



Value for Money Working Group
Background Paper

Measurement Obstacles to Achieving Value for Money at the Global Fund: A Problem Statement

Rachel Silverman and Amanda Glassman

Background paper prepared for the
Working Group on Value for Money:
An Agenda for Global Health Funding Agencies

Global Health Policy Program
Center for Global Development

October 2013

This report is a product of the Global Health Policy Program of the Center for Global Development, and a background paper for the Value for Money Working Group. The findings, interpretations, and conclusions expressed in this volume do not necessarily reflect the views of the Center for Global Development or Working Group members. Portions of this background paper have been incorporated into the final report of the Working Group, released September 2013. The authors are grateful to Victoria Fan, Mead Over, and Denizhan Duran for helpful comments and input, and to Global Fund staff for helping to clarify existing processes. All errors or omissions are our own.

Abstract

The Global Fund is currently finalizing design and implementation of its New Funding Model (NFM), which includes a focus on strengthened measurement and an impact-based investment strategy. To help frame the Fund's reform agenda in this area, we first consider the inherent measurement challenges faced by programs to address HIV, tuberculosis, and malaria, including those that arise from the competing demands for data by a broad range of constituencies. Next, we attempt to define a clear problem statement by outlining the Global Fund's heavy reliance on data and measurement as core components of its business model, while identifying critical deficiencies in the Global Fund's historical system of measurement. Finally, we situate the Global Fund within different perspectives on how measurement in health financing should operate, including the benefits and limitations of each vision.

List of Acronyms and Abbreviations

AIDS	Acquired immunodeficiency syndrome
API	Application programming interface
ART	Antiretroviral therapy
ARV	Antiretroviral
COD	Cash on Delivery
DQA	Data quality audit
FPM	Fund portfolio manager
HIV	Human immunodeficiency virus
HMIS	Health management information system
IHP+	The International Health Partnership+
IRS	Indoor residual spraying
ITN	Insecticide treated net
LFA	Local Fund Agent
M&E	Monitoring and evaluation
MERG	Monitoring and Evaluation Reference Group
MSM	Men who have sex with men
NFM	New Funding Model
OIG	Office of the Inspector General
OSDV	On-site Data Verification
PBF	Performance-based financing
PEPFAR	President's Emergency Plan for AIDS Relief
PMI	President's Malaria Initiative
PMTCT	Prevention of mother to child transmission
PQR	Price and quality reporting tool
PR	Principal Recipient
RSQA	Rapid Services Quality Assessment
TB	Tuberculosis
TRP	Technical Review Panel
UNAIDS	Joint United Nations Programme on HIV/AIDS
USAID	United States Agency for International Development
WHO	World Health Organization

1. Introduction

Over the past two years, the Global Fund to Fight AIDS, Tuberculosis and Malaria (the ‘Global Fund’ or the ‘Fund’) has undergone an intensive reform process, culminating in the design and implementation of its New Funding Model (NFM). Among a broad range of reform priorities, the Global Fund has repeatedly emphasized the importance of better measurement and increased transparency as part of a comprehensive investment agenda. For example, the five-year strategy for 2012-2016 included a provision to “systematically invest in high-quality data through baseline and progress surveys, data modeling, and...increased transparency of financial data,”¹ while the Consolidated Transformation Plan recognized that better outcome and impact measurement was essential to “further inform strategic decision-making.”² Likewise, upon taking office in January 2013, incoming Global Fund Executive Director Mark Dybul listed “impact measurement” as one of six priority areas that the Global Fund would address under his leadership. “It is good management to know about impact and to evaluate and integrate it into operations,” he wrote. “Donors to global health efforts need to know what impact their investments are having, and need to assure taxpayers that their investments are effective.”³

The Fund’s newfound emphasis on better measurement is responsive to the 2011 High Level Panel Report, which found that “the culture of the Global Fund has become one driven by the measurement of documentation, and not by health impact.”⁴ Historically, the Global Fund has focused its measurement and reporting on those ‘results’ that could be easily documented (i.e. number of training sessions, condoms distributed, etc.),⁵ rather than the most epidemiologically important – albeit more difficult to measure – outcome and impact indicators (i.e. patients retained on antiretroviral therapy (ART), lives saved, or infections averted). Yet this challenge has not been unique to the Global Fund; to the contrary, many global health funders have grappled with the difficulty of improving measurement and moving toward a results-based focus. Previous such efforts include the Joint United Nations Programme on AIDS (UNAIDS) Monitoring and Evaluation Reference Group, which included participation from the Global Fund and has defined 30 core indicators;⁶ the World Economic Forum’s Global Health Data Charter;⁷ and the International Health Partnership+ (IHP+) in cooperation with the World Health Organization (WHO), which proposed a “country-led [monitoring and evaluation (M&E)] platform for information and accountability.”⁸

While the Global Fund has committed to strengthened impact measurement under its New Funding Model, many details remain undefined. This paper aims to contribute to the ongoing dialogue around measurement at the Global Fund in three key ways. First, it considers the inherent measurement challenges

¹ The Global Fund (2011). The Global Fund strategy 2012-2016: Investing for impact. Strategic action 3.1.

² The Global Fund (2011). Consolidated transformation plan. Board Decision GF/B25/4. Twenty-fifth board meeting. November 21-22. Accra.

³ Dybul M (2013). The big push to defeat AIDS, TB, and malaria. The Blog. Huffington Post. Accessed 2 September 2013 at http://www.huffingtonpost.com/amb-mark-dybul/the-big-push-to-defeat-ai_b_2516889.html.

⁴ The Global Fund (2011). Turning the page from emergency to sustainability. The final report of the High-Level Independent Review Panel on Fiduciary Controls and Oversight Mechanisms of the Global Fund to Fight AIDS, Tuberculosis, and Malaria.

⁵ See the Global Fund (2012). Strategic investments for impact. Global Fund results report 2012. Page 15.

⁶ UNAIDS (2011). Global AIDS response progress reporting 2012. Guidelines for construction of core indicators for monitoring the 2011 political declaration on HIV/AIDS. Accessed 3 October 2013 at

http://www.unaids.org/en/media/unaids/contentassets/documents/document/2011/JC2215_Global_AIDS_Response_Progress_Reporting_en.pdf

⁷ World Economic Forum (2010). Global health data charter. Geneva: World Economic Forum in collaboration with Deloitte Touche Tohmatsu.

⁸ IHP+ and the World Health Organization (2011). Monitoring, evaluation, and review of national health strategies. A country-led platform for information and accountability. Geneva: World Health Organization.

faced by programs to address HIV, tuberculosis, and malaria, including those that arise from the competing demands for data by a broad range of constituencies. Next, it attempts to define a clear “problem statement” by outlining the Global Fund’s heavy reliance on data and measurement as core components of its business model, while identifying critical deficiencies in the Global Fund’s historical system of measurement. Finally, it situates the Global Fund within different perspectives of how measurement should operate, including the benefits and limitations of each vision.

2. Inherent Measurement Challenges

Competing Data Demands

One inherent challenge of measurement for the Global Fund is that the organization is asked to generate many different kinds of data, for many different uses, to satisfy the data needs of many different constituencies. By satisfying one audience or ‘need’ for measurement or data, the Global Fund risks neglecting or aggravating another. However, it may not be affordable or feasible for the Global Fund to fully satisfy the data needs of every potential constituent. At worst, such a broad and unfocused approach could lead the Fund to spend significant resources on high quantity but low-quality data collection that is ultimately of little use to any of its relevant constituencies.

For this reason, it is important to think critically about the different potential constituents for and uses of Global Fund data (Table 1). This presentation is simplified and non-comprehensive; in reality, different actors are likely to have multiple demands for data, which often overlap with each other. Nonetheless, it clearly speaks to the many potential uses for data, and the different kinds of measurement required to meet them all.

Table 1: Competing Demands for Data from Global Fund Programs

Constituent	Motivation	Data Needs
Beneficiaries/ Watchdogs	<ul style="list-style-type: none"> • Are grants funding the right things? • Is money being spent most effectively? 	Grant data, results, impact
Principal/ Sub-Recipients	<ul style="list-style-type: none"> • Can we manage our facilities/grants? 	Health management information system (HMIS) data, routine monitoring data
Recipient Governments	<ul style="list-style-type: none"> • Do we know what we need to inform national strategies and planning? 	Geographic and service delivery area allocation data, monitoring and surveillance data
Secretariat	<ul style="list-style-type: none"> • Can I effectively manage the grants in my portfolio? • Can I detect problems in implementation or management? • Can I implement performance-based financing? 	Grant performance, contractual compliance, fiduciary controls, results
WHO/UNAIDS	<ul style="list-style-type: none"> • Can we report overall trends and progress toward the Millennium Development Goals? 	Detailed country-level statistics,

	<ul style="list-style-type: none"> • Can we fulfill our disease modeling and surveillance responsibilities? 	allocation data
Donors	<ul style="list-style-type: none"> • Can we demonstrate accountability to our taxpayers? 	Spending, fiduciary controls, results, impact

Outcome Measures are Unavailable or Lagged

To effectively manage its grants, the Global Fund needs real time information on their performance. This need, however, is in tension with the organization’s expressed desire to move towards a focus on outcomes and impact. Particularly for HIV/AIDS, the most important outcome and impact measures (i.e. adherence, mortality, and incidence) can only be observed with follow-up after several years. Further, as Laga et al. (2012) argue, “direct estimation through follow-up of a cohort is complex, costly, and unsustainable outside of research settings.”⁹ Regular population-level surveys can help to assess national or regional trends in mortality and prevalence, but may not be sufficient for some essential indicators. When the outcome of interest is a rare event, as with new HIV infections in most countries, population-level surveys would require unmanageably large sample sizes to detect a statistically significant change, and are likely to underrepresent most at risk populations.¹⁰ Finally, measurement challenges can be exacerbated by the lack of valid and reliable measurements instruments, such as incidence assays to measure the frequency of new infections.^{11,12}

Service Delivery is Complex

Even at the output level, measurement of service delivery can be quite challenging. Many health services are complex and difficult to describe, and different observable indicators could describe equally high quality of care. Work is currently underway to systematically link the observable characteristics of service delivery to patient outcomes (i.e. ART retention); however, while such research is instructive for high-level policy-making, it does not yet provide conclusive guidance on best practice at the clinic level.¹³ Historically, the Global Fund has circumvented this challenge through an output ‘tally’ approach that does not attempt to detect service delivery quality, i.e. with indicators such as the number of people receiving ART, or number of mothers treated to prevent mother to child transmission of HIV (PMTCT).

Budgets Are Intertwined

A final inherent challenge to measurement, which will reemerge later in this paper, is the question of attribution when budgets are intertwined. That is, how do you measure the Global Fund’s contribution to ART, for example, if the Global Fund is only financing one component of the health production function

⁹ Laga M, Rugg D, Peersman G, Ainsworth M (2012). Evaluating HIV prevention effectiveness: The perfect as the enemy of the good. *AIDS* 26: 000-000.

¹⁰ UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance (2011). When and how to use assays for recent infection to estimate HIV incidence at the population level.

¹¹ Laga M, Rugg D, Peersman G, Ainsworth M (2012). Evaluating HIV prevention effectiveness: The perfect as the enemy of the good. *AIDS* 26: 000-000.

¹² World Health Organization (2012). WHO Working Group on HIV Incidence Assays statement on the use of HIV-1 incidence assays for surveillance or epidemic monitoring.

¹³ Over M, Schneider M, Velayudhan T (Forthcoming). Explaining the variation in on-site AIDS treatment costs. Econometric results and policy implications the MATCH study of 161 facilities from five countries: (Multi-country Analysis of Treatment Cost for HIV). Draft reviewed October 2012.

(i.e. training)? If a single service delivery output is the product of several different funding sources, it may be near impossible to isolate the unique contribution or impact of Global Fund financing.¹⁴

3. Problem Statement

Already, the Global Fund understands that accurate data collection is essential for realizing at least four key aspects of its business model:¹⁵

1. **Sustainability and efficiency** — that is, strengthening national health information systems and other data collection to improve program management by the principal recipient and build sustainable health systems. To this end the Global Fund recommends that recipients allocate 5–10 percent of their budgets to M&E activities.¹⁶
2. **Resource allocation**, within and across countries, and both for Global Fund grants and for the national strategy plans that provide a starting point for grant negotiations. The Global Fund’s 2012–2016 strategy calls for “strategic investment” in “the highest-impact interventions and technologies suitable to the country situation,” and for “appropriate targeting of most-at-risk populations.”¹⁷ This can only be done with robust data on the efficacy of interventions (including efficacy for particular subgroups) and on the size and characteristics of a country’s epidemics, including high-risk groups and geographic “hot spots” of transmission.
3. **Grant management** by the Global Fund, encompassing risk mitigation, regular oversight, performance incentives, and iterative reprogramming as challenges or opportunities for greater impact arise. Under the Global Fund’s performance-based financing (PBF) system, accurate performance data are needed to inform the magnitude of subsequent disbursements. But this is only one aspect of the Global Fund’s reliance on data for grant management purposes. Beyond PBF, the Global Fund requires data and measurement to detect and deter fraud, assess overall epidemiological trends, revise its funded activities, and coordinate funding with other partners and national governments. All require real-time measurement of financial flows, implementation progress, and other aspects of grant performance.
4. **Accountability** — between principal recipients and the Secretariat, between the Secretariat and the Board, and between the Board and donor governments. Just as the Global Fund is responsible for preventing misuse of its funds, it also provides implicit promises to its stakeholders about what it will achieve with those resources— “[save] 10 million lives and [prevent] 140–180 million new infections from HIV/AIDS, tuberculosis, and malaria between 2012 and 2016.”¹⁸ To be accountable to its Board and donors (and to mobilize additional resources in future replenishments), the Global Fund must track progress on those goals and document the appropriate use of its resources to achieve health impact.

¹⁴ McCoy D, Jensen N (2012). Quantifying the Global Fund’s Contribution to Saving Lives: Methodological and Policy Issues. Nairobi: Aidspan.

¹⁵ The following section is also included and elaborated upon in the full report of the Working Group. See Glassman A, Fan V, Over M, et al. (2013). More health for the money: putting incentives to work for the Global Fund and its partners. Washington: Center for Global Development.

¹⁶ The Global Fund (2011). Monitoring and evaluation toolkit. Accessed 20 June 2013 at <http://www.theglobalfund.org/en/me/documents/toolkit/>

¹⁷ The Global Fund (2011). The Global Fund strategy 2012-2016: Investing for impact.

¹⁸ The Global Fund (2011). The Global Fund strategy 2012-2016: Investing for impact.

Given this system wide reliance on data and measurement as an integral input to core Global Fund and country-level health objectives, it is not surprising that “data quality” has been a recurring concern in Global Fund policies. The Global Fund has responded to this challenge with several procedures designed to assess and improve the accuracy and reliability of the information on which it bases many key decisions. For example, the Global Fund has adopted a “risk management approach” to implementing Data Quality Audits (DQAs) among its grants, to “provide an in-depth assessment of data quality and M&E systems” for grant recipients.¹⁹ The Global Fund has also planned “country reviews” for recipients of its largest grants, designed to “evaluate disease outcome and impact, review program progress, and provide practical recommendations on where to achieve the greatest impact,” which are expected to inform program design under the new funding model.²⁰

Beyond these initiatives, routine performance validation by Local Fund Agents (LFAs) has long been part of Global Fund oversight practice. Principal Recipients (PRs) provide the Global Fund with periodic reports on grant implementation, including progress on country-chosen indicators and targets. These indicators often emphasize easily documented inputs and outputs (people trained, condoms distributed)²¹ rather than downstream health effects (outcomes, impacts). Once submitted, these reports are forwarded to the Global Fund’s designated LFA, typically an audit or consulting firm, which the Global Fund contracts to “independently oversee program performance” and “verify results.”²² For most periodic reports, LFA “verification” is conducted through a desk review of data sources, in which aggregate results are compared with the underlying documentation from facilities and program managers.²³ LFAs also conduct annual site visits for each disease area and principal recipient to verify data sources and to assess the quality of health services, both as described in official policy (usually at the Ministry of Health) and as followed in practice (at health facilities).²⁴

Yet despite this attention, the Fund’s current system of measurement does not enable it to optimize value for money in its grant-making. Below we identify six key problems with current practice that require urgent attention: 1) fragmented data collection; 2) low-utility data; 3) inability to attribute results or impact; 4) ‘audit’ focus; 5) perverse incentives; and 6) data inaccessibility.

Fragmented Data Collection

Already, the Global Fund and other donors finance and mandate a wide array of data collection. If the Fund consistently enforced its recommendation that 5-10% of grant money be budgeted for M&E, approximately \$250 million of Global Fund financing would have been allocated to M&E activities in 2012.^{25,26} In addition, LFA fees totaled an extra \$41 million, or 38% of the Secretariat’s overall annual budget.²⁷ Using these resources, PRs must report results against set targets every 3, 6, or 12 months, subject

¹⁹ The Global Fund. Data quality tools and mechanisms. Accessed 3 October 2013 at <http://www.theglobalfund.org/en/me/documents/dataquality/>

²⁰ The Global Fund (2013). Transition manual for the New Funding Model of the Global Fund.

²¹ The Global Fund (2012). Strategic investments for impact: Global Fund results report 2012.

²² The Global Fund. Local fund agents. Accessed 3 October 2013 at <http://www.theglobalfund.org/en/lfa/>

²³ The Global Fund (2013). Progress update/disbursement request guidelines. Accessed 3 October 2013 at http://www.theglobalfund.org/documents/core/guidelines/Core_PUDDR_Guidelines_en/

²⁴ The Global Fund (2011). LFA guidelines for On-site Data Verification (OSDV) and Rapid Services Quality Assessment (RSQA) implementation.

²⁵ The Global Fund (2011). Monitoring and evaluation toolkit. Accessed 20 June 2013 at <http://www.theglobalfund.org/en/me/documents/toolkit/>

²⁶ The Global Fund (2013). Global Fund news flash: issue 13. Blog, 17 January 2013. Accessed 3 October 2013 at <http://www.theglobalfund.org/en/blog/31214/>

²⁷ Global Fund (2012). Report of the General Manager. Twenty-Seventh Board Meeting.

to LFA “verification.”²⁸ Beyond this routine reporting, there have also been significant investments in household surveys, surveillance, and operational research using Global Fund financing, as well as through funding from other donors such as the US President’s Emergency Plan for AIDS Relief (PEPFAR), the President’s Malaria Initiative (PMI), UNAIDS, the United States Agency for International Development (USAID), and others.

Despite these sizable investments and past efforts to streamline data collection and harmonize indicators, however, the Global Fund’s data collection remains fragmented, both in policy and practice. The Global Fund has defined a “Top Ten Indicator” list for routine reporting and program impact (Appendix A), and updated its official M&E guidance in 2011. However, these two documents are not entirely consistent with each other, nor do they perfectly align to the indicators reported within the Fund’s 2012 results report (see Table 2, for example). Very few indicators are perfectly consistent in definition or wording across all three documents, and many indicators – two of the “Top 10,” 121 in the M&E guidance, and three in the 2012 results report – appear in only one of the three documents.²⁹

Table 2: Sample Indicator Inconsistencies within Global Fund Guidance

Top 10 ³⁰	Results Report ³¹	M&E Guidance ³²
<ul style="list-style-type: none"> Number of condoms distributed 	<ul style="list-style-type: none"> Condoms distributed 	<p>Routine Outputs</p> <ul style="list-style-type: none"> None <p>Impact/Outcome/Coverage</p> <ul style="list-style-type: none"> Percentage of women and men aged 15-49 years who have had more than one sexual partner in the past 12 months reporting the use of a condom during their last sexual intercourse Percentage of male and female sex workers reporting the use of a condom during penetrative sex with their most recent client Percentage of men reporting the use of a condom the last time they had anal sex with a male partner Percentage of people who inject drugs who reported the use of a condom the last time they had sexual intercourse
<ul style="list-style-type: none"> Number of women and men aged 15–49 years who received an HIV test in the last 12 months and who know their results 	<ul style="list-style-type: none"> HIV testing and counseling sessions provided 	<p>Routine Outputs</p> <ul style="list-style-type: none"> Number of people tested and counseled for HIV and who received results Number and percentage of pregnant women attending ANC whose male partner was tested for HIV <p>Coverage/Outcome/Impact</p> <ul style="list-style-type: none"> Percentage of women and men aged 15-49 who received an HIV test in the last 12 months and who know their test results Percentage of sex workers that received an HIV test in the last 12 months and who know their test results Percentage of men who have sex with men that received an HIV test in the last 12 months and who know the results Percentage of people who inject drugs that received an HIV test in the last 12

²⁸ Global Fund (2012). Performance-based disbursements. Accessed 28 December 2012 at <http://www.theglobalfund.org/en/performancebasedfunding/grantlifecycle/3/>

²⁹ This figure should be considered an estimate, due to ambiguity in the number of indicators based on various definitions and disaggregations. However, this is a relatively generous count, as indicators are classified as “consistent” even if they have different wording, or if one is coverage and the other is just the numerator.

³⁰ Global Fund (2011). Top ten indicators for routine Global Fund reporting. Accessed 3 October 2013 at http://www.theglobalfund.org/documents/monitoring_evaluation/ME_TopTenIndicators_Card_en/

³¹ The Global Fund (2012). Strategic investments for impact: Global Fund results report 2012.

³² The Global Fund (2011). Monitoring and evaluation toolkit. Accessed 20 June 2013 at <http://www.theglobalfund.org/en/me/documents/toolkit/>

Likewise, the Global Fund’s M&E guidance is not fully harmonized with other international organizations, particularly PEPFAR and the UNAIDS/Monitoring and Evaluation Reference Group (MERG) Core Indicators for the Global AIDS Response. Three of the Fund’s HIV-specific Top 10 indicators are not found in the official guidance for either of its aforementioned counterparts.³³ For example, while all three documents contain at least one indicator relating to HIV/TB, none are comparable, even given relatively lax standards for substantive rather than literal harmonization (Table 3).

Table 3: Sample Indicator Inconsistency Between Global Fund, PEPFAR, and UNAIDS

Global Fund Top 10 ³⁴	PEPFAR Essential Reported Indicators ³⁵	UNAIDS/MERG Core Indicators for Global AIDS Progress Reporting ³⁶
<ul style="list-style-type: none"> Number of TB patients who had an HIV test result recorded in the TB register 	<ul style="list-style-type: none"> Percent of HIV-positive patients who were screened for TB in HIV care or treatment settings Percent of HIV-positive patients in HIV care or treatment (pre-ART or ART) who started TB treatment 	<ul style="list-style-type: none"> Percentage of estimated HIV-positive incident TB cases that received treatment for both TB and HIV

Even if the Global Fund’s official guidance were clear and consistent, the Fund would still face substantial barriers to disciplined and systematic data collection due to inherent tension between portfolio-wide requirements and the Global Fund’s core principal of country ownership. In line with the Fund’s deference to country systems, it “does not impose its own indicators and targets” on recipient countries, but rather “uses those defined by countries themselves.”³⁷ This practice accounts for some of the fragmentation in the Global Fund’s internal reporting, as countries are free to propose indicators outside the scope of official guidance. A partial review of Global Fund grant scorecards conducted by Fan et al. (2013) reveals that many grant indicators are unrelated to the Global Fund’s M&E guidance (Table 4).

Table 4: Sample Indicators from Grant Scorecards that are Unrelated to Official M&E Guidance³⁸

HIV	Number of radio programs aired on the 10 radio-networks per year
HIV	Number of trainees using e- learning modules
TB	No. of advocacy materials produced and distributed (Flip Charts, Flash Charts, Bill Board, Cinema Slides, Leaflets, Pamphlets, TV, Radio Spots, Street Dramas, Folk Songs etc.)

³³ Appendix A, Indicators 7, 9, and 10.

³⁴ Global Fund (2011). Top ten indicators for routine Global Fund reporting. Accessed 3 October 2013 at http://www.theglobalfund.org/documents/monitoring_evaluation/ME_TopTenIndicators_Card_en/

³⁵ PEPFAR (2009). Next generation indicators reference guide. Accessed 3 October 2013 at <http://www.pepfar.gov/documents/organization/81097.pdf>

³⁶ UNAIDS (2011). Global AIDS response progress reporting 2012. Guidelines for construction of core indicators for monitoring the 2011 political declaration on HIV/AIDS. Accessed 3 October 2013 at http://www.unaids.org/en/media/unaids/contentassets/documents/document/2011/JC2215_Global_AIDS_Response_Progress_Reporting_en.pdf

³⁷ Global Fund (2012). Challenges. Performance based funding. Accessed 28 December 2012 at <http://www.theglobalfund.org/en/performancebasedfunding/challenges/>

³⁸ Fan V, Duran D, Silverman R, Glassman A (2013). Grant performance and payments at the Global Fund. CGD Policy Paper 031. Washington: Center for Global Development.

Low-Utility Data

As described above, data generated by Global Fund programs is expected to have many uses for many different constituencies. However, the current system of data collection and utilization is poorly equipped to leverage three key value for money practices: a) performance-based financing; b) costing and benchmarking; and c) impact-oriented measurement.

a. Performance Based Financing

The Global Fund implements a system of performance-based financing, wherein grant disbursement amounts are based on the Fund's assessment of grant performance. Theoretically, measured progress towards indicator targets is intended to inform a grant rating, which then determines how much money a grant receives during the next phase or disbursement period.

However, because the Global Fund does not implement performance-based funding exclusively on the basis of demonstrated results, the relationship between measured 'performance' and financing is ambiguous. As the Global Fund website explains, "financing is therefore not only (nor automatically) linked to the achievement of indicator results. Performance-based funding supports decisions based on a comprehensive and transparent assessment of program performance which also includes grant management, contextual challenges and, ultimately, impact."³⁹ LFAs are responsible for assessing PR performance and assigning a preliminary grant rating at every disbursement request, and during the grant renewal process. According to guidance, the LFA is tasked with assigning several different ratings, including for program results as compared to targets; financial management; program management; procurement and supply management; and an overall grant performance score that incorporates all of the aforementioned sub-ratings.⁴⁰ LFA grant ratings themselves are not final, but are subject to further adjustment by the Secretariat. The final process by which grant scores are assigned is not replicable or based upon clear, transparent criteria; rather they are subject to LFA and Secretariat discretion at every stage.

After a rating is assigned, the Secretariat issues a funding recommendation for the next disbursement period, or for Phase 2 if the grant is undergoing renewal. The recommended funding amount is theoretically based upon the corresponding 'indicative range' for each grant rating. However, because the grant ratings are not based exclusively on objective, measured progress against the most important indicators, there is only a loose relationship between performance against indicator targets and the grant ratings. Discretion by the Secretariat and LFAs creates the possibility of bias; for example, one 2007 study found that some LFAs were associated with systematically higher grant scores than others.⁴¹ The Secretariat and Grant Renewals Panel also holds enormous discretion over the performance-adjusted disbursement or commitment amounts, as the

³⁹ The Global Fund. Decision-making. Accessed 29 January 2013 at <http://www.theglobalfund.org/en/performancebasedfunding/decisionmaking/>

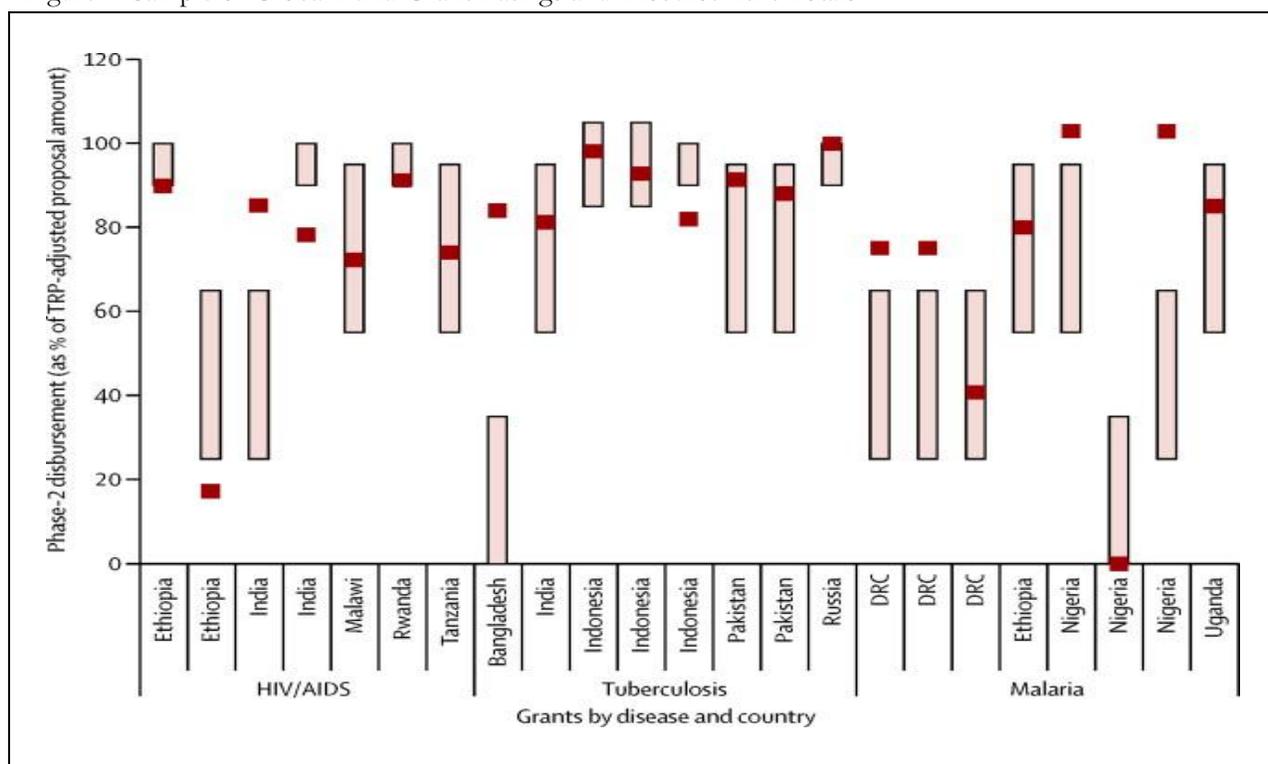
⁴⁰ The Global Fund (2011). LFA Manual. Section E. Accessed 4 October 2013 at http://www.theglobalfund.org/documents/lfa/LFA_Manual07SectionE_Manual_en/

⁴¹ Radelet S, Siddiqi B (2007). Global Fund grant programmes: an analysis of evaluation scores. *Lancet* 369: 1807-13.

indicative ranges are considered a “starting point” which does “not replace the judgment and active decision-making” of Global Fund bodies.⁴²

Such use of discretion dilutes the relationship between grant performance and final disbursement amount; in a recent analysis of Global Fund grants between 2003 and 2012, Fan et al. (2013) found no statistical association between grant ratings for Phase I and Phase II disbursement amounts.⁴³ Likewise, Secretariat and LFA discretion often leads to grant ratings and funding decisions well outside of the indicative ranges. Examining a subset of recent grants, Fan et al. further report that “at least 42% of [sampled] grants had final phase-2 amounts that were outside the expected indicative disbursement range derived from the grant rating, suggesting manual adjustment by Global Fund staff” (Figure 1).⁴⁴

Figure 1: Sample of Global Fund Grant Ratings and Disbursement Totals⁴⁵



Part of the observed discrepancy between grant scores and disbursements may be due to pressure from the Global Fund’s Board to ensure that Phase 2 commitments average around 75%, in addition to the Board-mandated 10% reduction on renewals.⁴⁶ Such a policy would provide a *de facto* incentive to normalize disbursements around 75%, which could further divorce financing decisions from measured performance.

⁴² The Global Fund (2012). The Global Fund operations policy manual. Accessed 4 October 2013 at http://www.theglobalfund.org/documents/core/manuals/Core_OperationalPolicy_Manual_en/

⁴³ Fan V, Duran D, Silverman R, Glassman A (2013). Performance-based financing at the Global Fund to Fight AIDS, Tuberculosis and Malaria: an analysis of grant ratings and funding, 2003-2012. *Lancet Global Health* 1(3): e161-68.

⁴⁴ Fan V, Duran D, Silverman R, Glassman A (2013). Performance-based financing at the Global Fund to Fight AIDS, Tuberculosis and Malaria: an analysis of grant ratings and funding, 2003-2012. *Lancet Global Health* 1(3): e161-68.

⁴⁵ Fan V, Duran D, Silverman R, Glassman A (2013). Performance-based financing at the Global Fund to Fight AIDS, Tuberculosis and Malaria: an analysis of grant ratings and funding, 2003-2012. *Lancet Global Health* 1(3): e161-68.

⁴⁶ Correspondence with Global Fund staff.

In sum, the broad discretion exercised by LFAs and the Secretariat suggestions that grant scores and financing decisions may not be clearly tied to measured results. Given this ambiguity, it is also unclear whether the intended performance incentive is fully transferred to the executor (i.e. the PR). Overall, the existing system of performance-based financing thus represents a missed opportunity to leverage value for money from measurement.

b. Costing and Benchmarking

Another important use for measurement is to relate outputs to money spent and inputs purchased in order to assess the cost-effectiveness of interventions. Inevitably, such an attempt to relate budgets to targets would be more indicative than precise because of cost variation between sites, patients, etc. At the very least, however, such data would help Fund Portfolio Managers (FPMs) and the Global Fund's Technical Review Panel (TRP) to understand whether budgets are reasonable given output targets. At most, a robust system for expenditure reporting and analysis could, in the long run, help to reveal areas for efficiency gains and the relative cost-effectiveness of interventions, particularly if expenditure data is made public for 'crowd-sourced' expert analysis.

To some extent, the Global Fund has exercised leadership in this area, particularly through its initiative to track unit costs for key commodity inputs – including antiretrovirals (ARVs), bed nets, and malaria treatment – via a price and quality reporting tool (PQR). PRs are required to report all relevant commodity purchases to the online PQR database, which is searchable and publicly accessible. The PQR has the potential to serve as an invaluable resource for procurement negotiation and costing; however, it is limited to a select number of commodity inputs and does not include other components of procurement (i.e. shipping, supply chain management, etc.) or service delivery. Further, the Global Fund has not consistently enforced PQR reporting requirements, and the database thus does not offer a complete or representative sample of purchase data.

Beyond commodity purchases, it is also not clear that the Global Fund currently has the capacity to relate service delivery outputs to unit costs. Table 5 provides an excerpt from a tuberculosis grant agreement in Bangladesh, which is among the most detailed information available in any grant agreement (many grant agreements contain no relevant budget information by service delivery area). To our knowledge, this is the only public place in which costs are related to interventions. According to the Global Fund's budgeting guidance, more detailed budgets may exist⁴⁷ – though they are not available online and their structure is unknown. Similar data may also exist for expenditures through the Global Fund's system for Enhanced Financial Reporting. However, the responses to our requests for data from the Secretariat suggest that such data is not currently in a usable form.

Table 5: Service Delivery Areas in Bangladesh⁴⁸

⁴⁷ The Global Fund (2011). Guidelines for budgeting in Global Fund grants.

⁴⁸ The Global Fund (2009). Program grant agreement. CCM Bangladesh Round 8 Tuberculosis Proposal with Health Systems Strengthening Cross-Cutting Interventions.

#	Macro-category	Objectives	Service Delivery Area	TOTAL Phase 1	%
1	TB Treatment	Pursue high quality DOTS expansion and enhancement	Improving diagnosis & Patient Support	17,536	1%
2	TB: Supportive Environment	Pursue high quality DOTS expansion and enhancement	Procurement and supply management	6,815	0%
3	TB: Supportive Environment	Pursue high quality DOTS expansion and enhancement	Monitoring, evaluation, supervision and	225,502	8%
4	TB: Supportive Environment	Pursue high quality DOTS expansion and enhancement	Human Resource Development	1,230,176	44%
5	TB/HIV Collaborative Activities	Address TB/HIV, MDR-TB and other challenges	TB/HIV	4,920	0%
6	TB Treatment	Address TB/HIV, MDR-TB and other challenges	MDR TB prevent and control and infection	135,766	5%
7	TB: Health Systems	Contributing to Health System Strengthening and Introduce	PAL	0	
8	TB Detection	Engage all care providers (PPM ISTC)	All care providers	157,804	6%
9	TB Detection	Empower people with TB and Community	ACSM (Advocacy, communication and social	382,864	14%
10	TB: Supportive Environment	Enable and promote Research	Operational Research	408,212	14%
11			Overhead	256,960	9%
			TOTAL*	2,826,555	100%

While Table 5 offers the most comprehensive publicly available service delivery cost data, it does not provide nearly enough detail to know if these interventions are sensible or cost effective. Most obviously, the totals are not related to itemized budgets (i.e. \$2,500 per course of second-line drugs * x patients, plus staff time, etc.). This is problematic because multiple funding sources contribute to a single program and grant performance is assessed in terms of outputs; accordingly, the only way to ensure reasonable cost levels is to demand precision about the translation of inputs into results, the portion of inputs covered by the Global Fund, and the cost of those inputs. Correspondence with Secretariat staff also suggests that excessively expensive cost components could historically be ‘hidden’ within the official budgets because their format was not amenable to the calculation of indicative unit costs. In addition, these tables represent only the planned budget; to our knowledge, no measurement of actual expenditures is done with service delivery outputs as the unit of analysis.

Within the current system, wherein the Global Fund directly reimburses cost components based on the expectation that they will translate into results, it is essential that costs are properly understood and assessed. However, if the Global Fund were to design a new system based on an accepted reimbursement rate for a specific output or outcome, such detailed documentation of cost components would not be necessary. That is, the Global Fund could either 1) work to better link costs with outputs, or 2) institute reimbursements that are independent of actual costs based on its ‘willingness to pay.’ These two options are not necessarily far apart – either way, budgets would need to include some unit cost per output (either a reimbursement rate or actual estimated cost) that are based on clear Global Fund guidelines.

c. Impact- and Coverage- Oriented Measurement

The Global Fund’s Top Ten Indicators (Appendix A) focus on inputs and outputs rather than outcomes and impact. In part this focus is understandable, as outcome and impact figures are difficult to assess through routine program monitoring. Nonetheless, the top ten indicator list is problematic for several reasons. First, several of the indicators have little value for assessing epidemiological progress, or even best practice in program implementation. The number of condoms distributed (indicator 7), for example, is meaningless without further knowledge about their correct and consistent use by high-risk or target groups. Likewise, indicators 8 and 10, which assess people benefiting from community-based programs and number of people trained, do not clearly relate to desired outcomes and impact. Even if the Global Fund limits its top 10 indicators to outputs, there are potential indicators, such as ART retention, which would relate more directly to epidemiological impact.

Second, all ten indicators are defined as numeric values without denominators, which limits their epidemiological utility. For example, indicator 3 measures the number of people receiving insecticide-treated nets (ITNs) and the number of households receiving treatment with indoor residual spraying (IRS), both of which are highly effective vector control strategies to prevent malaria transmission. However, the indicator does not require a denominator for the relevant targeted population, which is essential for understanding coverage and thus the potential for transmission disruption. The same holds true for several other indicators, such as ART provision (indicator 1), TB detection and treatment (indicator 2), and voluntary testing and counseling (indicator 5). Without denominators, these indicators fail to provide usable information on intervention coverage to inform epidemiological modeling and program planning.

Inability to Attribute Results

As discussed in Section 2, many Global Fund grants finance only one portion of the health production function. In any given grant, the Fund may finance one or more of the requisite inputs such as drugs and other commodities; training; supply chains; capital costs for new or improved labs and clinics; and the direct costs of service delivery such as health worker salaries. In such cases, the remainder of the cost components may be paid by recipient country governments, bilateral donors such as PEPFAR and the PMI, or other international organizations such as UNITAID and the World Bank.

Due to this comingling of funds, it is extremely difficult to discern the directly attributable impact of Global Fund financing. Indeed, one could argue that attribution itself is unnecessary, with grant performance best assessed via overall countrywide progress. Historically, the Global Fund has addressed this problem by reporting all results for which it “provides significant support to the national ARV program...[or] supports a more restricted project.” National ARV results financed by comingled funds could be reported if the Global Fund grant “supports an essential element of ARV treatment on a national scale” and “contribute[s] significant financial resources to the national effort,” defined as at least \$50 million to HIV programs over three years.⁴⁹

More recently, Global Fund representatives have described a shift away from ‘project-based aid toward ‘investment in the national program’ — suggesting that the Global Fund would be moving even further away from direct attribution of outputs and impacts. Even so, the Global Fund is committed to using its funds properly from an audit perspective, meaning that at the very least it will continue to require attribution of inputs. This speaks to the different purposes that attribution can serve, purposes closely related to the areas of measurement described above—particularly resource allocation, program management, and accountability.

Given the Global Fund’s current structure, its lack of clear attribution can be problematic for several core components of its business model. Without a clear relationship between Global Fund grants and observed results, the potential for PBF to incentivize and reward strong grant performance is questionable. For example, if a government received a Global Fund grant for ART, it could theoretically shift the cost of a single input (i.e. training) to Global Fund financing for people already on treatment. The PR could thus report that it had met its ART targets for the Global Fund grant without having started any additional people on treatment. If results are fungible between funders, one can imagine many such ways to game reporting

⁴⁹ The Global Fund (2011). Global Fund results fact sheet. Mid-2011. Other criteria include adequate grant performance and the absence of significant data quality issues.

requirements. For example, an Office of the Inspector General (OIG) audit report for a malaria grant in Madagascar found that “net results reported to the Global Fund included UNITAID LLINs (and yet the indicator results were tied to funding).”⁵⁰ There may be checks on such gamesmanship within the Secretariat, but it is not clear from publicly available documents that the implications of ambiguous attribution for PBF and grant management have been fully taken into account.

Further, the lack of attribution makes it near impossible for the Fund to learn whether a specific program component does and does not work—that is, whether a financed intervention is effective, and whether it should be eligible for future financing. Even if the overall national program is demonstrating strong epidemiological progress, it is still wasteful for the Fund to invest scarce resources in an ineffective component. Impact evaluation, as defined by the International Initiative for Impact Evaluation (3ie), thus seeks to isolate the causal effect of an intervention from overall trends and other confounding factors, to “measure the net change in outcomes for a particular group of people that can be attributed to a specific program.”⁵¹ This definition is thus quite distinct from that of the Global Fund’s Technical Evaluation Reference Group, which emphasizes “the importance of contribution and assessing causation and competing explanations rather than narrow attribution to one source of financing and single intervention”⁵² – and which is thus insufficient for measuring intervention efficacy.

‘Audit’ Focused Measurement

Most LFAs are accounting firms with core expertise in *audit*, not monitoring and evaluation. Stemming from 2008 changes in Global Fund guidance, all LFA teams must now include a monitoring and evaluation expert with at least seven years of relevant experience.⁵³ Nonetheless, their assigned scope of work and role within the Global Fund architecture does not support rigorous monitoring and evaluation, nor does it ensure thorough verification of reported outputs by PRs.

While the LFAs’ current scope of work includes a ‘data verification’ function, it is unrealistic to expect the LFAs’ ‘audit’ approach to ensure high data quality and complete accuracy. For each PR progress report, the LFA is tasked with verifying results through a desk review of data sources.⁵⁴ The exact nature of the verification is not specified, but appears to focus on the consistency of topline results figures with the underlying documentation from facilities, program managers, etc. In addition, LFAs typically conduct an on-site data verification (OSDV) once annually per PR and for each disease type. But the OSDVs, conducted by a single LFA staff member over 6-12 days, are too limited in scope to detect and deter all measurement problems. According to Global Fund guidance, the OSDV need only consider three indicators and eight sites to reach conclusions about overall accuracy; even when discrepancies are identified, any overstatement by <30% is considered a “minor data quality issue.” In addition, the three-step OSDV procedure is itself quite superficial, including a “bottom-up audit trail comparing recorded events in primary records to aggregated reports; cross-verifications of programmatic data with other sources of information, i.e. ARV patient records

⁵⁰ The Global Fund (2011). Audit of Global Fund grants to Population Services International Madagascar. Audit & investigation reports issued by the Global Fund’s Office of the Inspector General on 1 November 2011.

⁵¹ International Initiative for Impact Evaluation (2008). Founding document for establishing the International Initiative for Impact Evaluation. Accessed 4 October 2013 at <http://www.3ieimpact.org/media/filer/2012/05/17/3iefoundingdocument30june2008.pdf>

⁵² The Global Fund (2012). TERG work plan 2013-2014. 5th Strategy, Investment, and Impact Committee meeting.

⁵³ The Global Fund (2011). LFA manual. Accessed 5 October 2013 at http://www.theglobalfund.org/documents/lfa/LFA_ManualComplete_Manual_en/

⁵⁴ The Global Fund (2011). Guidelines for completing the PR “ongoing progress update and disbursement request”, and LFA “ongoing progress review and disbursement recommendation.”

and pharmacy reports; and spot-checks of actual delivery of services and/or commodities with beneficiaries.”⁵⁵

A strengthened verification system for outputs and coverage measures might be sufficient for some proven interventions, such as ARV treatment/retention or ITNs, because there is established evidence of a clear relationship between the intervention and health impact. But for many interventions supported by the Global Fund, effectiveness is unclear or untested, or may vary dramatically depending on the details of program implementation (i.e. behavior change and condom distribution for HIV prevention, generalized AIDS education, and experimental delivery strategies). For these interventions, it is insufficient to simply verify that people were reached, because there is no established relationship between those outputs and broader health impact.

For example, consider a condom distribution and education program targeted toward commercial sex workers. By reviewing invoices, LFAs could easily assess how many condoms were purchased by the PR and at what price; whether trainings were held with community outreach workers; and whether the total number of condoms reported as distributed is consistent with the number distributed by each outreach worker. Such methods would help to detect fiduciary irregularities and identify obvious delays or problems with program implementation. But even with M&E-specific expertise on staff, the LFA as an institution is not appropriate to assess measures of program quality such as outcomes, coverage, and impact. For example, were the condoms received by the target group (an output)? Were they used properly, or at all (an outcome)? What percentage of commercial sex workers used a condom regularly and appropriately with their last clients (coverage)? And did condom distribution result in lower HIV incidence among commercial sex workers (impact)? As indicators get closer to outcomes and impact, the measurement tools needed – i.e. surveys, impact evaluation, biomarkers, surveillance sites, etc. – are well beyond the LFAs’ core competencies in audit and document verification.

These problems do not stem from poor LFA performance, but from the extensive and unrealistic demands placed upon them. Over time, more and more tasks have been outsourced to LFAs, such that they now have essential core responsibilities at every stage of the grant cycle.⁵⁶ Illustrative of their proliferating roles and responsibilities, the Global Fund now lists 70+ guidelines and tools for LFA use, including a 256 page manual.⁵⁷ Rather than adding even more M&E responsibilities, it may be wise to consider paring back the LFA mandate and creating a new institution type with core expertise in measurement and evaluation.

Perverse Incentives

Because performance-based financing is intended to reward the achievement of input and output targets based on self-reported results, PRs may face strong incentives to increase their input and output numbers – possibly at the expense of program impact and country data systems. This challenge is not unique to the Global Fund. Lim and others (2008) found that the GAVI Alliance’s results-based immunization services support program (currently being phased out) caused countries, on average, to inflate their official immunization statistics—an effect neither prevented nor predicted by the GAVI Alliance’s use of Data

⁵⁵ The Global Fund (2011). LFA guidelines for On-site Data Verification (OSDV) and Rapid Services Quality Assessment (RSQA) implementation.

⁵⁶ Global Fund. Local Fund Agents. Accessed 5 October 2013 at <http://www.theglobalfund.org/en/lfa/>

⁵⁷ The Global Fund (2011). LFA manual. Accessed 5 October 2013 at http://www.theglobalfund.org/documents/lfa/LFA_ManualComplete_Manual_en/

Quality Audits.^{58,59} Similarly, the production of more health services does not necessarily translate into better health. For example, the ‘fee-for-service’ model common in the U.S. health care system incentivizes providers to perform unnecessary but costly procedures with little to no health benefit (and possibly net harm).⁶⁰

Among the Top 10 indicators, several examples clearly demonstrate how perverse incentives to maximize inputs and outputs could undermine program effectiveness. Indicator 7 asks grantees to report on the number of condoms distributed, but not to whom they were distributed, or whether they were used effectively or at all. This could lead to ineffective targeting to easily reachable low-risk groups, or distribution to groups who do not understand the purpose of condoms and thus never use the product. In either case, the incentive to distribute a maximal number of condoms would run in opposition to the most effective strategies, for example targeted condom distribution and social marketing to high-risk, hard to reach groups such as commercial sex workers and men who have sex with men (MSM). Likewise, Indicator 10 asks for the number of people trained with grant funds, but not whether the training had any impact on clinical practice, quality of care, or other measures, or even whether the health workers underwent standard training programs with known effectiveness. This could incentivize PRs to hold unnecessary training sessions; to offer general training to many people versus specialized training targeted toward the most relevant health workers; or to have the same people undergo many duplicative trainings. Such distortions would be wasteful but natural consequences of a measurement strategy where accountability is primarily defined as meeting specific input or output targets.

Because of these dynamics, self-reported data should be treated with caution and robustly verified to manage and mitigate perverse incentives. But as described above, LFAs lack the mandate, resources, and staff capacity to ensure representative, credible, and rigorous verification of results reported by recipients. The OSDV and Rapid Services Quality Assessment (RSQA) provide yearly on-the-ground spot checks of program performance. Given their limited scope and frequency, however, they are also unable to offer a representative sample for all but the smallest programs, even if sites are selected through random sampling (as recommended in Global Fund guidance, though not commonly implemented). Further, selected sites are notified of the LFA visit a week before, giving time to prepare data sources.⁶¹ In contrast, Rwanda’s highly successful PBF scheme also uses an audit approach, but auditors verify results at all facilities once each quarter.⁶² If the incentive to inflate results is not fully contained by the existing LFA checks, overstated results could infiltrate both formal reporting to the Global Fund and the countries’ own routine monitoring systems, leading to poor data quality and a degradation of local monitoring and surveillance capacity.

⁵⁸ Lim SS, Stein DB, Charrow A, Murray CJL (2008). Tracking progress toward universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage. *Lancet* 372(9655): 2031-46.

⁵⁹ According to personal communications with GAVI Alliance staff, GAVI Alliance was aware of the likelihood of discrepancies between administrative and survey data at the time the immunization services support program was launched. However, the initial program design was borne from a conscious decision to endow countries with responsibility for measurement in line with principals of country ownership and health system strengthening, and as an effort to avoid creating new parallel systems.

⁶⁰ See, for example, Abelson R, Creswell J (2012). Hospital chain inquiry cited unnecessary cardiac work. *New York Times*, 6 August 2012. Accessed 4 October 2013 at http://www.nytimes.com/2012/08/07/business/hospital-chain-internal-reports-found-dubious-cardiac-work.html?hp&_r=1&

⁶¹ The Global Fund (2011). LFA guidelines for On-site Data Verification (OSDV) and Rapid Services Quality Assessment (RSQA) implementation.

⁶² Basinga P, Gertler PJ, Binagwaho A, et al. (2010). Paying primary health care centers for performance in Rwanda. Policy Research Working Paper 5190, World Bank.

Data Inaccessibility

The Global Fund ranks among the most transparent donor agencies,⁶³ particularly with respect to its financial data and grant portfolio. The website offers a document database⁶⁴ of grant proposals, agreements, performance reports, and scorecards (among others), though some such documents are missing or incomplete. The Fund also offers detailed excel-format data for grant- and disbursement-level variables such as disease area, implementer, country, dates, performance ratings, and disbursement amounts. A few important document types are *not* publicly shared, particularly grant budgets, work plans, and LFA reports.

Despite the availability of core program documents, however, program *data* remains largely inaccessible and difficult to interpret. Specific grant service delivery areas and activities can only be identified by sifting through 40-page PDF documents; in many cases, such information is not available at all.⁶⁵ Likewise, results are typically found within publicly posted grant scorecards and within the annual portfolio-wide results report, which in 2012 provided cumulative aggregates (2002-2012) for 17 indicators – despite the fact that indicators are not universally standardized, and thus are unlikely to produce comparable data. The most recent report disaggregates the cumulative results by region, but does not provide data at the annual or country-level units of observation.⁶⁶ For the layperson, including watchdogs, external experts, and civil society/beneficiaries in developing countries, it has historically been quite challenging to understand the activities and performance of Global Fund programs. This inaccessibility has impeded public accountability and oversight, not just for the Global Fund itself but also for its implementing partners, including country governments.

In recent months, the Global Fund has taken several positive steps towards more open and accessible data. The Fund's recently revamped web portal offers a user-friendly, interactive interface for browsing basic grant information for individual grants (grant scores, disbursements, etc.), but cannot be easily exported or aggregated across grants or countries. Most recently, the Global Fund has released some of its internal data through a Web application programming interface (API),⁶⁷ though manipulation and interpretation of the data requires significant expertise and effort. The Global Fund should continue its movement toward greater transparency and data accessibility over the coming years to further enhance accountability and coordination.

4. Competing Visions of Measurement

To simplify, we can frame two competing visions of measurement: 1) hierarchical reporting, and 2) population-based measurement. Historically, most donors (including the Global Fund) have used a hierarchical reporting system (Figure 2). Under the Global Fund's current system of hierarchical reporting, program and financial reports flow upwards from facilities to sub-recipients; from sub-recipients to principal recipients, from principal recipients to LFAs; from LFAs to the country portfolio managers; and ultimately into portfolio-wide results figures.

⁶³ Publish What You Fund (2012). 2012 aid transparency index. Accessed 5 October 2013 at <http://www.publishwhatyoufund.org/index/2012-index/>

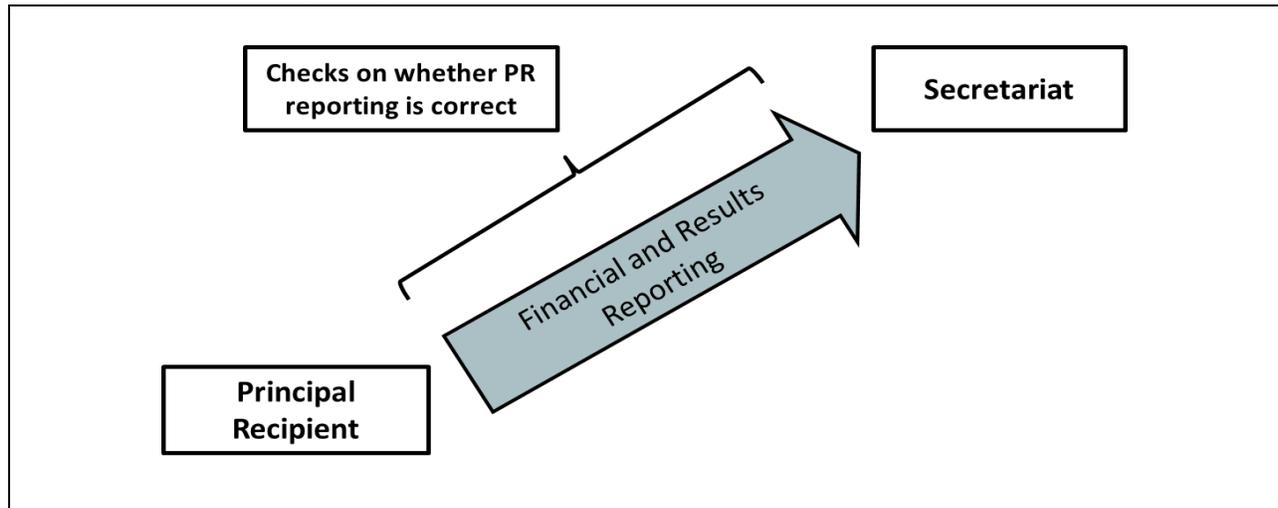
⁶⁴ Global Fund (2012). Document search. Accessed 30 December 2012 at <http://portfolio.theglobalfund.org/en/Search/DocumentSearch#>

⁶⁵ PGAs are not universally posted online. Posted documents are often missing relevant and important information. Citation: work in progress by Fan V, Silverman R, Glassman A.

⁶⁶ Global Fund (2012). Strategic investments for impact. Global Fund results report 2012.

⁶⁷ See the Global Fund (2013). Welcome to the Global Fund data site. Accessed 5 October 2013 at <http://web-api.theglobalfund.org/Home/Index>

Figure 2: Hierarchical Reporting



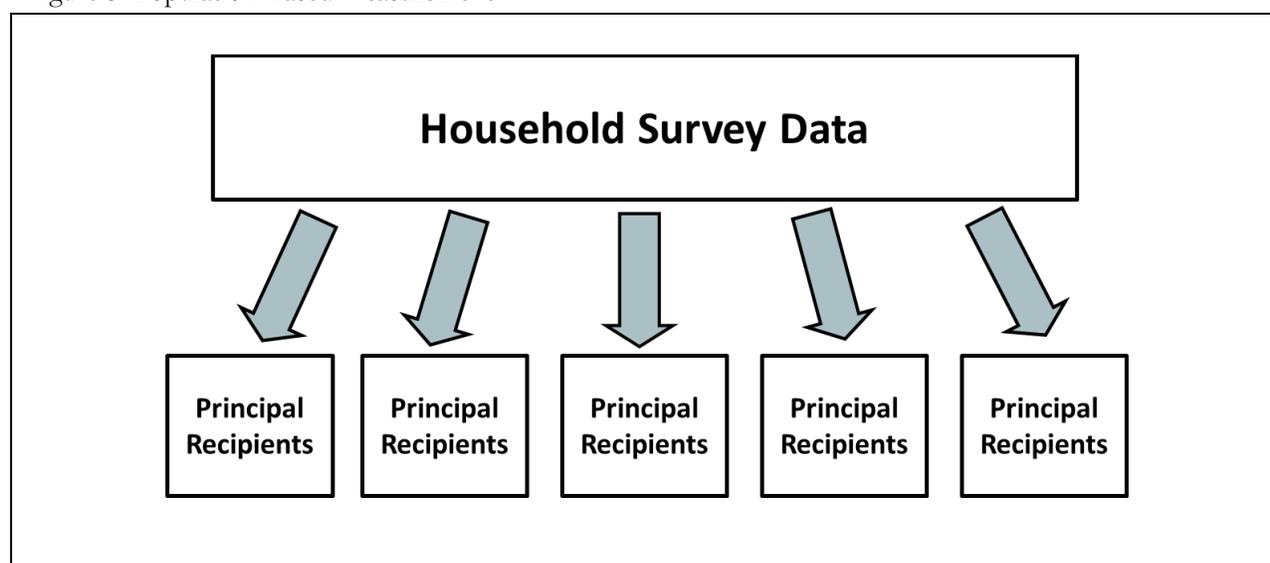
Within a hierarchical reporting system, results are ultimately generated by PRs themselves, albeit with external spot checks and documentation reviews for completeness and accuracy. Accountability is achieved through an audit strategy to verify if funds were spent correctly and PRs truly completed the reported activities. The hierarchical approach offers several advantages, including strong fiduciary accountability and the ability, at least theoretically, to assess unit costs through financial reporting. Some also argue that hierarchical reporting promotes country ownership, as the PR takes principal responsibility for implementing and reporting on the program's measurement strategy, ideally through a national HMIS system.

These advantages, however, are countered by important limitations. Because hierarchical reporting relies primarily on the routine monitoring data generated by the programs themselves (i.e. receipts, facility records, etc.), it is subject to the principal-agent problem described in Section 3, wherein PRs are responsible for reporting their results but hold a vested interest in reporting the 'right' results to maintain funding. Under such an incentive scheme, self-reported results are easy to exaggerate or manipulate, though LFA spot checks may provide a partial deterrent. Hierarchical reporting can also be hampered by the limited M&E capacity of many PRs, particularly country governments. Hierarchical reporting is also expensive – substantial costs are borne by the PR itself in building and maintaining M&E systems, and by the Global Fund in verifying and auditing PR reports. Finally, hierarchical reporting does not provide information on program impact, as its scope is typically limited to the input and output indicators collected via routine program monitoring.

Under population-based measurement (Figure 3), the data used to evaluate programs is collected from the populations served by a program rather than the program itself. Population-based measurement is independent of the programs being measured – data may be collected via household surveys, impact evaluation, vital registration records, etc. Accountability is ensured because the donor knows whether the funded program achieved its ultimate objectives. Cash on Delivery (COD) Aid, for example, represents a 'pure' population-based measurement approach, wherein funding is tied explicitly and exclusively to independently verified results.⁶⁸

⁶⁸ Birdsall N, Savedoff WD. Cash on delivery: a new approach to foreign aid. Washington: Center for Global Development.

Figure 3: Population-Based Measurement

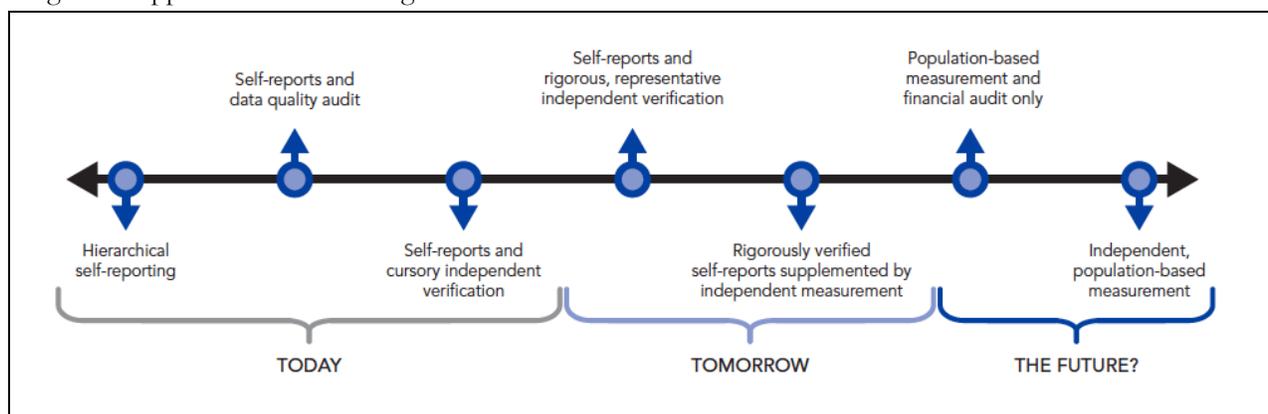


Population-based measurement avoids many of the deficiencies inherent to hierarchical reporting. This approach can detect impact, i.e. whether Global Fund programs are achieving their desired objectives such as high intervention coverage and reduced disease prevalence. Independent measurement also eliminates the principal-agent problem; PRs cannot ‘game’ the results without actually having the desired impact on overall epidemiological trends. In addition, the data generated by independent household surveys (while potentially expensive) is a global public good; beyond its utility to the donor agency, it is also invaluable to international organizations such as UNAIDS and the WHO; to other donor agencies such as PEPFAR and the PMI; and for country governments.

However, because population-based measurement is divorced from the programs themselves, attribution of results to a specific donor or grant is extremely difficult. Likewise, it may not be possible to clearly link changes in epidemiological trends (for better or ill) to any particular factor or strategy. For these reasons, and because there will be a significant lag between the occurrence of implementation problems and the detection of poor program impact, a population-based measurement approach may offer limited utility for proactive program management. This approach also cannot ensure fiduciary accountability, as a program may achieve strong impact despite significant fraud or misuse of funds.

While we have thus far framed hierarchical reporting and population-based measurement as diametrically opposed, the two approaches are really two extreme ends on a spectrum (Figure 4). Few funding agencies adopt either extreme approach. Rather, they choose among hybrid approaches in the middle of the spectrum, where self-reported results are subject to increasingly rigorous verification and supplemented by population-based measurement to assess the coverage, outputs, and impacts of supported programs.

Figure 4: Approaches to Assessing Grant Performance



The Global Fund’s current measurement sits toward the left end of the spectrum, with principal recipients’ self-reports supplemented by cursory independent checks for accuracy and data quality. Without making radical changes, the Global Fund could greatly strengthen data by supplementing self-reporting with rigorous, representative, and independent data collection. The benefits of robust, independently verified data are fourfold:

- First, as a recent World Bank report notes, “the very existence of the verification process is a key improvement in the governance of the health system” through its ability to both promote health system accountability and encourage national dialogue on health service results.⁶⁹
- Second, independent data sources and rigorous verification improve the quality of administrative data, critical to promoting sustainable M&E systems in recipient countries and improving in-country program management. Even the best-performing countries will gain if they can regularly test their administrative reporting systems against independent robust and reliable data. For the lowest capacity countries, such independent verification may be the only way to have accurate data until the substantial time and investment in reporting systems begins to pay off. Indeed, when programs financed by the performance-based Health Results Innovations Trust Fund implemented verification in participating facilities (at times alongside penalties for over-reporting), the World Bank observed a clear and rapid jump in the accuracy of self-reported data on quantity of services delivered.⁷⁰ In Cameroon, for example, independent verification helped significantly reduce over-reporting of outpatient consultations.⁷¹ Still, there remains much to learn about the optimal strategy for measuring and verifying service quality.
- Third, robust performance verification is critical to informed program management by the Secretariat. Without on-the-ground staff who can regularly interact with beneficiaries and observe program implementation, independent data are crucial for ensuring that the Secretariat has an accurate assessment of the returns to its financial investments. In turn accurate data assure that performance-based payments reward real improvements, rather than administrative reporting errors or intentional manipulation. For this reason alone verification of programmatic data deserves substantial investment by the Secretariat, likely equaling or exceeding the amounts spent on LFAs.

⁶⁹ Health Results Innovation Trust Fund (2013). Information to improve value for money in health. Africa Health Forum: Finance and Capacity for Results.

⁷⁰ Cashin C, Vergeer P (2013). Verification in results-based financing (RBF): The case of the United Kingdom. Washington: World Bank.

⁷¹ The World Bank (2013). Using results-based financing to achieve maternal and child health – progress report 2013. Washington: World Bank.

- Fourth, high-quality data are global public goods that can be coordinated with other stakeholders and inform the work of national governments, donors, and independent researchers. To improve accountability around the Family Planning 2020 commitments, the Bill and Melinda Gates Foundation and others will support direct data collection in 69 countries, including baselines and annual follow-ups to estimate modern contraceptive users. Such large data collection efforts merit joint support, and connections with funders like the Global Fund.

5. Recommendations

Please see the full report of the Working Group.⁷²

⁷² Glassman A, Fan V, Over M, et al. (2013). More health for the money: putting incentives to work for the Global Fund and its partners. Washington: Center for Global Development.

Appendix A: Top Ten Indicators for Routine Global Fund Reporting⁷³

	Disease	Indicators for routine Global Fund reporting
1	HIV	Number of adults and children with advanced HIV infection currently receiving antiretroviral therapy
2	TB	Number of (a) new smear-positive TB patients detected, (b) new smear-positive TB patients who were successfully treated and (c) laboratory-confirmed MDR-TB patients enrolled in second-line anti-TB treatment
3	Malaria	Number of (a) insecticide-treated nets or re-treatment kits distributed to people and (b) households (or structures or walls) in designated target areas sprayed by indoor residual spraying in the past 12 months
4	Malaria	Number of people with fever receiving antimalarial treatment according to national policy (specify artemisinin-based combination therapy versus other therapy)
5	HIV	Number of women and men aged 15-49 years who received an HIV test in the last 12 months and who know their results
6	HIV	Number of HIV-positive pregnant women who received antiretrovirals to reduce the risk of mother-to-child transmission
7	HIV	Number of condoms distributed
8	HIV, TB and malaria	Number of people benefiting from community-based programs: specify (a) care and support including orphan support, home-based management of malaria and directly observed therapy (DOT); (b) behavior change communication outreach activities including specific target groups; and (c) disease prevention for people most at risk (except behavior change communication)
9	HIV/TB	Number of TB patients who had an HIV test result recorded in the TB register
10	Strengthening health systems for HIV, TB and malaria	Number of people trained

⁷³ The Global Fund. Top Ten Indicators for Routine Global Fund Reporting. Available at http://www.theglobalfund.org/documents/monitoring_evaluation/ME_TopTenIndicators_Card_en/