

*Known limitations to data may result in somewhat limited analysis
+Known strengths in data may result in more extensive analysis

Key

	Mixed methods analysis of results chains
	Targeted impact analysis
	Comprehensive impact modeling

5 Motivation for Differentiated Approaches

There are several reasons why the PCE has opted to take a different approach by disease and country. First and foremost is in an effort to provide results that are relevant to the broad array of audiences who the results might reach. While an overall assessment of the steps leading from Global Fund investment to impact may be relevant and useful to stakeholders in the Global Fund Secretariat, a more tailored analysis focused on isolated sections of the results chains may be more useful to national programs.

Second is the feasibility of adding value through impact evaluation. Differences in data availability (indicators tracked and level of disaggregation), data quality, and basic differences in epidemiology limit our ability to apply a generic approach to all diseases in all countries. Furthermore, extensive impact evaluation has already been (or is being) conducted by other organizations for certain diseases in certain countries. For these reasons, the PCE seeks to evaluate different diseases in different ways in order identify applications of existing data that are both possible given known limitations and actually useful to stakeholders.

6 Mixed Methods Analysis of Results Chains Alone

The first approach to assessing impact is descriptive analysis of indicators in a sequential fashion along the results chain, with accompanying narrative to explain and contextualize the trends. This approach is now familiar to the PCE and its audiences at time of writing, as it has been presented in each of the 2018/19 annual country reports and their synthesis. We caveat this approach with *alone*, because all approaches fundamentally provide analysis along the results chain in that they use it as their basic analytical framework.

The advantages of this approach are its breadth. While the dose-response analysis (described below) provides substantial detail about the relationships between variables, it is constrained to those variables which (according to theory) have a corresponding indicator before or after it that can also be measured. Approaching the results chains in a more narrative way enables deeper dives into subpopulations and indicators that may be more isolated quantitatively, but are important nonetheless.

7 Targeted Impact Analysis within Results Chains

The second approach to impact evaluation constitutes the most varied category. These analyses will focus on a key section or sections of the results chain without necessarily analyzing all sections comprehensively, instead proving further depth on that area. It may also be specific to subnational areas, subpopulations or country policies and characteristics that cannot easily be replicated in other countries. Examples of these approaches are still emerging in each country but may include:

- Mixed methods analysis of findings arising from the subnational resource tracking study in Uganda. In this context, stakeholders identified a need for understanding financial flows at the district level. The PCE selected six districts to explore in depth to understand how resources are allocated, spent and translating into outputs at the local level. These districts may not be generalizable to the entire country however, so catered analyses will be developed to understand the relationship between inputs, outputs and impact these districts alone.

- Malaria elimination analysis in Guatemala. Given the epidemiological status of malaria in the Guatemalan context, prevention and community case management activities are focused on very specific hot spots where parasite prevalence remains higher. Separate modeling approaches are being designed to understand the second half of the results chain in this case specifically.

- Rationalization analysis in DRC. In the DRC context, international donors and the government have undergone a process known as rationalization, whereby donors coordinate which geographic areas their investments will be directed towards, and limit their activities elsewhere. This policy offers an unusual opportunity to evaluate donor investment as a "natural experiment", and analyses may be structured for this situation specifically.

These examples are indicative of the types of country-tailored impact evaluations that the PCE is exploring, and may not be the final analyses that the evaluators and stakeholders agree upon.

8 Comprehensive Impact Modeling along Results Pathways

The final approach will be to apply models that analyze the relationships among adjacent indicators along all sections of the results chain. This may be described as dose-response analysis with the idea that greater intensity of investment in one intervention is expected to produce more outputs in its associated indicator (and lower investment intensity should produce fewer outputs). What makes this more complex than a simple comparison of time trends is the fact that each intervention may realistically be expected to produce outputs in multiple indicators, and each activity/output indicator is expected to be impacted by investments in multiple interventions. For example, investment in facility-based treatment of malaria is expected to result in increases in RDTs used and ACTs used (not just one or the other), and so is integrated community case management. Without more detailed data that precisely tracks outputs that are the direct response to certain investments, this sort of many-to-many relationship is necessary to reflect in a model.

Thus, it will be essential to develop a model of the complex relationships between variables at each point in the results chain. Where similar disaggregations of data are available throughout the results chain, we will apply structural equation models. Where less detailed data are available at some levels, other models and causal inference approaches (e.g. g-computation) will by applied. Figure 2 displays an example of what such a model is expected to look like, focusing on the first half of the results chain.

Figure 2 Example of Impact Modeling along Results Pathways

*Inputs with no exiting arrows affect all activities
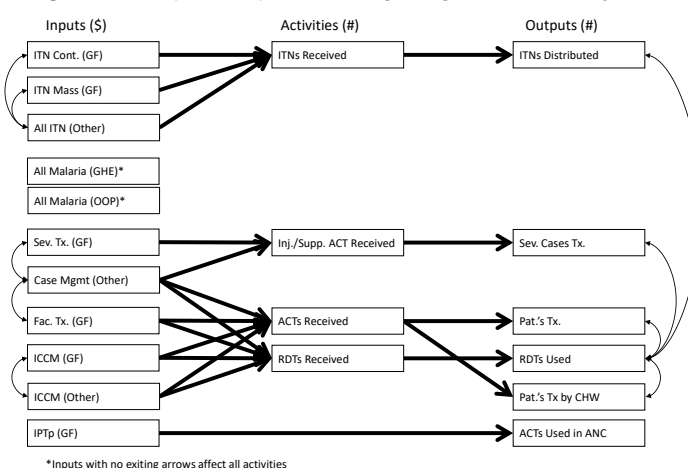
In the figure, each box represents an indicator that is being tracked by the PCE, and the arrows between them represent regression coefficients (i.e. correlations) describing how the indicators co-vary in time. Curved arrows represent correlated error terms, i.e. correlations without an explicit direction.

In this way, we can bring together the indicators the PCE is tracking along the results chains and represent their relationships using a structure that accurately

reflects their relationships according to our theory of change. For example, this allows us to specify which inputs are intended to impact which outputs, but also which inputs are intended (according to the PCE theory of change) to be allocated in partnership with other donors. A similar model relating indicators between outputs, outcomes and burden of disease will be constructed to represent the second half of the results chain, taking advantage of subnational data as much as possible to find correlations.

The PCE is already in the process of measuring these correlations, exemplified in Figure 3. This only displays three of the many interrelated indicators in the full model however, hence the need for a more complex approach.

Figure 3 Example of Pairwise Correlations as Part of Impact Model
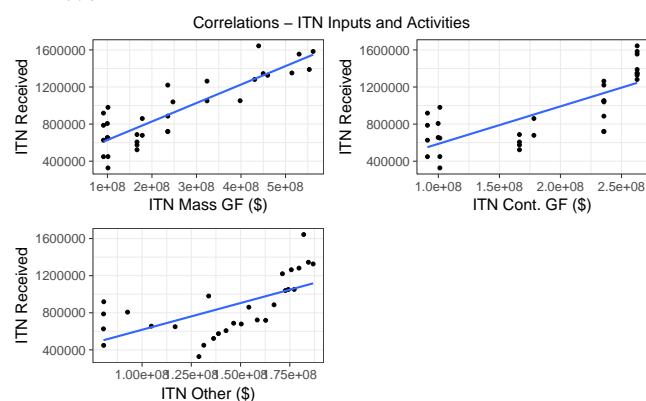
The final product of the analysis will focus less on the measurement of coefficients however, and more on the implications of the estimates. We will use counterfactual analysis to describe the anticipated changes in outputs that the model predicts will result from an alternative Global Fund investment. In other words, the PCE will construct a realistic alternative grant budget, and use the model to evaluate the changes in outputs, outcomes and burden of disease that would be expected to result from it.

This model relies on several conditions to be reliable. First, the breadth of indicators included in it serves as the most critical constraint to its generalizability. Because most available program data tracks countable commodities, many of the non-commodity interventions may be left out of the model and therefore not reflected in the counterfactual analyses. Second, including specific control variables will be essential. Indicators of varying data quality, varying unit costs and confounding socio-demographic indicators will be

included when possible in order to control for trends that may otherwise explain the relationships between sections of the results chain.

8.1 Extension to RSSH

In evaluating the contribution of catalytic and system-wide Global Fund investments (such as investments to strengthen resilient and sustainable systems for health (RSSH)), an additional layer of controls may be necessary. This is because such investments are intended to operate in addition to, or synergistically with, other program areas, not to result in outputs on their own.

Essentially, this amounts to controlling for spending on RSSH as well. The model will be tested with and without the inclusion of RSSH investments and the difference will be assessed in terms of expected outputs and outcomes, as well as total variance in the indicators explained by the model.

8.2 Relationship with Value for Money

Impact model has a natural relationship with value for money (VfM) assessment. Focusing on the first half of the results chain, the regression coefficients in the model can be interpreted as cost per output, or efficiency. Focusing on the second half of the results chain, the coefficients can be interpreted as impact per output, or effectiveness. The model will be explored through the lens of VfM in order to explore the relative efficiency and effectiveness of some investments compared to others.

9 Mixed Methods

Each of the above approaches to quantifying impact will be complemented by qualitative evaluation techniques which take a "deep dive" approach. While a complete description of those approaches is not provided here, complex, unexpected, or otherwise noteworthy observations from the results chains and impact evaluations will be identified by the PCE, and explored in greater depth in order to understand the "why" behind the indicators and arrows connecting them. This qualitative follow-up serves the essential function of understanding the reasons behind the trends and correlations seen in the quantitative data. This follow-up process will simultaneously serve to understand the ways in which PCE thematic areas (e.g. gender and human rights, sustainability transition and co-financing etc.) intersect with the Global Fund's investments. This process will take place regardless of the impact evaluation approach used for each country and disease.

10 Limitations of Impact Evaluation

Many of the limitations to the impact evaluation approaches have been described above, but are important to highlight.

First, the quantitative indicators measured through the results chains often focus on countable commodities and may miss outputs from other types of interventions. More generally however, all of the impact evaluation approaches are limited by the quality and quantity of available data. As the PCE has already learned and discussed in annual country reports, data systems vary between countries and diseases, and offer greater or weaker insight into the ongoing national efforts to fight the three diseases depending on the number of indicators they track (and make available), the level of geographic and demographic detail reported, the quality of the reports and the timeliness with which data are reported. While the PCE will take every effort to correct for data quality, limited data availability and quality in some cases may require the PCE to use either existing model estimates or aggregates instead of direct program reports. In other cases, analysis may necessarily focus on the areas of the results chain with the most reliable data.

Another limitation relates to the mixed methods nature of the results. The PCE plans to follow some deep dives in each country, but cannot feasibly perform qualitative follow-up to explain the "why" behind every noteworthy observation from the results chains and impact evaluation results. A process will be followed to identify deep dive areas that satisfy the needs of country stakeholders, fill information gaps, and explain significant portions of the results chains, but will not comprehensively span all parts of the results chain.

Finally, although the second and third approaches are intended to provide methods that accurately reflect reality, they are necessarily a simplification. Dynamics such as time-lagged feedback between prevention and treatment activities, burden of disease and subsequent spending may be too complex to fully account for in the model, let alone the full range of activities. For that reason, the counterfactual results must be interpreted as instructive information that accounts for as much as possible, but not necessarily as causal effects. In the end, lack of stronger study design may make it unrealistic to confidently make broad statements such as "per dollar invested by the Global Fund, X lives are saved" using this approach.