

Evans School Policy Analysis and Research (EPAR)

Evaluating Country-Level Government M&E Systems
EPAR Request No. 299

Pierre Biscaye, Melissa LaFayette, Andrew Martin,
Monica Richardson, Zoë True

C. Leigh Anderson

*Prepared for the Development Policy and Finance Team
of the Bill & Melinda Gates Foundation
April 22, 2015*

Professor C. Leigh Anderson, Principal Investigator

Abstract

This report is intended to support the Development Policy and Finance (DPAF) team's work on improving the results information that informs aid allocations. Aid results information is often not comparable, since monitoring and evaluation frameworks, information gathering processes, and definitions of "results" differ across donors and governments. This report reviews approaches to results monitoring and evaluation used by governments in developing countries, and highlights trends and gaps in national monitoring and evaluation (M&E) systems. We collect evidence on 42 separate government M&E systems in 23 developing countries, including 17 general national M&E systems and 25 sector-specific national M&E systems, with 14 focused on HIV/AIDS, 8 on health, and 3 on agriculture. The evidence review includes external case studies and evaluations of M&E systems, government M&E assessments, M&E plans, strategic plans with an M&E component, and multi-country reviews of M&E, accountability, and aid effectiveness. We evaluate harmonization of government and development partner M&E systems, coordination and institutionalization of government M&E, challenges in data collection and monitoring, and analysis and use of results information. We also report on key characteristics of M&E systems in different sectors. The report is based on an accompanying spreadsheet that contains the coded information from the 42 M&E systems.

1. Introduction

Bilateral and multilateral aid organizations are interested in understanding whether the funds they provide to recipient countries are used efficiently and achieve desired outcomes. Indicators of "effectiveness" of aid funding can include quantities of outputs or services provided as a result of aid funds, changes in outcomes that are related to those outputs or services, and measures of the cost of providing outputs or achieving certain level of impact on target outcomes.

To measure effectiveness and efficiency in achieving desired outcomes, donor organizations rely on monitoring and evaluation (M&E) systems for collecting and analyzing data on costs, activities,

Highlights of Findings on Government M&E Systems

- Governments face several challenges with institutionalizing and coordinating M&E systems, including defining and clarifying roles and leadership, aligning and coordinating across sectors, and building internal staff capacity.
- In many countries, strong demand from elected officials is supporting improved coordination of M&E.
- Data collection challenges include inadequate staffing, high staff turnover, infrequent training for data collection skills, duplication of efforts, delays in data collection and submission, and limited data verification.
- Many systems do not report rules or standards for data collection, aggregation, or verification. An increasing number of systems, however, are using electronic tools and systems to improve data collection and are sharing publicly available databases.
- Almost all systems have strategic frameworks, often expressed as a theoretical causal chain outlining activities, outputs, and outcomes, but there is a greater focus on tracking outputs of programs than evaluating their outcomes or impacts.
- Few systems consistently use M&E data for decision-making around strategy, budgeting, or program management.
- Efforts to align donor and government M&E systems include the use of common indicators, technical support from donors, public dissemination of M&E data, and systems for mutual accountability.
- Harmonization between donors and governments is limited by donors' ongoing use of parallel implementation and reporting systems, but the number of these separate systems is falling in many countries.

EPAR's innovative student-faculty team model is the first University of Washington partnership to provide rigorous, applied research and analysis to the Bill and Melinda Gates Foundation. Established in 2008, the EPAR model has since been emulated by other UW Schools and programs to further support the foundation and enhance student learning.

NOTE: The findings and conclusions contained within this material are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

outputs, and outcomes. These M&E systems may be set up specifically by the donors to monitor the programs they support, or the systems may be overseen by the recipients of their funding. Holvoet, Gildemyn, & Inberg (2012) suggest one reason for donors' hesitancy to use recipient government M&E systems as concerns about the strength and reliability of these systems, due to differing data collection methods and varying degrees of information quality. Donors often set up parallel M&E systems in an attempt to ensure measurement of the specific indicators they are concerned with, but these parallel systems can hinder development of local capacity for M&E and increase reporting burdens (Holvoet, Gildemyn, & Inberg, 2012; OECD, 2011).

Agreements such as the Paris Declaration on Aid Effectiveness in 2005 and the subsequent Accra Agenda for Action in 2008 tie governments and aid organizations together through commitments to improve aid effectiveness. These commitments are intended to create better harmonization between aid donors and recipients by aligning aid to recipient country priorities, increasing use of country-led monitoring and evaluation (M&E) systems to measure results, and improving mutual accountability. A central tenet of these aid agreements is that improving aid results management does not lie solely in donor organizations, but must be based on recipient-led demand for results information (Porter & Goldman, 2013) and ownership of M&E systems by recipient governments (Segone, 2010). Donor organizations are encouraged to provide technical support to recipient governments and to work closely with them to develop M&E systems that reliably collect and analyze data on aid effectiveness (OECD, 2011).

This report addresses the need for better understanding of current government M&E systems, including how these systems are organized and how results information is gathered, processed, shared, and ultimately used. We review case studies and government documentation describing 42 country-level M&E systems in 23 developing countries, including 17 general national M&E systems, 14 systems focused on HIV/AIDS, 8 on health, and 3 on agriculture. Section 2 of this report describes our literature search process and results, and Section 3 outlines our review methodology. In Section 4, we divide our key findings into four sections, each examining an overarching question that evaluates the various components and complexities of government M&E systems:

1. M&E Institutionalization and Coordination: What is the level of institutionalization and coordination of government M&E systems?
2. Data Collection: How well-established are the data collection and monitoring systems of government M&E systems?
3. Analysis and Use of Results Information: To what extent are governments analyzing and using the monitoring and results information generated by their M&E systems?
4. Harmonization: What is the level of alignment and harmonization between government and development partner M&E systems?

In each section, we draw on relevant questions from our review framework spreadsheet, which accompanies this report¹. We summarize general trends across all M&E systems, as well as highlight good practices and recurring challenges. In Section 5 we evaluate general characteristics and key trends specific to each type of M&E system we review: general national systems, health systems, HIV/AIDS systems, and agriculture systems.

2. Literature Search Process and Results

Our search aimed to identify national government M&E systems in developing countries, such as those tracking progress on Poverty Reduction Strategies or in the HIV/AIDS sector. We conducted one general Google search for government results monitoring & evaluation (M&E) systems². This search captured multiple types of documents referring to government systems designed to monitor and evaluate results in domains relevant to international aid flows. We also performed individual Google searches for specific countries' government M&E systems³, targeting the following Foundation focus geographies: Bangladesh, Burkina Faso, Ethiopia, Ghana, India - Bihar, India - Odisha, Kenya, Malawi, Mali, Nepal, Nigeria, Rwanda, Tanzania, and Uganda.

We elected to use Google rather than academic databases for our searches with the understanding that much of the literature we were looking for would be grey, or unpublished, including reports and evaluations of governments, donors, and non-profit organizations. We used Google results as a starting point for additional searches, looking for relevant documents on the sites that came up in the results lists, especially when these were government sites. We also conducted

¹ The questions in the review framework are listed in Appendix A.

² We used the following search string: ("monitoring and evaluation" OR "performance management" OR "performance monitoring" OR "results management") AND national AND (system OR framework) AND ("economic development" OR "international development" OR aid) AND "case study" We chose to append "case study" to the search string as the most relevant results from our initial exploratory searches were framed as case studies. However, we did not find many "case studies" of government M&E systems, so the documents we retrieved include a variety of reports and evaluations of these systems, often commissioned by national governments or donor organizations.

³ To conduct these searches, we found that using an abbreviated search string, ("monitoring and evaluation" OR "M&E") AND national AND (system OR framework) AND "Country Name", yielded more relevant results than the longer, more general search string.

supplemental searches of the Scopus and PAIS International databases to ensure that we captured prominent academic literature.

We screened the titles and summaries of the results for all of our searches and retrieved all documents that met the following screening criteria:

- Describes an M&E system that is administered by a national government or ministry;
- Describes either a national M&E system or an overall health/agricultural sector M&E system;
- Describes framework OR monitoring/evaluation results for a specific M&E system; and
- Does not describe general principles or guidelines for M&E that are not applied to a specific M&E system.

In some cases evaluations or reports of government M&E systems were not available and we elected to include national M&E frameworks and strategic plans in our body of evidence. These documents provide insight into demand for M&E, challenges, and planned M&E activities, although they could not be treated as evidence that these activities were ever carried out. As the scope of this request is country-level, we excluded documents describing sub-national M&E systems (e.g. region-, program-, and project-level M&E systems) and M&E systems administered by non-government development partners. We included documents describing national-level M&E systems that were specific to the health or agricultural sector, as some countries lacking comprehensive national M&E systems have been supported by development partners to establish systems in these sectors. Finally, we also included documents describing country M&E systems not on our list of target countries, and documents that evaluated multiple countries.

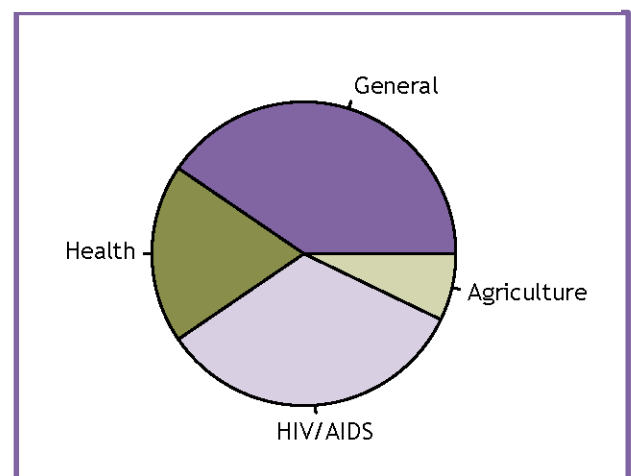
Appendix A includes further detail outlining our search process, including a table breaking down the number of results reviewed and documents retrieved for each of the searches we conducted, a map illustrating the countries and types of systems included in this review, and a timeline illustrating the year of publication of the documents we review.

Our searches yielded 108 relevant documents. These documents describe 42 different government M&E systems in 23 different countries. Of these 42 national systems, 17 are general national M&E systems that aim to monitor and evaluate progress across many sectors. The remaining 25 systems are national in scope but sector-specific, including eight health-sector M&E systems (of which one focuses on WASH), 14 separate HIV/AIDS M&E systems, and three agriculture M&E systems.

For each of these 42 government M&E systems, we collect and review all of the documents describing or evaluating the system. These documents include case studies and evaluations of M&E implementation, internal M&E assessments, specific M&E plans, general strategic plans with an M&E component, and multi-country reviews of M&E, accountability, and aid effectiveness. The multi-country documents include:

- A 2014 OECD report on making development cooperation more effective;
- A 2013 case study on growing demand for M&E in Africa;
- A 2012 review of M&E arrangements related to Poverty Reduction Strategy Papers (PRSPs);
- A 2012 United Nations Economic and Social Council review of mutual accountability for development results;
- A 2012 review of health aid effectiveness;
- Three OECD reports (most recently from 2011) on progress in implementing the Paris Declaration for aid effectiveness;
- A 2009 UNICEF review of capacities for country M&E systems;
- World Bank reviews of results-based national development strategies (2007), HIV/AIDS M&E in Africa (2007), institutionalizing M&E in Latin America and the Caribbean (2006), and institutions for monitoring poverty reduction strategies (2006); and
- Two evaluations of HIV/AIDS M&E in sub-Saharan Africa (Peersman, et al., 2009; Ekouevi, Karcher, & Coffie, 2011).

Figure 1. Types of M&E systems reviewed



Appendix B includes a table summarizing the body of evidence for each of the 42 government M&E systems we reviewed. We rated the strength of the body of evidence for each M&E system based on the relevance and quality of the information included in the documents. A “low” rating indicates that fewer than two documents are available and that these provide

limited detail on recent M&E planning or implementation. A “medium” rating indicates that two to six documents are available, but there are gaps in the information provided. These gaps include lacking either a specific M&E plan or specific M&E evaluation. A “high” rating indicates that at least three documents are available that also provide a high level of relevant and recent information on M&E planning and implementation. Of the 42 M&E systems we review, 24 were given a “medium” or higher strength of evidence (including all but one of the general national M&E systems) and 18 were rated “medium-low” or lower. Of these 18 weaker bodies of evidence, all but 1 (India) describe a sector-specific M&E system.

3. Review Methodology

Prior to reviewing the selected documents, we developed a review framework and set of questions to categorize information about different components of government M&E systems based on criteria commonly used to evaluate or categorize M&E systems in the case studies and evaluations we retrieved⁴. The majority of the questions are either dichotomous (Yes/Not specified) or rate M&E systems on a scale from low to high, to allow for comparative analysis. We organized the questions (93 in total, listed in Appendix C) into the following categories:

- Basic Descriptive Questions (9 questions)
- Demand for M&E (11 questions)
- Institutional Design (18 questions)
- Strategic Framework (13 questions)
- Budgeting (5 questions)
- Monitoring and Data Collection (17 questions)
- Analysis (11 questions)
- Reporting and Use of M&E Data (9 questions)

To answer the review framework questions for each government M&E system, we reviewed all of the documents describing aspects of that system in order to aggregate relevant information and provide an accurate portrait of the system⁵. After reviewing all the documents, we summarized the information for each question into the review framework spreadsheet⁶.

In the spreadsheet, the majority of questions are coded as either “Yes,” “Somewhat,” “Planned,” or “Not Specified.” “Yes” means there is clear evidence in the documents reviewed. “Somewhat” means the evidence indicates there are some limitations to the M&E system. “Planned” means that planning documents address the question but there is no evidence from M&E reports or evaluations. Finally, “Not specified” means that there is no information in the documents reviewed that addresses the question.⁷ In some cases, we coded “Mixed evidence,” which indicates that there is contradicting evidence either within individual documents or across documents, and it was not possible to determine that one source was likely to be more valid.

For certain questions, we rated M&E systems on a scale from low to high, but the ratings criteria are specific to the individual questions. We describe the criteria used at the same time as we present our findings for these questions. The only questions that do not have categorical responses are those asking which offices are responsible for different M&E activities, but for most of these questions the type of office provides some level of categorization.

The review framework spreadsheet is included as a separate deliverable for this project. Appendix D includes tables listing the relevant questions from the review framework for each section of our findings, and notes which questions are specifically addressed in the report, which is intended to highlight the main findings captured in our review framework. As our findings are often based on several documents across multiple systems, we generally name the systems the findings apply to and include the source information in the spreadsheet, rather than citing each document individually in this report. However, we do include specific citations when presenting examples of findings from individual documents. All references consulted in our research are included at the end of this report.

⁴ Our review framework questions are contained in an excel spreadsheet to aggregate the information from the 42 government M&E systems. The spreadsheet also includes short qualitative descriptions justifying the coding decisions for each of the review questions.

⁵ For documents specifically describing the particular M&E system, we read the entire text, while for documents reporting on multiple different M&E systems, we used the search function to locate descriptions of the M&E system we were reviewing.

⁶ The information from the different documents for each country is also available in a single notesheet, organized by question.

⁷ The proportion of M&E systems coded as “Not specified” will indicate gaps in our body of evidence, where we lack information to be able to describe different aspects of M&E systems. We did not code “No,” as it is possible that our literature search did not identify all documents describing a particular M&E system, and many of the documents we reviewed are several years old, meaning there may have been changes to the M&E systems in the last few years.

4. Findings Across M&E Systems

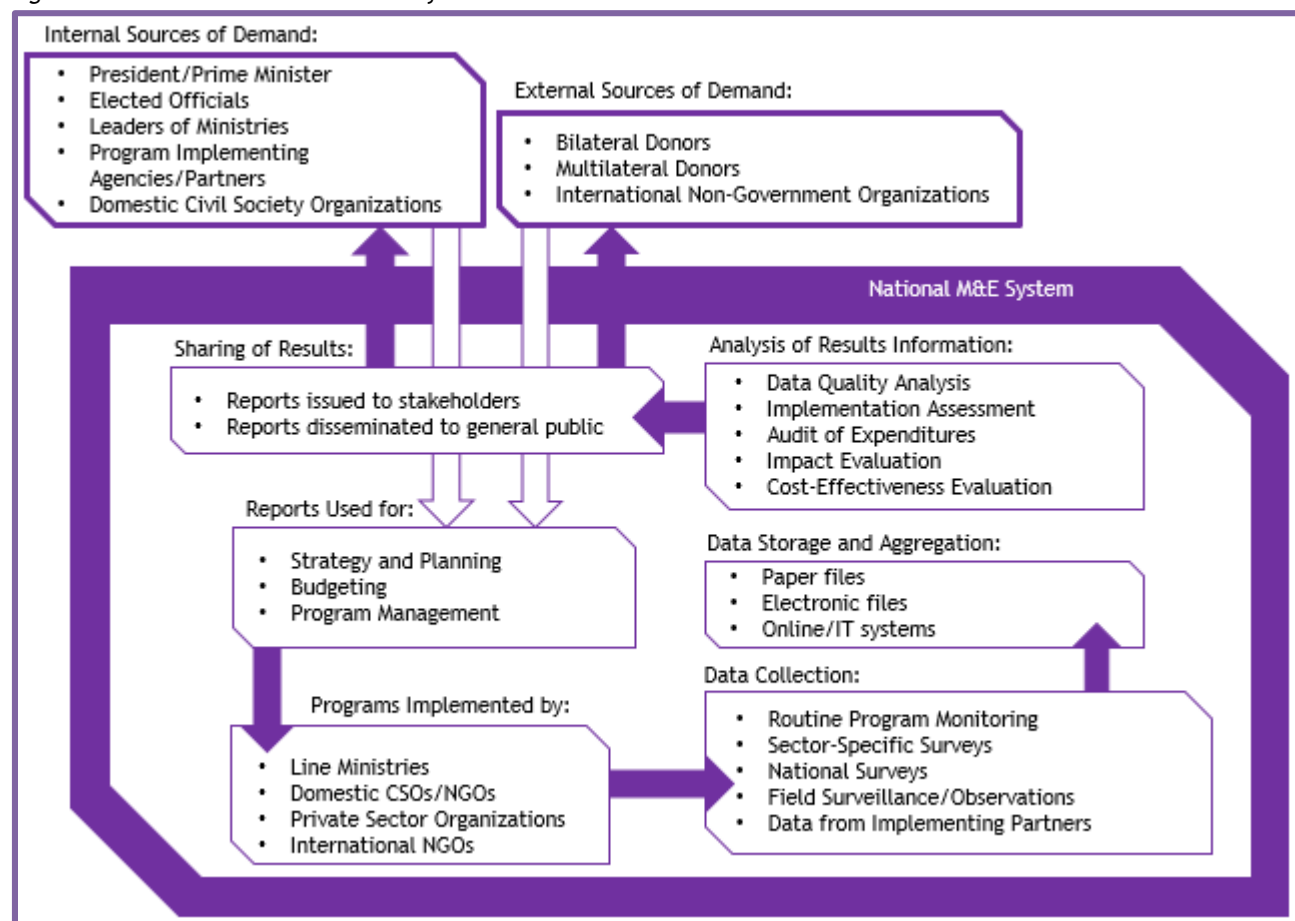
We evaluate the evidence on government M&E systems through four overarching questions that cross over the eight categories of M&E system components coded in our review framework:

1. M&E Institutionalization and Coordination: What is the level of institutionalization and coordination of government M&E systems?
2. Data Collection: How well-established are the data collection and monitoring systems of government M&E systems?
3. Analysis and Use of Results Information: To what extent are governments analyzing and using the monitoring and results information generated by their M&E systems?
4. Harmonization: What is the level of alignment and harmonization between government and development partner M&E systems?

For each of these questions, we analyze the relevant evidence from our review framework and summarize the general trends across all of the 42 M&E systems reviewed. We also highlight recurring challenges and examples of best practices for each overarching question. Some distinctions among the different types of M&E systems (general vs. sectoral) are identified, though the particular characteristics of these systems are covered more thoroughly in section 5 of this report.

The general flow of information in the M&E systems we review is illustrated in *Figure 2*. Section 4.1 of this report covers internal sources of demand and how national M&E systems are organized in terms of institutional roles and harmonization of M&E activities across government offices and levels of government. In section 4.2 we examine processes for data collection, storage, aggregation, and verification. Section 4.3 reviews government systems for analyzing M&E data and for using results information. Section 4.4 covers harmonization of internal and external sources of demand for M&E data and processes for alignment, technical support, and mutual accountability.

Figure 2. Flow of Information in M&E Systems



4.1 M&E Institutionalization and Coordination

Key Findings:

- Elected officials, and especially Presidents or Prime Ministers, appear to be drivers of demand for national M&E systems.
- Internal demand¹ for M&E information is rated as “high” or “medium” in 25 systems. 13 systems receive a “high” rating for demonstrating high-level political commitment, engaging a wide range of stakeholders in planning processes, and focusing on reviewing results and performance to inform planning.
- 19 systems have “high” clarity of M&E roles and responsibilities, indicating that documents specifically define which office is responsible for the various M&E activities and reports indicate that these offices do carry out their responsibilities. 13 have a “medium” rating, and nine have a “low” rating, indicating limited evidence of clear M&E roles.
- Harmonization and coordination of M&E activities are either rated as “high” or “medium” in 20 systems. 11 systems have “low” levels of alignment of M&E activities, indicating undefined roles and responsibilities, weak coordination, and/or overlap or confusion in implementing M&E activities across poorly-aligned M&E systems. Another 11 systems do not have enough evidence to assess the harmonization of their M&E activities.
- The main challenges to institutionalizing and coordinating M&E systems include defining and clarifying roles and leadership within M&E systems, aligning and coordinating M&E activities across sectors, and building capacity for M&E activities within recipient governments.

The ability of an M&E system to support decision-making depends on how well the system is institutionalized and coordinated. According to Holvoet, Gildemyn, & Inberg (2012), government M&E systems require appropriate institutional structures for coordination, support, central oversight, and feedback that incorporate different stakeholders. In this section we analyze the level of institutionalization and coordination of government M&E systems. We consider internal demand for M&E, institutional design of M&E systems, and clarity and harmonization of M&E roles and responsibilities. *Figure 2* is a heat map illustrating our findings for several key elements of M&E institutionalization and coordination, subdivided by the different sectors of government M&E systems.

Figure 3. Institutionalization and Coordination of Government M&E Systems



4.1.1 Internal Demand for M&E

Our rating for the level of internal demand⁸ for M&E, column 1 in *Figure 3*, is based on evidence of the number of government offices demanding M&E information, the nature and extent of their demand for M&E, and the involvement of civil society in demanding M&E information. The evidence of internal demand is primarily from case studies and evaluations of government M&E systems reporting on internal M&E demand. We also consider evidence from strategic plans and M&E-specific plans outlining the processes for developing M&E systems, which often include descriptions of the government offices responsible for promoting M&E. Systems with little evidence of government demand or that only describe plans to increase demand are rated “low,” those that demonstrate consistent demand for M&E from more than one government office are rated “medium,” and those that demonstrate a strong culture of government demand for M&E are rated “high.”

In our ranking, M&E systems with “high” internal demand are characterized by a wide range of stakeholders engaged in strategic planning processes (e.g. Nepal, Nepal HIV/AIDS, Nepal Health, Nigeria HIV/AIDS, Ethiopia Health), a focus on reviewing results and performance to inform planning (e.g. Nepal HIV/AIDS, Mexico, Chile), and high-level political commitment, such as from the president or prime minister (e.g. Mexico, South Africa, Ethiopia, Chile). Systems with low levels of demand for M&E sometimes describe an intent to follow these guidelines, and several have plans to invoke internal demand, but there is limited evidence that they are carried out. In some of these cases (e.g. Malawi HIV/AIDS, Kenya Health), it appears that the main driver for M&E is compliance with international agreements and reporting requirements, rather than internal demand for M&E.

For three of the systems in this report, the evidence we reviewed does not describe sources of internal demand. Eleven of the 39 systems whose documentation does discuss internal demand for M&E demonstrate low levels of internal demand. In Benin, internal demand is low but the government is “introducing a range of tools to increase commitment by Cabinet, the president or prime minister and sector departments” (Porter & Goldman, 2013). A 2011 review finds that the Rwanda Health system downsized its M&E task force and delayed finalizing the health M&E policy (Holvoet & Inberg, 2010). 12 systems are rated “medium,” including Rwanda where M&E demand is shifting from donor-led to government-led (Murray-Zmijewski & Gasana, 2010). For example, in Kenya the 2010 Constitution includes accountability requirements promoting M&E (CLEAR, 2012), and in South Africa HIV/AIDS demand has been pushed by a presidential HIV/AIDS campaign (Kawonga, Blaauw, & Fonn, 2012). 13 systems demonstrate high levels of internal demand, such as Chile where there is a 20 year history of implementing M&E (Guzman, Irarrazaval, & de los Rios, 2014) and Nepal where stakeholders from all relevant ministries and from civil society are involved in developing M&E guidelines and in supporting monitoring activities (National Planning Commission of Nepal, 2013). The remaining three systems that discuss internal demand (Mali, Malawi, Ghana) have mixed evidence, with different sources indicating different levels of internal demand.

For most general national M&E systems (nine of 12 systems where documents discuss internal demand sources), the office of the president or prime minister is described as a main driver of internal demand, sometimes (five systems) with support from certain ministries. For HIV/AIDS M&E systems demand generally comes from multi-sectoral agencies established to monitor progress in achieving HIV/AIDS outcomes (five of eight systems where documents discuss internal demand sources). Demand for M&E information in Health and Agriculture M&E systems is usually driven by the ministries responsible for these areas. In 15 systems, stakeholders from civil society, such as academic institutions and local non-profit organizations, contribute to internal demand for M&E information (column 2 in *Figure 3*). Three systems describe plans for increasing civil society demand. In some cases however (e.g. Mali, Uganda, Malawi Agriculture), M&E systems appear to be implemented mainly as a response to donor reporting requirements. The review spreadsheet includes additional detail on internal and external sources of demand for M&E.

Overall, 21 of the 42 systems we review describe challenges for increasing internal demand for M&E. Challenges described in several M&E systems include:

- limited internal capacity for responding to increased demand for M&E activities (e.g. Belize HIV/AIDS, Ethiopia, Ethiopia Health, Kenya HIV/AIDS, Malawi, Malawi Agriculture),
- managing the pace and sequencing of M&E reforms to overcome internal resistance (e.g. Mexico, Chile),
- separate systems for monitoring budgets and outcomes that make it difficult to incentivize use of M&E (e.g. India, Ghana),
- fear that M&E information will jeopardize funding (e.g. Rwanda Health), and

⁸ We define “internal demand” as demand from stakeholders inside the country, including government officials and civil society, as opposed to “external demand” from international development partners.

- lack of funding and resources dedicated to M&E and reliance on donor support (e.g. Ethiopia, Ethiopia HIV/AIDS, Ethiopia Health, Uganda, Uganda HIV/AIDS, Malawi, Malawi Agriculture, Malawi HIV/AIDS, Kenya, Kenya HIV/AIDS, Belize HIV/AIDS, Colombia).

Six systems (Rwanda, Colombia, Uganda, Tanzania HIV/AIDS, Kenya HIV/AIDS, Chile) report increasing communication with and involvement of stakeholders as a method and recommendation for increasing internal demand for and use of M&E information. These systems describe a need to promote awareness of monitoring information and evaluation findings and making this information more accessible to government offices at different levels, civil society organizations, international development partners, and the general public. Some reports describe increased bottom up participation of civil society as a key means of promoting accountability and sustained demand for M&E. The expectation is that as more stakeholders are involved in an M&E system, they will increasingly demand results information from the government officials that are accountable to them, thereby increasing those government officials' demand.

However, there are limitations to externally communicating government results information. We evaluated whether data collected by M&E systems were open and public, meaning that the raw monitoring information collected is available and accessible beyond a summary in distributed reports. Twelve systems have evidence that data are open and public⁹. These systems generally publish M&E data on government websites or accessible databases. Eight systems describe plans for sharing their results information openly and publically, but it is not clear from the documentation whether they have begun to do so.

Other methods for increasing demand for M&E described in the systems are encouraging ministers to play a leading role in championing the usefulness of the M&E information (e.g. Chile) and adoption of results-based management and incentives tied to performance (e.g. Benin). In addition, 23 of the M&E systems involve plans or ongoing activities to increase internal capacity for M&E, in order to be able to respond to greater demand for M&E.

4.1.2 Clarity of Institutional M&E Roles

Another component of M&E systems is the clarity of institutional roles (column 3 in *Figure 3*). In most cases, M&E roles and responsibilities are described in M&E-specific plans, and we use evidence from M&E reports and evaluations to assess whether the roles are carried out as planned. For the systems with no M&E-specific plan available, our assessments are based on M&E reports or strategic plans that include an M&E component. We rate clarity of institutional roles as “high” if the systems specifically define which office is responsible for the various M&E activities, and reports indicate that these offices do in fact carry out their responsibilities. A “medium” rating indicates that the systems specify which offices are responsible for a limited number of M&E roles, but do not delegate responsibilities for all M&E activities. A “low” rating indicates that the systems do not specify which offices are responsible for M&E activities, or are characterized by confusion and overlap over roles and responsibilities.

Nearly half of the systems we reviewed (19) are coded as having “high” clarity, with well-defined roles and responsibilities. For 14 of these systems, we were able to retrieve an M&E-specific plan describing institutional roles, although for all 19 there was additional evidence to demonstrate that the roles and responsibilities described in the plans had been implemented. We find that 13 systems have “medium” clarity of M&E roles, with some gaps in the planning or allocation of M&E responsibilities. Four of these systems have M&E-specific plans that we reviewed, but these plans do not sufficiently clarify certain M&E roles. For example, for Nepal Health, the Public Health Administration Monitoring and Evaluation Division (PHAMED) appears to be responsible for most M&E activities but this is not stated anywhere in their M&E plan (Nepal Ministry of Health and Population, 2012). In Malawi, a 2006 World Bank report finds that although there is a plan for a coordinated M&E system, it “has not been implemented, and many stakeholders are not aware of their roles” (Bedi, et al., 2006). Finally, we find that nine systems have “low” institutional clarity and lack defined M&E roles. For example, for Ethiopia Health the M&E system appears to be coordinated by the Federal Ministry of Health but it is not clear which offices are responsible for specific M&E activities. In Mali, a 2006 World Bank report finds that “There are too many actors in a highly dispersed, loosely coordinated system, and it is not entirely clear where the overall leadership is located” (Bedi, et al., 2006). In South Africa, a 2012 review finds “overlapping mandates and unclear boundaries for M&E functions among government departments” (Goldman, et al., 2012).

Many M&E systems do not clearly specify which office is responsible for particular activities. In our review, we recorded which government offices are responsible for different M&E activities, including overall M&E coordination, overseeing

⁹ Links to the online databases for these twelve systems are included in the coding spreadsheet that accompanies this report.

monitoring, budget monitoring, evaluation, reporting, and planning (including strategy, activities, and budgets). While 37 of the 42 M&E systems describe who is responsible for coordinating M&E activities, 37 describe who is primarily responsible for overseeing monitoring activities, and 32 describe who is primarily responsible for national-level reporting, fewer systems specify who is responsible for budget monitoring (24 systems), evaluation (26), or planning (23).

To analyze the diversity of government offices involved in M&E activities, we categorize the offices responsible for different activities into four main groups: Ministries (where no further detail is given on who within a ministry was responsible), Ministry Sub-Committees (specific national-level offices within ministries), Multi-Sectoral Agencies (independent agencies with representatives from different government offices and other stakeholders, sometimes including civil society organizations), and other Independent Agencies (independent government offices without a clear multi-sectoral element).

For the majority of HIV/AIDS M&E systems (eight of 14), a single multi-sectoral agency, usually a form of National AIDS Committee, is responsible for coordinating M&E and for most M&E activities. Only one other system (Nepal Health) features a similar multi-sectoral agency. Among the nine systems with this type of institutionalization of M&E roles, seven have a “high” level of institutional clarity while the other two have a “medium” level. Similarly, Mexico’s National Council for the Evaluation of Social Policy (CONEVAL) is an independent agency responsible for coordinating M&E activities, and the system also has a high rating for clarity of roles. Four M&E systems assign M&E coordination and responsibility for most M&E activities to a particular ministry sub-committee. In Chile it is the Budget Directorate in the Ministry of Finance (DIPRES); in Colombia it is the Directorate for Evaluation of Public Policy (DEPP) in the Department of National Planning; in Rwanda Health it is the Planning, Policies, and Capacity Building Unit in the Ministry of Health; and in Uganda Health it is the Quality Assurance Department in the Ministry of Health. Of these systems, two have “high” clarity of M&E roles and responsibilities and two have “medium” clarity. The remaining M&E systems (mostly general national systems) either have one ministry or a variety of different offices responsible for M&E activities, but there is no clear association with the clarity of M&E roles and responsibilities, which may be “high,” “medium,” or “low.”

In 25 M&E systems, the documentation we reviewed describes particular roles for national statistics offices or departments. The main role of statistics offices is in supporting data collection. In 13 M&E systems the statistics office is specifically responsible for conducting surveys and coordinating official government statistics, and in a further three M&E systems the statistics office provides other data collection and management support. In six M&E systems, the statistics office provides coordination support to the office that manages the M&E system, generally through membership in a working group or multi-sectoral agency. In three M&E systems, the statistics office plays a supporting role in data analysis and evaluation. Thus, while national statistics offices are generally not the main office responsible for any M&E activities, they often have a role in supporting government M&E systems.

4.1.3 Harmonization of Government M&E Activities

To further evaluate the coordination of government M&E systems, we analyze the level of harmonization of their M&E activities (column 4 in *Figure 3*). Drawing on Porter & Goldman (2013), the Centre for Learning on Evaluation and Results (CLEAR, 2012), and Holvoet, Gildemyn, & Inberg (2012), we consider harmonization of M&E activities according to the following criteria:

- whether there is overlap or confusion in implementation of M&E activities,
- whether the office responsible for coordinating M&E is active in managing the different M&E activities, and
- whether there is overlap among different uncoordinated M&E systems.

Systems with clearly assigned *and implemented* M&E roles that are coordinated under a unified government M&E system are rated “high.” A unified M&E system indicates that data collection and reporting activities of different users of M&E information are aligned, so implementing offices are not repeating the same M&E activities for different government systems. Systems that demonstrated some overlap or confusion in implementing M&E activities but that are generally coordinated and implemented as planned are rated “medium.” Systems where unclear roles and responsibilities, weak coordination, and/or confusion among responsibilities for multiple separate government M&E systems lead to significant overlap or confusion in implementing M&E activities are rated “low.” For 11 systems, the documents we reviewed do not provide enough information to assess the harmonization of their M&E activities.

Six systems demonstrate a “high” level of harmonization in M&E, with clearly coordinated and aligned activities supporting a unified M&E system. Of these six systems, five are HIV/AIDS M&E systems where multi-sectoral government agencies have clear responsibilities for coordinating M&E activities. The only exception is Mexico’s general national M&E system, where

M&E activities are managed by an independent agency, CONEVAL. All of these systems also have an institutional framework for combining and coordinating M&E functions, usually operationalized in an M&E plan placing that mandate on the coordinating agency. 23 of the 42 systems we review have such frameworks, and another seven describe plans for implementing one.

14 systems demonstrate a “medium” level of harmonization. In Kenya, M&E roles are clearly assigned, but monitoring data is required to feed into multiple different sources with different requirements, leading to multiplication of effort (Porter & Goldman, 2013). In Colombia, a 2007 review finds that the activities of different offices are not well-integrated (Independent Evaluation Group). In Uganda, a 2012 study finds a well-defined organizational structure for M&E but a need to converge existing M&E systems “to reduce, harmonize and minimize duplication and resultant waste of human and financial resources” (CLEAR). 11 of these 14 systems either have or plan to implement an institutional framework for M&E.

11 systems have a “low” level of M&E harmonization. For example, for Kenya Health a 2010 USAID review reports that “coordination and supervision within and between levels are weak” (Blumhagen, et al., 2010). In Malawi, a 2013 review finds that “Linkages among various Ministries with respect to the exchange of information are weak or non-existent” (CLEAR-AA, 2013). The Kenya HIV/AIDS system has a clear institutional framework with one multi-sectoral body responsible for M&E coordination, but M&E activities are decentralized to such an extent that they are almost impossible to manage, leading to issues with parallel M&E systems, inadequate data quality assurance systems, gaps in data collection, and delays in reporting (Karani, Bichanga, & Kamau, 2014; Government of Kenya, 2011). The South African general national M&E system is characterized by a variety of M&E mandates across departments resulting in overlapping and uncoordinated functions and a disconnect between M&E and planning activities. These issues may be related to the absence of a legal basis for much of this system and consolidation across the multiple actors although this system has a “Government Wide Monitoring and Evaluation (GWM&E) Policy Framework” (Goldman, et al., 2012). Of these 11 systems, five have institutional frameworks for M&E and four have plans to implement one.

The systems we review describe many challenges with M&E coordination and harmonization. The most common challenge, mentioned in 18 systems, is a lack of clarity or mandate for the office in charge of coordination. In many cases (e.g. South Africa, Malawi Agriculture, Mali, Kenya), the office charged with M&E coordination does not have a firm legal basis, while in others (e.g. Benin, Malawi, Rwanda Health, Kenya Health, Chile, India), there is confusion over how to coordinate M&E activities across multiple sectors and levels of government. Eight systems also describe lack of capacity for M&E coordination, both in terms of number of staff and their training and skills with respect to M&E activities. Further, four (Senegal, Kenya, India, Ethiopia Health) describe lack of resources for M&E coordination, with this limited funding contributing to capacity issues and reducing the coordinating office’s ability to manage M&E activities.

Two HIV/AIDS M&E systems present potential approaches for overcoming these challenges. The Nepal HIV/AIDS system has a multi-sectoral agency with several staff responsible for M&E coordination, regional officers tasked with coordinating M&E functions, clear plans for building regional M&E capacity, and provisions for coordinating monitoring activities with regional health offices (Nepal Ministry of Health and Population, 2013). The Bangladesh HIV/AIDS system emphasizes building M&E capacity both within the central M&E coordination unit and within implementing partners (Bangladesh NASP, 2007).

4.2 Data Collection

Key Findings:

- 37 out of 42 systems have a list of indicators based on a strategic framework to guide data collection.
- We find significant gaps in the evidence for government data collection and monitoring processes: 17 systems do not specify rules or standards for data collection, 18 do not specify rules for aggregation of data, and 22 do not specify a process for data verification.
- Data collection challenges include inadequate staffing, high staff turnover, infrequent training for data collection skills, duplication of efforts, and delays in data collection and submission.
- Evidence of data verification is limited. 10 systems demonstrate operational processes for data verification, but there are gaps in their implementation. Nine additional systems specify plans for implementing data verification processes.
- 15 systems demonstrate use of an IT system or tools to support M&E. 14 other systems specify plans for establishing IT M&E systems. The nature of these IT systems varies, from simple excel-based forms to more complex data collection and aggregation databases. 12 systems have a public online database of M&E data.

Data collection and monitoring capacities appear to be a challenge in the development of M&E systems, with only a few systems demonstrating advanced capacities. In this section we consider the current trends and challenges for data collection and monitoring in M&E systems, looking at indicator selection, data sources and standards, data management, verification, and aggregation, coordination of data collection, and capacity for monitoring. *Figure 4* sorts the M&E systems we review according to their ratings on the different data collection and monitoring questions.

Figure 4. Data Collection and Monitoring

Documented rules for data collection	Documented rules for data aggregation	Process for data verification	Functioning IT system		
				Ethiopia	National M&E Systems
				Mexico	
				Ghana	
				Rwanda	
				Senegal	
				Tanzania	
				Burkina Faso	
				Benin	
				Chile	
				India	
				Mali	
				Nepal	
				Malawi	
				South Africa	
				Colombia	
				Uganda	
				Kenya	
				Malawi	HIV/AIDS M&E Systems
				Kenya	
				Nepal	
				Bangladesh	
				South Africa	
				Vietnam	
				Belize	
				Ethiopia	
				Tanzania	
				Uganda	
				Ghana	Health M&E Systems
				Rwanda	
				Nigeria	
				Cote d'Ivoire	
				Malawi WASH	
				Malawi	Health M&E Systems
				Ethiopia	
				Kenya	
				Uganda	
				Rwanda	
				Nepal	Agriculture M&E Systems
				Mozambique	
				Rwanda	
				Malawi	Agriculture M&E Systems
				Ghana	

☐ Not specified / Mixed evidence
☐ Planned
☐ Somewhat
☐ Yes

4.2.1 Selection of Indicators

A list of indicators serves to clarify what an M&E system is intended to monitor and evaluate (Holvoet, Gildemyn, & Inberg, 2012). A large majority (37 of 42) of the systems we review have a list of indicators based on a strategic framework and linked to either a general or sector-specific national strategy. For the five systems where we do not find a list of indicators specified, the strength of the body of evidence for those systems is “low” or “medium-low” (with the exception of Benin, for which the seven documents we review do not include an M&E plan).

On-the-ground data collection may not always be aligned, however, with the official list of indicators. Only 14 of the systems specified that routine monitoring indicators were aligned with the official strategic indicators. For the Ethiopia Health system, a 2007 evaluation noted that monitoring and reporting was not yet based on the official strategic indicators (Ethiopian Ministry of Health, 2007), and more recent documents do not clarify whether reporting based on official strategic indicators has begun. It is difficult to determine if our finding of a low number of systems specifying alignment of routine and strategic indicators is driven by real implementation difficulties, or a lack of reporting on this issue in our evidence base. 32 of 42 systems collect data for their indicators in part through routine monitoring processes of implementing agencies, which indicates at least some alignment of routine monitoring and national indicators.

Methods for selecting indicators were diverse across the systems we reviewed. For some systems, indicators are selected during the collaborative process of updating the overall strategic plan (e.g. Rwanda HIV/AIDS, Ghana), or through other participatory processes. In other systems, government agencies specializing in M&E (e.g. Mexico’s CONEVAL, South Africa’s Department of Performance Monitoring and Evaluation) are solely responsible for selecting indicators. There are no evident differences (including the alignment of routine monitoring indicators with the official list) between systems that selected indicators through a collaborative processes and those that used a specialized agency.

We do not review indicator lists for different M&E systems within a given country to assess whether the same indicators are used in sectoral and general national systems. Further, while most indicators are drawn from national strategic frameworks, we do not assess whether the sector-specific strategic frameworks are aligned with the general multi-sector frameworks, as the number of indicators can be in the hundreds for each system. However, countries with more harmonized M&E activities may be more likely to align indicators across systems, as a challenge for countries with uncoordinated M&E activities is overlapping and inconsistent data collection across separated M&E systems.

4.2.2 Basic Capacities: Data Collection and Aggregation

Following agreement on what indicators to collect data on, the most basic capacities necessary for a functioning monitoring system are data collection and data aggregation. To evaluate data collection capacity, we first examine whether M&E systems have documented rules for data collection. Data collection rules include what data to collect, who collects data, and the frequency of data collection. As displayed in column 1 of *Figure 4*, 19 systems have documented rules for data collection, four have some rules in place but not consistently applied, two have plans for implementing rules, and 17 do not specify any data collection rules.

M&E systems with documented data collection rules include the Rwanda HIV/AIDS system, where separate rules for data collection for health care facilities and for community-based organizations are established, and the Malawi HIV/AIDS system, which specifies acceptable data sources in addition to specifying responsibility for and frequency of data collection. Data collection rules are often laid out in tables where the methods and responsibilities for data collection are outlined for each indicator. This approach is common in the HIV/AIDS M&E systems we review (e.g. Belize HIV/AIDS, Ethiopia HIV/AIDS, Nigeria HIV/AIDS, Nepal HIV/AIDS, Bangladesh HIV/AIDS). Issues with systems that “somewhat” have documented data collection rules include local organizations using their own data collection forms instead of official ones (e.g. South Africa HIV/AIDS), or official standards that are too vague to guide the actual practice of data collection by line ministries (e.g. Ethiopia).

We also evaluate data collection capacity through the extent of local data collection coverage. We rate local data collection coverage from high to low, with a “high” rating indicating consistent coverage in all areas, and “low” indicating significant gaps in data collection for certain areas of the country, populations, or sectors of government. We only have enough evidence to evaluate the extent of local data collection coverage in 14 systems, with four systems rated “high”, six rated “medium”, and four rated “low.” All four of the systems with high local data collection coverage are HIV/AIDS systems. For systems with “low” or “medium” data collection coverage, data collection is limited for rural areas or areas

with poor infrastructure (e.g. Ethiopia Health, Uganda Health, Belize HIV/AIDS), for non-government implementing agencies (e.g. Malawi Agriculture), and for ministries with weaker data collection capacities (e.g. Ethiopia, Malawi).

Groups responsible for data collection also vary across M&E systems. Line ministries are frequently responsible for data collection (15 systems), often in collaboration with central government agencies. In the systems we review, implementing agencies on the ground, such as health clinics and other service providers, are tasked with collecting relevant information, but district- or region-level line ministry or M&E offices are responsible for overseeing these data collection activities. Some systems (13) also incorporate data collected by NGOs or civil society organizations, while others (11) plan to.

Data from local sources must be aggregated to measure indicators at the regional or national level. We evaluate whether M&E systems include documented rules for data aggregation, including tools for recording data and standardized reporting forms. Compared to data collection, we found fewer (14) systems with specified rules for data aggregation, and a greater number (six) of systems that planned to create rules. A common approach is developing specific forms for reporting data (e.g. Nepal, Nepal HIV/AIDS, Nigeria HIV/AIDS, Rwanda Agriculture, Malawi WASH, Malawi HIV/AIDS, Bangladesh HIV/AIDS). The majority of systems with documented rules for data aggregation (10 of 14) are HIV/AIDS systems. Included are the Nepal HIV/AIDS system, where data are aggregated monthly at a central agency, and the Tanzania HIV/AIDS system, which includes a single national database. Even otherwise well-developed general national systems (e.g. Mexico, South Africa) have noted difficulties in data aggregation due to variations in data formats at regional and local levels (Castro, et al., 2009; Engela & Ajam, 2010)). We did not find any evidence of rules for data aggregation for 18 of the systems.

Several evaluations of M&E systems mention challenges with increasing the capacity for monitoring, especially challenges with staffing. Some systems have high staff turnover (e.g. Uganda Health, Ethiopia), and others lack funding to hire enough qualified workers (e.g. Malawi Health, Bangladesh HIV/AIDS, Uganda HIV/AIDS, Tanzania). For other systems, staff training is infrequent due to cost constraints, especially for peripheral workers at the lowest levels of the M&E system (e.g. Nigeria HIV/AIDS, Ethiopia Health). Inadequate staff capacity can result in duplication of efforts (e.g. Nigeria HIV/AIDS) and a low percentage of data being turned in on time (e.g. Cote d'Ivoire HIV/AIDS).

4.2.3 Advanced Capacities: Data Verification and Functioning IT Systems

Beyond data collection and aggregation standards are more sophisticated monitoring capacities, including established processes for verifying data quality and integrating data collection, aggregation, and verification in an IT M&E system (Ethiopian Central Statistics Agency, 2011). Data verification processes include periodic data audits and supervisory visits from centralized agencies to local data collection centers as well as comparison of similar measures from different sources. Most systems (22 of 42) do not describe any data verification processes, and only four, all of which are HIV/AIDS systems, include clear rules for the process of data verification. The Nigeria HIV/AIDS system, for example, uses a tool developed by the Global Fund at least biannually for assessing data quality. The National Agency for the Control of HIV/AIDS also reviews the quality of data submitted from local agencies on a monthly and quarterly basis (MEASURE Evaluation, 2014). The Kenya HIV/AIDS system specifies the groups responsible for ensuring data quality from government sources, and from CBOs and NGOs (the National AIDS Control Council and Constituency AIDS Control Committees, respectively).

Seven systems have somewhat established processes for data verification, where there are gaps in the implementation of these processes. For example, in Rwanda, data verification processes exist but are not well-understood or used often at the sector level (Murray-Zmijewski & Gasana, 2010). In the Malawi HIV/AIDS system, data verification is only established for one type of data, antiretroviral treatment information (Lowrance, et al., 2007). In Colombia, there is a team responsible monitoring data quality, but no system of regular data audits (Independent Evaluation Group, 2007). More systems (nine) describe plans to implement data verification processes. 13 systems, including six that do not describe any data verification processes, note that the role of the national statistics office is to verify official national statistics. These systems may therefore have verification processes carried out by the national statistics office that are not described in their M&E documentation.

Some evaluations of M&E systems report on challenges with ensuring data quality¹⁰. Several systems report inconsistent data quality from different sources (e.g. Mexico, Rwanda HIV/AIDS, Malawi HIV/AIDS, Ethiopia Health). For example, in the Ethiopia Health M&E system, data from routine sources and population-based surveys are inconsistent, and in the Rwanda HIV/AIDS system, data from different CBOs vary in quality. Reviewers of the Uganda and Kenya Health systems recommend the development and use of IT M&E infrastructure in order to decrease the likelihood of data errors.

¹⁰ We do not review primary data for any of the M&E systems. Our findings are based on internal and external assessments of data quality.

One of the targets for transparency and accountability for development results of the OECD Busan Partnership for Effective Development Cooperation is implementing the common open standard for electronic publication of M&E information (OECD & UNDP, 2014). Electronic publication is facilitated by the use of electronic data collection and storage. 14 systems use information technology (IT) tools for data collection, though in some cases data are first collected and later stored electronically, rather than immediately collected electronically. Another 14 systems use electronic data storage to a limited extent, such as only within some ministries or offices, or have plans to store and manage data electronically. In many cases (e.g. Uganda Health, Rwanda Health), a hybrid system using both electronic databases and paper record-keeping exists, with intentions to use electronic databases more fully in the future.

Eleven of 42 systems use IT M&E systems that integrate data collection, aggregation, and verification, as shown in column 2 of *Figure 4*. The Malawi WASH system and the Belize HIV/AIDS system use standardized MS Excel forms as the basis for their IT systems. A few systems describe recent upgrades to their IT data management systems, such as Colombia which introduced a new software package with increased functionality (Independent Evaluation Group, 2007), and South Africa where the system was redesigned to be more user-friendly (Goldman, et al., 2012). Additional detail on specific IT M&E systems is included in the review spreadsheet. The use of IT M&E systems can reduce the potential for errors from manual data collection aggregation (e.g. Kenya Health, Nepal HIV/AIDS). Thirteen systems describe plans to develop IT M&E systems. Twelve systems include public online databases of M&E information, and another eight describe plans to develop such databases. A significant challenge in the use of either IT systems or basic electronic data storage is the availability of electronic infrastructure, such as computers and internet access (e.g. South Africa, Ghana, Malawi Health, Malawi HIV/AIDS), as well as local capacity to use these tools (e.g. Ghana, Rwanda, Ethiopia Health).

4.3 Analysis and Use¹¹ of Results Information

Key Findings:

- Almost all (39 out of 42) systems have strategic frameworks that guide M&E activities. Seventeen are expressed as some kind of theorized causal chain (logframe, results chain, or theory of change).
- While 34 systems involve tracking changes in indicators over time, just 13 of these explicitly describe how they use before and after data from time series to evaluate the effectiveness of their programs. Four systems (Nepal, Chile, Rwanda, and Mexico) use control groups in addition to time series to evaluate impacts of specific programs.
- Six systems go beyond measuring changes in outcome indicators and also evaluate the cost-effectiveness of their activities.
- Most M&E systems we review have well-developed output indicators, which allow governments to assess their performance in implementing their strategy. 14 systems conduct or plan to conduct audits of expenditures against budgets to assess implementation performance.
- 18 systems have clear, established processes for reviewing M&E data to inform strategic planning and program management, such as periodic performance reviews. 14 other systems have plans to implement such processes.
- In 14 systems, information gathered from M&E systems helps to guide budgeting through performance-based budget allocations.
- 17 systems include joint or coordinated progress or performance reviews with stakeholders inside and outside of government.

Many M&E systems are criticized for concentrating on the monitoring component and ignoring evaluation (Porter & Goldman, 2013), and for collecting and reporting information on various indicators without analyzing what the results mean for the effectiveness of government activities and applying that analysis to budgeting and planning. Holvoet, Gildemyn, & Inberg (2012) evaluate the M&E systems of 20 countries in SSA that had finalized a second Poverty Reduction Strategy Paper (PRSP) as of May 2010. They find that about half do not make any distinction between monitoring and evaluation, and those that do “fail to explain how ‘monitoring’ will feed into ‘evaluation.’” Similarly, the authors find that there is “a general tendency for monitoring to crowd-out evaluation,” and add that “it would be naïve to assume that political leaders and bureaucrats are always eager to learn from evidence.” Fewer than half of the M&E systems we reviewed specify who is responsible for budget monitoring, evaluation, or planning activities in the context of M&E, though many of these systems

¹¹ For the purpose of this review, we define “use” of M&E information as using the outputs of monitoring and evaluation activities to inform decision-making around government strategy and planning and around program management.

may still conduct activities in these areas. In this section we assess whether and how governments are analyzing and using monitoring and results information. We consider the development of strategic frameworks, use of different types of evaluation, attention to impact evaluation, rules for reporting and dissemination, and evidence of performance management and results-based decision-making.

4.3.1 Strategic Frameworks and Drawing Meaning from M&E Information

A strategic framework details planned activities and outputs, which allow governments to evaluate how well they have done in implementing their strategy, as well as expected outcomes, which allow governments to evaluate how well their strategy is doing in achieving their goals. M&E systems that lack any of these components limit a government's ability to draw meaning from the monitoring information they generate and to assess their performance. Nearly all (39 of 42) of the M&E systems we reviewed have a strategic framework that guides their M&E activities, most of which are related to national development plans. For seven of the general national M&E systems, the strategic framework is a Poverty Reduction Strategy Paper (PRSP). The frameworks for sectoral M&E systems may be a component of or related to larger national strategic frameworks, but in all but one case (Malawi WASH), the health, HIV/AIDS, and agriculture M&E systems have strategic frameworks specific to their sector. Many strategic frameworks take the form of multi-year plans which are periodically assessed and refined, highlighting the need for strong monitoring and evaluation of performance to inform the development of strategies and the targeting of particular outcomes. The review spreadsheet that accompanies this report includes further detail on the strategic frameworks for each M&E system we review.

A strategic framework may attempt to formalize a theory or set of hypotheses about the relationships between activities, outputs, and outcomes. M&E systems based on such strategic frameworks can then be used to test the hypotheses underlying government strategy and programs. M&E data can indicate areas where further attention, resources, or time are needed, or where the assumptions and theory may be flawed. Holvoet, Gildemyn, & Inberg (2012) note that the articulation of the causality chain clarifies program theory and improves its "evaluability." They find that indicators in countries with less PRSP experience are rarely integrated into a causal chain. On the other hand, Rwanda provides "a detailed description of the causal chain in explaining results, accompanied by illustration through concrete examples" (Holvoet, Gildemyn, & Inberg, 2012). 17 of the 39 strategic frameworks we review are expressed as a theory of change, results chain, or logframe, with clear connections between planned activities, expected outputs of those activities, and desired outcomes resulting from those outputs. These causality chains take different forms in M&E system documentation, including flow charts, detailed tables, and lists of outcomes and activities carried out in pursuit of outcomes.

Causality chains make a clear distinction between outputs, which are the direct results of activities, and outcomes, which are the changes or impacts that result from those activities. For example, for the activity of conducting HIV/AIDS screenings, the output would be the number of HIV/AIDS screenings conducted, and an outcome would be the change in the prevalence of HIV/AIDS. Differentiating between outputs and outcomes allows evaluators to assess two components of performance: how they have implemented their strategy and planned activities (implementation), and how their strategy and activities have affected their desired outcomes (impact). Holvoet, Gildemyn, & Inberg (2012) distinguish these types of evaluation as assessing whether you are "doing things right" (implementation) and "doing the right things" (impact).

In spite of the importance of distinguishing between outputs and outcomes for evaluation, 15 of the 42 systems we review do not clearly do so. For six of these systems, we were only able to retrieve and review one document describing their M&E system, so it is possible that this distinction may be made in other documentation. Among the 15 systems that do not differentiate outputs and outcomes are several that focus primarily on outputs and do not generate information on outcome indicators (e.g. South Africa HIV/AIDS) and others that combine output and outcome indicators together without specifying how these indicators will be used differently (e.g. Rwanda Health). The 27 systems that do differentiate between output and outcome indicators do not always use this terminology, but the distinction can be inferred in how they organize the indicators. For example, some strategic frameworks refer to outputs as short-term or intermediate outcomes, while others do not refer to "outcomes" but have indicators for strategic objectives or goals that measure the effectiveness of their strategy.

4.3.2 Implementation Assessment

Indicators in strategic frameworks determine what information M&E systems will generate, and therefore what type of performance is evaluated (Holvoet, Gildemyn, & Inberg, 2012). Most M&E systems we review have well-developed output indicators, which allow governments to assess their performance in implementing their strategy. This assessment is useful for determining whether the human and financial resources allocated to different activities were sufficient to implement

them as planned, and also provides a simple and quantifiable way of measuring the performance of those responsible for carrying out activities and achieving specific outputs - whether government is “doing things right” (*ibid.*).

A frequent challenge among the M&E systems we reviewed, however, is an overemphasis on outputs as opposed to outcomes (e.g. Chile, Ghana). This focus appears to be especially pronounced in health M&E systems, which appear to regularly track a wide variety of output indicators describing quantities of services provided, but do not track as many outcome indicators measuring effects of those services. In some systems (e.g. Ethiopia Health, Kenya Health), the main function of M&E appears to be to monitor outputs - the quantitative services provided by government offices. For example, a 2012 report finds that the South Africa HIV/AIDS M&E system “largely generates data on service quantity” rather than outcomes (Kawonga, Balaauw, & Fonn, 2012). In the Ethiopia Health system, there are 155 indicators but “No indicators in the national minimum core indicator set are linked to the relevant short (1 year), medium (3-5 years), and long-term (10-15 years) targets” for outcomes (Ethiopian Ministry of Health, 2007). In the Kenya Health system, there are between 200 and 500 indicators in use, nearly all focused on outputs (Blumhagen, et al., 2010). This focus on outputs allows governments to monitor what they are doing and whether their activities match up to their plans, but not to evaluate the impacts of their activities.

Another type of implementation assessment is budget or audit evaluation, which analyzes the degree to which expenditures on activities match the planned budgets for those activities. Less than half of the M&E systems we review clearly indicate processes for tracking disaggregated budget and expenditure figures in their M&E documentation. After looking for indicators related to spending and for descriptions of government audits, resource flow analysis, or expenditure frameworks, we find that 16 systems describe or have planned protocols for tracking disaggregated budget figures and 21 systems describe or have planned protocols for tracking disaggregated expenditure figures. Often, these figures are aggregated by specific financial departments, such as a ministry of finance, budget department, comptroller general, auditor general, or treasury, but in some cases the figures are aggregated by an M&E office.

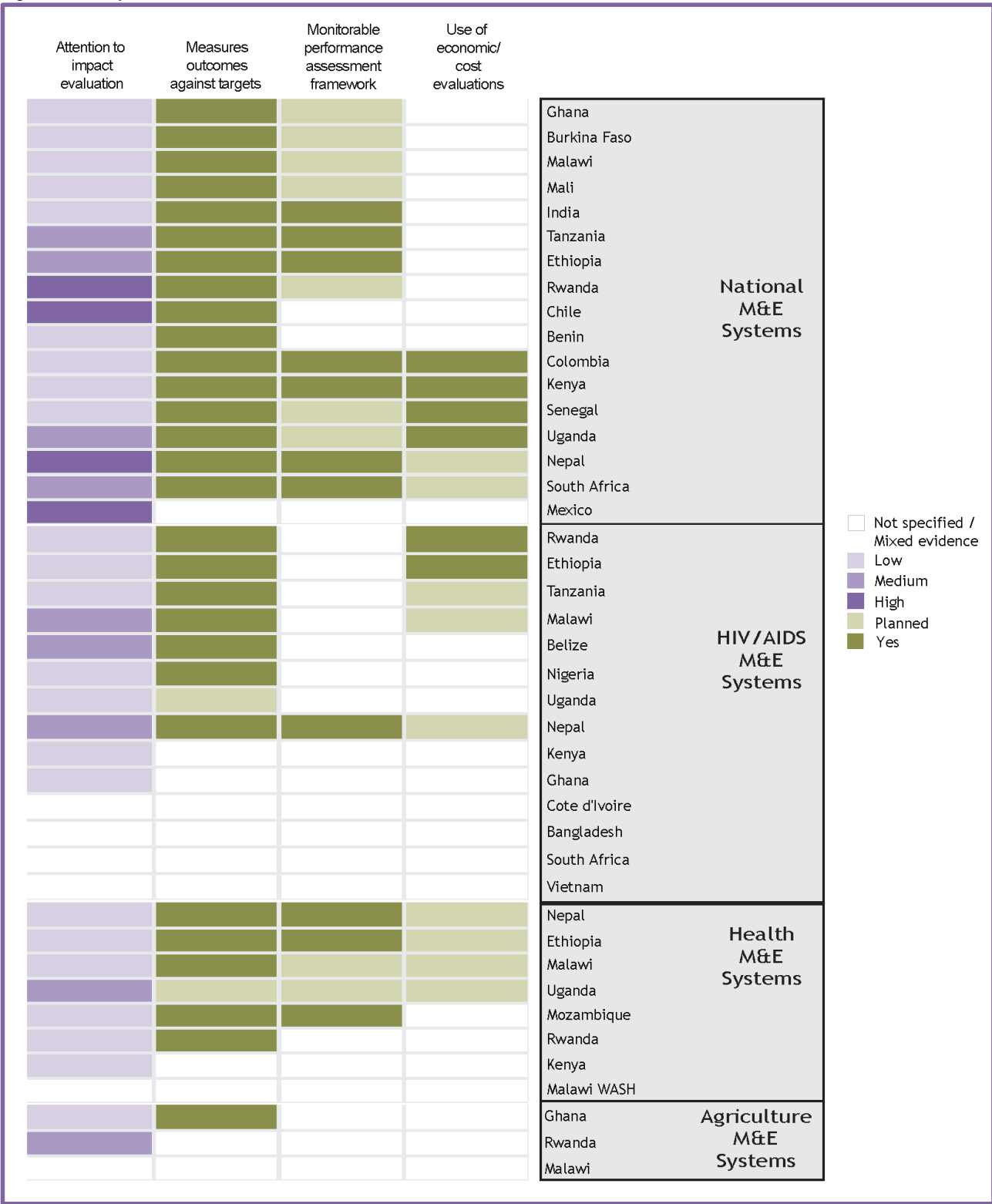
Of particular interest to international donors is whether recipient governments record aid flows and monitor expenditure of these funds on different activities. Donors often establish parallel M&E systems to track how the funds they provide are spent and the resulting impacts, but international agreements such as the Paris Declaration are increasingly promoting harmonized M&E systems to reduce monitoring and reporting demands and redundancy (OECD, 2011). However, based on the documents we review only 21 of the M&E systems clearly indicate that they attempt to record even aggregated aid flows, and in many of these cases there are discrepancies between the amounts scheduled by donors and the amounts reported by governments (OECD, 2011). Even fewer M&E systems appear to clearly track funding sources for different activities, as we found evidence of this in just four of the systems. Given that our review generally does not include detailed budget or accounting documents, however, it is possible that more of these systems do monitor funding sources. Few of the systems we review appear to consider accounting and tracking of funding as a component of M&E.

While the emphasis on reporting of budget allocations and expenditures varies within these M&E systems, there is evidence that nine of them conduct regular audits of expenditures, and another five describe plans to do so. Audits are most common among general national M&E systems, with nearly half of them (8 of 17) conducting or planning to conduct them. These generally take the form of regular expenditure reports that are reviewed by specific finance offices, usually the same office charged with tracking budget and expenditure information (e.g. Kenya, Senegal, Chile, Nepal, South Africa). These audits are also an important accountability mechanism for ensuring that government funds are actually spent on the activities to which they were allocated rather than diverted or embezzled (e.g. Mali, India). The low prevalence of audits in the M&E documentation we review may indicate that these types of implementation assessment are not usually considered a part of M&E and are described elsewhere, potentially in the plans for budget, finance, or audit departments.

4.3.3 Impact Evaluation

A second type of performance evaluation is impact evaluation, which assesses the effectiveness of activities in achieving desired outcomes, or “doing the right things” (Holvoet, Gildemyn, & Inberg, 2012). Such evaluations can take many forms, including comparing changes in outcomes over time, comparing outcomes to targets, and comparing outcomes in groups or areas with different levels of government activities. For our review, we considered any M&E system that at a minimum tracked and compared outcomes over time as conducting some form of impact evaluation. *Figure 5* is a heat map where the M&E systems we review are ordered according to their ratings on different aspects of impact evaluation.

Figure 5. Analysis of M&E Information



Of the 42 M&E systems, 27 have evidence of tracking outcomes over time in order to evaluate effectiveness. We do not specifically review the data generated by these systems, so our findings on the use of time series is based on what is reported in the documents we review. This evidence includes strategic plans reporting results for different indicators in

previous years, M&E plans using previous results as a basis for setting performance targets, and M&E reports describing progress in the measurement of outcome indicators. A further ten M&E systems do not have evidence of actively comparing outcome indicators over time, but demonstrate plans to begin doing so. The most common form of impact evaluation is based on time series, or collecting data over time starting from a baseline, and measuring changes in outcomes.

We evaluate the 42 systems according to the attention they give to evaluating the effectiveness of their programs (column 1 in *Figure 5*). A “low” rating indicates the collection of data over time without specific attention to the potential causes of changes in outcome measures. A “medium” rating indicates that the systems specify how they will use before and after data from time series to associate the introduction of new strategies or programs with changes in outcomes. A “high” rating indicates that the system uses before and after data as well as control groups that do not receive certain government programs to evaluate the impact of these programs. We do not have sufficient evidence to rate attention to impact evaluation for eight systems.

Figure 6. Attention to Impact Evaluation

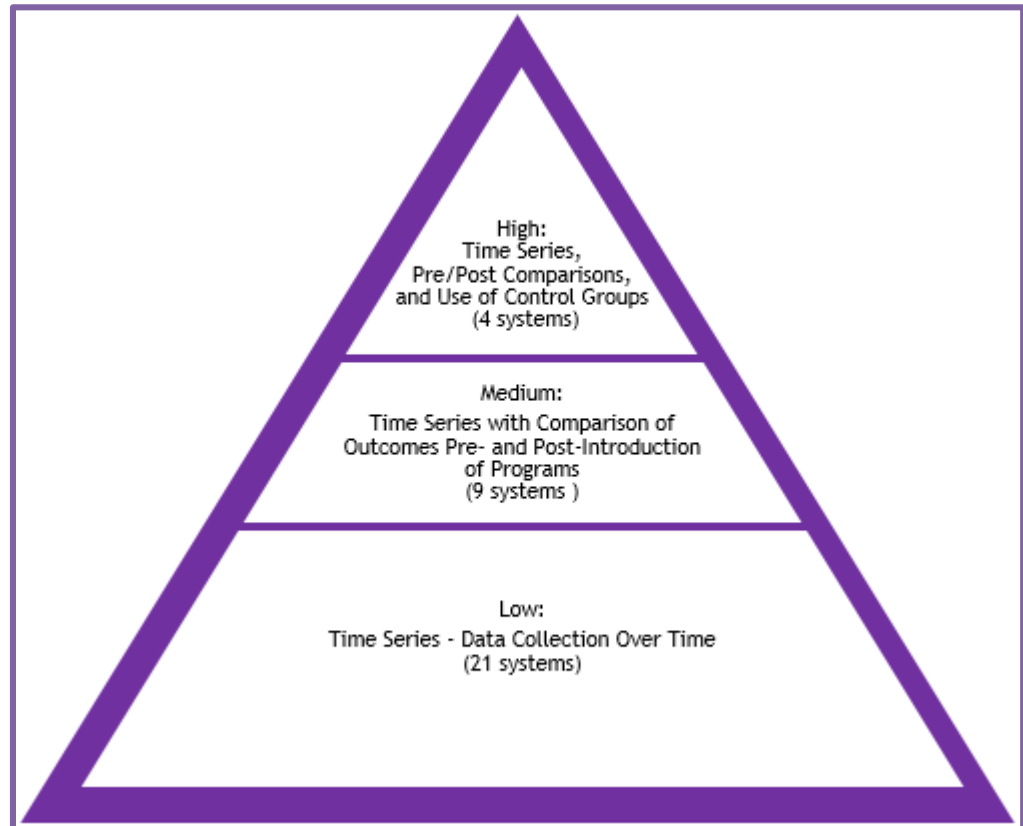


Figure 6 illustrates the different levels of attention to impact evaluation.

While 34 M&E systems use or describe plans to use time series data, just 13 of these explicitly describe how they use time series to evaluate the effectiveness of their programs. The other 21, which we characterize as having low attention to impact evaluation, do not clearly state how measuring outcome indicators over time would allow them to evaluate effectiveness of particular programs. For example, Colombia’s database system allows them to collect baseline performance, annual targets, and actual performance for comparison against targets (Independent Evaluation Group, 2007), but there is no evidence that they use their data to evaluate change over time. Nine M&E systems with “medium” attention to impact evaluation specify using time series to evaluate the effectiveness their programs by comparing outcome measures before and after certain programs are introduced. For example, the 2008 Belize HIV/AIDS M&E Plan specifically indicates that time series analysis are used to evaluate changes over time, and states how baseline information is used to accomplish this (Belize National AIDS Commission, 2008).

Four M&E systems (Nepal, Chile, Rwanda, and Mexico) with “high” attention to impact evaluation use time series in addition to demonstrating occasional use of control groups to try and associate changes in outcomes with specific programs. Control groups are not used to evaluate all programs. In Mexico they were used in the evaluation of the Oportunidades program (Castro, et al., 2009); in Chile they are used for new programs (Guzman, Irarrazaval, & de los Rios, 2014); and in Nepal they are planned for particular programs and projects (National Planning Commission of Nepal, 2013). Rwanda also sometimes conducts randomized controlled trials (Holvoet, Gildemyn, & Inberg, 2012).

A majority of systems also use targets for outcome indicators as a means of evaluating performance and impact (column 2 in *Figure 5*), with 23 systems measuring outcomes against targets for success and another eight systems planning to¹². Holvoet, Gildemyn, & Inberg (2012) find that while results are compared against targets in most M&E systems, reports “lack analytical depth, [...] leaving the reasons for the discrepancies unexplained.” The M&E systems that we reviewed do not go into a great deal of detail on the process of selecting targets, though in many cases targets are related to baseline measures for outcome indicators, and in a few cases they appear to be influenced by sets of international goals such as the Millennium Development Goals. For example, Ethiopia’s general national M&E system uses preset MDG targets for some indicators.

In many cases the targets are formalized in a performance or results framework (column 3 in *Figure 5*), which outlines the expected outcomes and targets for a given time period as a means of evaluating performance. A 2011 OECD report on implementation of the Paris Declaration on Aid Effectiveness includes an assessment of whether governments had developed a results-oriented framework for M&E, and a 2012 review of health aid effectiveness (Shorten, et al.) reports on whether health M&E systems in several countries include a performance assessment framework. Based on the information from these two reports and on supporting evidence from other M&E documents, we find evidence that eight M&E systems have largely developed performance assessment frameworks formalizing targets and expectations, and another 12 M&E systems have planned or taken some action towards developing such a framework. However, it is possible that several of the other M&E systems also have performance assessment frameworks, but do not report them as such. For example, while eight of the HIV/AIDS M&E systems use or plan to use targets for outcomes, just one of these systems (Nepal HIV/AIDS) clearly describes a formalized framework for using these targets to evaluate performance.

Six M&E systems go beyond measuring changes in outcome indicators to assess performance and also evaluate the cost-effectiveness of their activities, as shown in column 4 of *Figure 5*. Another nine M&E systems describe plans or processes for conducting cost evaluations that we cannot verify were implemented. Cost evaluations combine disaggregated information on government expenditures on different activities with measures of the outcome indicators connected to those activities, in order to assess the cost of achieving desired impacts. These cost evaluations are often described in terms of assessing “value for money” (e.g. Kenya, Uganda, Senegal). Holvoet, Gildemyn, & Inberg (2012) argue that cost effectiveness evaluations may appear counterproductive to governments if they continue to receive budget support without having to demonstrate effectiveness.

22 M&E systems describe specific challenges with their capacity for evaluation. Challenges include insufficient funding and staffing (e.g. Malawi Agriculture, Uganda Health, Ethiopia, Tanzania, India, Ghana, Kenya Health), low levels of training and experience (e.g. Nepal, Malawi (all four systems), Uganda, Senegal, Colombia, Kenya HIV/AIDS), and weak institutional support for results measurement (e.g. Rwanda, Mexico). Some systems address capacity challenges by supporting internal capacity building. Chile started with simple desk evaluations to build capacity for more complex evaluations, while South Africa established an Evaluation and Research Unit to drive the system and provide training and technical support. Other systems get support for evaluations from civil society or development partners. A case study of evaluation in Kenya recommends drawing on outside evaluation experience (Porter & Goldman, 2013), and a UNICEF evaluation finds that Mali strengthened evaluation by developing a formal national civil society association for evaluation to provide technical support and help institutionalize a culture of evaluation (Segone, 2010). Our review, however, indicates that just six M&E systems commission external independent evaluations, and these are the M&E systems that are also very strong in our other measures (Chile, Colombia, Mexico, Nepal, Nepal HIV/AIDS, South Africa). Further, reducing the number of separate donor M&E systems could free up trained and experienced evaluators to support government M&E systems (e.g. Malawi Agriculture, Uganda Health).

4.3.4 Use of M&E Information

The M&E systems we review vary in the extent to which they use monitoring and evaluation results to influence their planning and budgeting decisions (column 1 in *Figure 7*). We rate systems that consistently incorporate M&E information through regular review and planning processes as having a “high” level of use of results information, systems where there are gaps in established processes as having a “medium” level, and systems that do not demonstrate consistent use of results information as having a “low” level. Two of the M&E systems we review demonstrate “high” use of results

¹² While the use of targets in evaluating impact is common, the effects on performance of using targets are uncertain. According to Binnendijk (2000), setting targets can orient and motivate project managers and staff, clarify the results for which managers will be held accountable, and serve as subjective means of evaluating performance. However, targets may fail to inspire strong performance if they are set too low, inciting reduced effort after they are achieved, or too high, inciting reduced effort in anticipation of not achieving them. Unrealistic targets may also “set in motion perverse incentives to hide or distort the data” (Binnendijk, 2000).

information in strategic planning and program implementation. In the Malawi Health M&E system, results from annual reports influence the annual work planning and budgeting process, and mid-term reviews are conducted to evaluate progress on the strategic plan and influence the development of future strategic plans. The Performance Evaluation System (SED) in Mexico “continually produces M&E information regarding the extent to which goals and objectives of federal programs have been achieved” to inform planning decisions (Castro, et al., 2009).

Nine general national M&E systems demonstrate “medium” use of results information, with some established processes but gaps in how the information is used. For example, a 2013 report finds that in Malawi government offices have avoided policy dialogue forms out of fear of criticism (CLEAR-AA, 2013). The most common reported use of M&E information is in the development of multi-year strategic plans, which incorporate performance and lessons learned from implementation (e.g. Ghana, Ethiopia, Colombia). Colombia has implemented a constitutional mandate for evaluation and a presidential directive on results-based management, with the goal of promoting social accountability (Independent Evaluation Group, 2007). In some cases M&E information also influences annual work plans (e.g. Ghana, Colombia, Kenya).

Ten of the systems we reviewed (including eight HIV/AIDS M&E systems) outline plans for institutionalizing the use of results information in planning, but we cannot confirm whether these plans have been implemented. For example, Nepal plans to use M&E information to support accountability for service delivery, measure results, evaluate impacts, and improve effectiveness through results-oriented planning (National Planning Commission of Nepal, 2013). Another ten M&E systems do not discuss the extent to which M&E results influence planning, so we cannot assess use. 11 M&E systems demonstrate “low” use of M&E information in planning, either because results information is not made available during the planning process or because there are no established processes for incorporating it.

Figure 7. Use of M&E Information



Challenges for the use of results information include:

- lack of capacity for analysis (e.g. Rwanda Health, Rwanda HIV/AIDS, Kenya Health, Malawi Agriculture),
- data quality issues (e.g. Uganda, Uganda Health, Ethiopia Health, Mali, Rwanda HIV/AIDS),

- weak links between the M&E system and decision-makers (e.g. Uganda, Tanzania, South Africa, Rwanda, Nepal, Malawi HIV/AIDS), and
- lack of culture of performance-based management (e.g., Mali, Malawi, Malawi Agriculture, Chile).

Fourteen M&E systems appear to regularly use results data and evaluations to influence budget allocations (column 2 in *Figure 7*), generally as part of annual plans. In Malawi and Mali for example, the M&E plans describe how the national budgeting process for the next year begins with the completion and review of the annual sectoral review. Some M&E systems (e.g. Rwanda Health, Uganda, Nepal) have established processes for performance-based financing, with budget allocations clearly connected with M&E results information. Another ten systems describe plans for connecting budget allocations to results information. Some of the other M&E systems likely incorporate results information in their budgeting processes, but the documents we reviewed do not describe these processes.

Eighteen systems have clear, established processes for reviewing M&E information and developing policy and strategy recommendations based on that information, although 14 other systems have plans to implement such processes (column 3 in *Figure 7*). In most cases, the lack of clear processes for using M&E data appears to be mainly due to low levels of demand. While there is demand for M&E data for reporting purposes, this does not seem to translate into high demand for information for planning purposes. In Malawi, decision-making is more influenced by patronage rather than results information (CLEAR-AA, 2013), and in South Africa M&E data is often seen as a means to show positive impacts of government rather than a tool for improving performance (Goldman, et al., 2012). In Kenya, M&E information is feared as a policing tool rather than demanded as a means to improve efficiency and effectiveness (CLEAR, 2012).

In some cases (e.g. Mali, Uganda, Malawi Agriculture), M&E systems appear to be implemented mainly as a response to donor reporting requirements. In 17 systems, review processes include joint or coordinated progress or performance reviews with stakeholders inside and outside of government (column 4 in *Figure 7*). These reviews usually take place annually (though they may be more frequent) and can involve government bodies at different levels, private sector organizations, civil society organizations, and development partners. Their objective is to go over government performance and influence planning, sometimes resulting in the production of annual plans (e.g. Mexico, Kenya Health). Another 12 systems describe plans for joint progress reviews.

In some systems (e.g. Colombia, Kenya, Nepal), there is evidence that sharing results information with stakeholders supports government accountability and can encourage the use of M&E data in planning processes. The most common interval for primary national-level M&E reports is annual, with 22 systems reporting annually including 11 of 17 general national M&E systems. HIV/AIDS systems are more likely to produce more frequent reports, with one producing twice-annual national reports, four producing quarterly reports, and one producing monthly reports. Most systems have procedures for disseminating M&E reports internally to different offices and different levels of government, as 28 disseminate M&E information internally while four describe plans to do so. For the other ten systems, it is not clear to what extent they share performance reports within the government.

The number of systems disseminating reports to external stakeholders is greater, with 29 systems sharing M&E information with the public, civil society organizations, and development partners, and another five with plans to do so. In some cases countries share results directly with stakeholders. For example, Belize HIV/AIDS “issues a quarterly bulletin to key stakeholders” (Belize National AIDS Commission, 2008) and Rwanda shares information with donors through the Common Performance Assessment Framework (Murray-Zmijewski, & Gasana, 2010). Some countries make information available on websites. For example, Chile makes evaluation summaries available on the DIPRES website (dipres.goc.cl) and the Uganda Health system shares information on the MOH website (hris.health.go.ug), while 12 systems post M&E data on publicly available online databases. In other cases information is shared externally through joint sector reviews (e.g. Rwanda), or through the media (e.g. Uganda Health, Uganda HIV/AIDS).

4.4 Harmonization of Government and Development Partner M&E Systems

Key Findings:

- We rate seven systems as “high” rating for harmonization, indicating strong actions to align government and donor M&E systems and reduce redundancies such as parallel reporting and implementation systems. 13 systems have a “medium” rating, and 12 have a “low” rating, indicating increases in parallel systems between 2005 and 2011 or other evidence that donors have separate M&E systems.
- There is insufficient information to assess the level of mutual accountability for 18 out of the 42 systems we reviewed. National systems are rated based on Organisation for Economic Cooperation and Development (OECD) data, with nine out of 17 rated “high.” Other systems are assessed based on case studies and mutual assessments of M&E progress. Overall, 11 systems are rated “high” for mutual accountability between donors and governments, indicating the presence of partnership agreements and joint performance reviews. Ten systems are rated “medium” and three systems are rated “low.”
- 29 out of 42 government M&E systems disseminate information to external stakeholders through either periodic reports, public meetings, or joint reviews.
- 32 out of 42 systems use indicators from at least one external framework, with Millennium Development Goals as the most common.

Alignment is one of the key principles in the OECD’s 2005 Paris Declaration, an agreement on aid effectiveness reached between more than 100 donors and aid-receiving countries. In this report, “alignment” refers to both the use of common M&E systems and a common development agenda by donors and countries. Having separate, parallel donor and country M&E systems is thought to hinder the development of in-country institutions and overburdens government ministries with multiple data collection and reporting requirements (OECD, 2011).

In this section, we examine current trends, challenges, and best practices in donor-government M&E alignment, drawing on questions from our framework related to harmonization of M&E frameworks and indicators, technical and financial support from development partners, coordination of data collection and reporting, and mutual accountability. *Figure 8* is a heat map presenting our findings on donor-government alignment, where M&E systems are ordered according to their ratings on the different questions.

Figure 8. Alignment of donor and country M&E systems

Harmonization of M&E demands	Mutual accountability	M&E info externally disseminated	Support provided by donors		
				Malawi	National M&E Systems
				Ethiopia	
				Mali	
				Tanzania	
				Ghana	
				Uganda	
				Colombia	
				Benin	
				Rwanda	
				Nepal	
				South Africa	
				India	
				Senegal	
				Kenya	
				Mexico	
				Chile	
				Burkina Faso	
				Ethiopia	HIV/AIDS M&E Systems
				Malawi	
				Nepal	
				Nigeria	
				Kenya	
				Tanzania	
				Cote d'Ivoire	
				Uganda	
				Belize	
				Rwanda	
				Vietnam	
				Ghana	
				Bangladesh	
				South Africa	
				Rwanda	Health M&E Systems
				Malawi WASH	
				Uganda	
				Ethiopia	
				Nepal	
				Kenya	
				Malawi	
				Mozambique	Agriculture M&E Systems
				Rwanda	
				Malawi	
				Ghana	

☐ Not specified / Mixed evidence
☐ Low
☐ Medium
☐ High
☐ Planned
☐ Yes

4.4.1 Harmonization of M&E Demands

“Harmonization of M&E demands,” column 1 in Figure 8, refers to the extent to which common M&E frameworks, indicators, and reporting systems between countries and donors are used. In addition to information from case studies and

evaluations of specific M&E systems, we also used the number of parallel project implementation units (PIUs)¹³ as an indicator of harmonization of M&E demands. Decreasing numbers of PIUs parallel to existing in-country management systems indicates that donor and country systems are becoming more aligned. The OECD's 2011 Aid Effectiveness report includes changes in the numbers of parallel PIUs between 2005 and 2010 as an indicator of donor-government harmonization (OECD, 2011).

As displayed in *Figure 8*, HIV/AIDS M&E systems have the greatest proportion of “low” ratings (seven of 14 systems), and general national systems are most often rated “medium” (seven of 17). A “low” rating indicates increasing parallel PIUs between 2005 and 2011 or other evidence that donors have separate M&E systems, such as the HIV/AIDS system in Cote d'Ivoire where a heavy reliance on donor M&E systems was reported. A “medium” rating is used when the documentation reveals some efforts to harmonize donor and government M&E systems, as in Malawi, where parallel PIUs were reduced by 50% between 2005 and 2010, and Ghana Agriculture, where the Canadian International Development Agency uses a mixture of Ghana's and its own M&E frameworks. A “high” rating indicates strong actions to align government and donor M&E systems and reduce redundancies. For example, Senegal's general national system decreased parallel PIUs by 80% between 2007 and 2010. In Rwanda, a 2013 report finds that development partners are “progressively orienting their logical framework at the project level” to align with the government's national strategic framework (Newfarmer, Michele, & Vijili, 2013).

4.4.2 Mutual Accountability and Donor Support

Mutual accountability between donors and countries was measured by agreements about how aid should function, as well as regular assessments of the effectiveness of these agreements in achieving development goals. Our ratings of mutual accountability are displayed in column 2 of *Figure 8*. For national systems, these ratings are largely based on the OECD's report on effective development cooperation (OECD & UNDP, 2014). The OECD's five indicators for mutual accountability are:

- the existence of an aid or partnership policy between donors and countries on development goals,
- specific targets for development goals,
- periodic joint assessments involving countries and donors assessing progress towards the targets,
- involvement of non-executive stakeholders in reviews and assessments, and
- the public availability of the results of joint reviews.

We rate mutual accountability as “high” when at least four of the five OECD indicators were met, “medium” when two or three were met, and “low” when one or fewer were met. Based on these criteria, nine of 17 general national systems receive a “high” rating for mutual accountability. Rwanda's general national system is exceptional, meeting all five of the OECD's criteria for mutual accountability. The system includes a Common Performance Assessment Framework (CPAF) and a Donor Performance Assessment Framework (DPAF). Donors review government actions through the CPAF, and governments assess donor actions through the DPAF, both of which have specified indicators. Rwanda appears to be unique in including continual mutual performance review in its M&E plan, as opposed to periodic joint reviews.

For health M&E systems, Shorten, et al. (2012), assess mutual accountability through periodic mutual assessments of progress by countries and donors. Mozambique Health is rated “high,” as Shorten, et al. (2012) report that they have developed a global partnership agreement and conduct regular mutual assessments. Two systems (Ethiopia Health, Nepal Health) that show progress in the quality and frequency of mutual assessments are rated “medium.” In Nepal, more than 70% of development partners signed country health compacts (Shorten, et al., 2012). Ethiopia is an original signatory of the International Health Partnership and related initiatives (IHP+), and the strategic health plan stipulates that government and development partners “make a commitment and adhere to principles of harmonisation and alignment” (Ethiopian Ministry of Health, 2010). Agriculture and HIV/AIDS systems are rated based on case studies, when available. Uganda Health and Ghana Agriculture are rated “low” because of a lack of periodic reviews and the ineffectiveness of periodic reviews, respectively. For 18 systems, we do not have enough information to judge the level of mutual accountability.

The external dissemination of M&E information by countries and the provision of technical M&E support by donors (columns 3 and 4 in *Figure 8*), are also related to mutual accountability. In the majority of cases (29 of 42), we find that government M&E systems disseminate their results information externally, generally through periodic reports (mostly annual) but occasionally through public meetings or joint reviews. Another five systems describe plans but no evidence of follow-through for sharing results information with development partners.

¹³ PIUs are structures donors create for the management of projects funded by aid money.

In addition, we find evidence that donors provide technical support for M&E implementation in most cases (32 of 42). Donor support includes strategic guidance in developing M&E systems, funding for different M&E activities (especially for analysis and evaluation), evaluation and research support, and training and capacity building in different technical areas such as data collection and analysis. For example, in Mali donor groups are organized in different sectors to coordinate technical and financial assistance (including building national capacity on data collection), publishing annual action plans, and funding a large portion of data collection and analysis (Segone, 2010; Bedi, et al., 2006). In Uganda, donors are represented on the national evaluation sub-committee to provide evaluation expertise (IMF, 2010). Research and evaluations by development partners in South Africa help to inform government multi-year reviews (Goldman, et al., 2012). In Benin, evaluations are carried out by both the government and by development partners (Paul & Dossouvi, 2011).

4.4.3 Incorporation of External Frameworks

Standard lists of indicators and targets are a key component of the relationship between countries and donors. The use of common indicators allows for some comparison between the rates of progress in poverty alleviation and development in different countries. The use of indicators from at least one external framework is specified in 32 of the 42 systems we review. *Table 1* displays the number of M&E systems that specifically reference each type of external framework¹⁴, with Millennium Development Goals (MDGs) as the most common.

Table 1. External frameworks incorporated into strategic frameworks, by type of M&E system

External framework	Agriculture	Health	HIV/AIDS	General	Total
MDGs	2	7	9	11	29
African Health Policy/Strategy		2			2
Paris Declaration		1		1	2
PEPFAR			2		2
UNGASS (or other UN accord)			8		8
Global Fund			3		3
Universal Access			4		4
Others		4	4		8
Not specified	1	1	2	6	10

14 M&E systems use a combination of indicators and targets selected from several different external frameworks. Both Rwanda's and Nigeria's HIV/AIDS systems use indicators from the MDGs, UNGASS, PEPFAR, Global Fund, and Universal Access, and Mali's general national system includes indicators from the MDGs, the Paris Declaration and other frameworks. External frameworks are usually adapted to fit the specific goals of national and sectoral M&E systems. In Burkina Faso, MDG indicators are used but targets are adjusted, often downwards (IMF, 2012). On the other hand, Ethiopia's general national M&E system considers reaching MDG targets as a minimum threshold of success for its poverty reduction strategy, and therefore uses the preset targets (IMF, 2011). Since indicators may be adapted to national contexts, it is not clear if results information on indicators from the same external frameworks but different government M&E systems is comparable.

¹⁴ Some systems may use indicators from external frameworks but not specifically refer to them in their M&E documents.

5. Characteristics of Different M&E Systems

Key Findings:

- The evidence we review for general national M&E systems is stronger than the evidence for the sector-specific M&E systems, as 16 of 17 such systems have “medium” or stronger bodies of evidence, compared to two of eight health systems, six of 14 HIV/AIDS systems, and none of three agriculture systems.
- Internal demand for use of M&E results information is higher in general national M&E systems than in the sector-specific systems; all four systems that demonstrate high levels of internal demand are general national M&E systems (Colombia, Ethiopia, India, Nepal).
- Roles and responsibilities for M&E are poorly defined in health systems, with five of eight rated as “low” for clarity of M&E roles, due partly to difficulties in coordinating multiple vertical reporting systems within the health sector. In contrast, we categorize 10 out of 14 HIV/AIDS systems as having a “high” level of clarity around M&E roles and responsibilities, and just one with a “low” level (Cote d’Ivoire HIV/AIDS).
- General national M&E systems have lower overall ratings in data collection than HIV/AIDS and health M&E systems. For eight of 17 general systems there is no evidence of documented rules for data collection, aggregation, or verification. In contrast, ten out of 14 HIV/AIDS M&E systems have documented data collection rules, and ten out of 14 also have rules and standards for data aggregation.
- All four of the systems that specify using control groups in addition to time series to evaluate impact are general systems (Chile, Mexico, Nepal, Rwanda), and nearly all general national systems (16 of 17, with four “planned”) demonstrate the use of targets to evaluate impact, compared to 15 (with three “planned”) of 25 sectoral systems.
- General national systems are also more likely to have implemented or planned formal performance assessment frameworks (14 of 17 compared to six of 25 sectoral systems) and to carry out or plan audit evaluations examining budget allocations and expenditures (eight of 17 compared to six of 25).
- Nine of 17 general systems demonstrate “medium” or higher levels of using results information to influence planning, compared to one of 14 HIV/AIDS systems (eight of which do describe plans to use results information) and one of 11 health and agriculture systems.

Of the 42 government M&E systems we review, 17 are general national M&E systems that aim to monitor and evaluate progress across many sectors. The remaining 25 are national in scope but sector-specific, including eight health-sector M&E systems (of which one focuses on WASH), 14 separate HIV/AIDS M&E systems, and three agriculture M&E systems. In this section we analyze characteristics of these types of government M&E systems, and compare trends across the different types. Unless otherwise specified, all comparisons are between one particular type of government M&E system and all the other systems we review.

5.1 General National M&E Systems

The 17 general national M&E systems are designed to monitor and evaluate government performance across all sectors. We review seven systems from East and Southern Africa (Ethiopia, Kenya, Malawi, Rwanda, South Africa, Tanzania, and Uganda), five systems from West Africa (Benin, Burkina Faso, Ghana, Mali, and Senegal), three systems from Latin America (Chile, Colombia, and Mexico), and two systems from South Asia (India and Nepal). The bodies of evidence are generally of “medium-high” (six systems) or “high” (four) strength, with an average of over six documents reviewed per system. Only India, for which the only document we were able to retrieve is a 2013 evaluation of its general national M&E system, receives a “low” rating for the strength of the body of evidence.

Internal demand for M&E is higher for general national M&E systems than for the sector-specific systems, with “high” ratings for internal demand notably in countries like South Africa, Chile, and Mexico where development partners have a reduced role. The higher demand is often driven by strong interest from elected officials, as the office of the president or prime minister is an important driver of M&E demand in nine general national systems, and in many of these systems the executive office has recently pushed for M&E reforms (e.g. Mexico, Ghana, Uganda). Challenges for M&E demand in general national systems include coordinating parallel M&E systems, increasing awareness of results information, promoting internal ownership of the M&E system to reduce dependence on development partners, and establishing clear legal mandates for M&E coordination.

General national systems appear to have more challenges than sectoral systems with coordinating M&E activities, as only one (Mexico) has a “high” level of harmonization of M&E activities and five (South Africa, Senegal, Mali, Benin, Malawi) have “low” levels. A greater variety of offices are involved in general national M&E systems, including independent agencies, multi-sectoral agencies, ministries, and ministry sub-committees. In many cases multiple offices have overlapping roles and responsibilities. In contrast to the sectoral M&E systems, general national M&E systems are unlikely to have a single office tasked with most M&E responsibilities (six of 17). Of the 16 systems with multiple different offices with leading roles in M&E, 11 are general national systems. General systems are also more likely to have challenges with harmonizing M&E frameworks across government offices (nine of 17), since different sectors also have their own frameworks that need to be aligned with the national framework.

Effective data collection and monitoring appears to be a challenge for some general national M&E systems. The heat map in *Figure 4* in section 4.2 illustrates that general systems are much worse in this area than HIV/AIDS and Health M&E systems. For many general systems (eight of 17) there is no evidence of rules or standards for data collection, aggregation, or verification. These processes may exist, but they are not formalized in the documents we reviewed. Only Ethiopia, Mexico, Ghana, and Rwanda demonstrate either existing or planned processes for all three of these activities. In some cases (South Africa, Colombia, Kenya), a functioning IT M&E system for data collection, aggregation, and verification may reduce the need for clear rules, as these processes are automated by software tools. However, outside of these three countries, just three (Ethiopia, Ghana, Uganda) have somewhat implemented IT M&E systems to support data collection and management, while four (Mexico, Tanzania, Nepal, Malawi) describe plans for using an IT M&E system.

Another particular challenge of general national systems is indicator selection. Challenges include limiting the number of sector-specific indicators, focusing more on strategic and overarching indicators, and ensuring that routine monitoring systems in different sectors are aligned with these overarching indicators and generating the necessary information for measuring them. Rwanda and South Africa represent good practices for indicator selection and alignment. Rwanda instituted a process for developing an overarching national set of indicators, after which different sectors and sub-national offices created their own matrix of indicators to measure performance against overall national goals (Murray-Zmijewski & Gasana, 2010). South Africa signed performance agreements with different ministers, departments, and provinces after agreeing on the set of primary national indicators (Goldman, et al., 2012). Strong links with line ministries and sub-national offices are especially important for general national M&E systems, as the coordinating offices often depend on the capacity of sector and sub-national offices to produce quality information (e.g. Malawi, South Africa, Senegal, Colombia). Differences in data quality among these sources of data is a significant challenge for general M&E systems (e.g. Ghana, Rwanda, South Africa, Tanzania).

In spite of these data collection challenges, general national M&E systems rate higher than sectoral systems on most of our indicators for analysis of M&E data, as seen in *Figure 5* in section 4.3.3. All four of the systems that specify using control groups to evaluate effectiveness in addition to time series are general systems (Chile, Mexico, Nepal, Rwanda), and nearly all general national systems (16 of 17, with four “planned”) demonstrate the use of targets to evaluate impact. General national systems are also more likely to have implemented or planned formal performance assessment frameworks (14 of 17 compared to 6 of 25 sectoral systems) and to carry out or plan audit evaluations examining budget allocations and expenditures (eight of 17 compared to six of 25). In addition, five of the six M&E systems that commission external independent evaluations are general systems (Chile, Colombia, Mexico, Nepal, South Africa).

General national M&E systems also tend to rate higher on the use of results information for decision-making around strategy and program management, as shown in *Figure 7* in section 4.3.4. Nine of 17 general systems demonstrate “medium” or higher levels of use, compared to one of 14 HIV/AIDS systems (eight of which do describe plans to use results information) and one of 11 health and agriculture systems. Sixteen of 17 general systems make some effort to connect budget allocations to results information, though in the cases of Chile, South Africa, Ghana, and Kenya, there is mixed evidence on whether these efforts have a significant influence on budget allocations. Joint performance reviews involving external stakeholders are prevalent in general national M&E systems, with eight conducting regular joint reviews and five planning to do so. In Mali, a UNICEF report finds that “Participatory planning processes involving civil society, children donors, and decentralized regions provide a common consensus on the vision and strategic results and outcomes” (Segone, 2010). In Ethiopia, developing the latest strategic plan involved coordinated performance reviews and discussions with government bodies at different levels, private sector organizations, civil society organizations, and development partners (IMF, 2011). Nepal’s M&E plan states that “Public audits should be performed of the various development activities carried out by government agencies, local bodies, non-government organizations and communities with an objective to improve service delivery” (National Planning Commission of Nepal, 2013).

The general national M&E systems appear to have higher levels of mutual accountability than other M&E systems. Fourteen of 17 general national systems have established processes for sharing results information with development partners, and one other has plans for external dissemination. In addition, nine of the 11 M&E systems with “high” ratings for mutual accountability¹⁵ are general national M&E systems. The exception, within general national M&E systems, is countries with less dependence on foreign assistance, notably South Africa, India, Chile, and Mexico. South Africa is the only one of these systems with a “low” rating, as a 2011 OECD study reports that they do not have reviews of mutual accountability. These countries have lower levels of donor alignment and mutual accountability because their M&E systems are more oriented to internal demands for results information, rather than development partner demands. Their general national M&E systems rate highly on most of the measures in our review framework.

Among the other 13 general national M&E systems, seven are intended to monitor progress of Poverty Reduction Strategy Papers (PRSPs), which entail a high level of mutual accountability with the IMF and World Bank. The five general national M&E systems that rate the highest in our alignment heat map (*Figure 8*) are Ethiopia, Mali, Tanzania, Malawi, and Rwanda. The first three of these countries have PRSPs, while Rwanda has two mutual performance assessment frameworks (described in section 4.4.2.) and Malawi has well-developed mutual accountability systems.

5.2 Health M&E Systems

To increase aid effectiveness in the health sector, sector-wide approaches (SWAp) were introduced in developing countries in the 1990s. SWAp introduced policy frameworks with specific health sector priorities to improve coordination and harmonization between governments and donors, enhance national ownership, and increase government accountability. Holvoet & Inberg (2013) find that the impact of SWAp has been relatively limited as donors continue to rely on their own systems for M&E because they consider government systems to be weak. Health sector M&E systems are often characterized by having multiple vertical systems that can result in fragmentation and duplication of efforts. For example, a World Health Organization review of Uganda’s Health M&E system reports multiple disease-specific M&E systems (e.g. HIV/AIDS, malaria, TB) running parallel to each other, which increases the burden on health workers collecting M&E information and causes complications in aggregation (Boerma & Gore, 2010).

We review eight M&E systems specific to the health sector: Ethiopia, Kenya, Mozambique, Nepal, Rwanda, Uganda, and Malawi (for which we also review one health system specific to WASH programs). The strength of the body of evidence was rated “medium-high” for two of these systems (Uganda Health, Ethiopia Health), “low” for one system (Malawi WASH), and “medium-low” for the other five systems.

Challenges with institutionalization and coordination, shown in *Figure 3* in section 4.1, appear particularly widespread in health M&E systems. In general, the roles and responsibilities for M&E are less clearly defined and the harmonization of M&E activities appears lower than in other types of M&E systems. Uganda Health is the only health system rated “high” for clarity of institutional roles, with an M&E plan that clearly defines roles and responsibilities of key actors and a framework for combining and coordinating activities. Few health systems outline clear roles and responsibilities, or do not specify roles beyond general M&E coordination (e.g. Rwanda, Ethiopia). When specified, the coordinating bodies for M&E are typically located within the Ministry of Health (MoH). In Rwanda, the office responsible for coordinating M&E is the Planning, Policies, and Capacity Building Unit in the MoH, and in Uganda it is the Quality Assurance Department in the MoH. The Public Health Administration Monitoring and Evaluation Division (PHAMED) appears responsible for coordinating Nepal’s Health M&E system. These three systems receive “high” or “medium” ratings for clarity of M&E roles. The other systems with “low” ratings either do not specify an office responsible for coordinating M&E, or lack a coordinating office with a clear mandate and resources (e.g. Ethiopia). Only two health systems (Ethiopia, Malawi WASH) have plans to improve harmonization of M&E activities.

As in the other M&E systems, data collection and monitoring capacities appear to be challenges in health sector systems. Several systems indicate that the quantity of health indicators is too large for effective monitoring. For example, for Ethiopia Health a review recommended tracking a smaller number of core indicators rather than many disease-specific indicators (Ethiopia Ministry of Health & MEASURE Evaluation, 2012). Another challenge is satisfying the needs of different stakeholders within the health sector. In Kenya, stakeholders of programs with more resources and defined information needs (i.e. HIV/AIDS programs) have more influence to advocate for disease-specific indicators than those with less-developed programs (Blumhagen, et al., 2010).

¹⁵ The ratings for mutual accountability are based on aid or partnership policy, targets for development goals, periodic joint assessments, involvement of external stakeholders, and public availability of results of joint reviews.

Half of the health M&E systems we reviewed have documented processes for data collection, and two have rules for data aggregation, though three additional systems have plans to develop them. Just one health system (Uganda) has a “somewhat” developed process for data verification. For most systems (five of eight), data collection is decentralized and collected through local/routine M&E systems, often based in hospitals and health centers. In Kenya, for example, routine data are collected through a network of community units, as well as 6,034 health facilities throughout the country. Data are then transmitted from the district to the national level (Blumhagen, et al., 2010). For the Malawi WASH system, data are collected by Health Surveillance Assistants at the district level (Meek & Samanyika, 2013).

Challenges in data collection include lack of accountability for complete and accurate reporting, inadequate staff training, and conflicting or redundant information flows leading to issues with data quality. In Uganda, health facilities themselves do not use the data they collect, so staff lack motivation to improve data quality (Holvoet & Inberg, 2013). In Malawi, donors have created parallel data collection systems due to lack of trust in data generated by the government system (WHO, 2012). Three of eight health systems (Uganda Health, Malawi Health, Malawi WASH) use an IT M&E systems for data collection, storage, and management, and Rwanda Health has a somewhat developed IT M&E system. Three health systems have plans to develop IT M&E tools, but there are challenges. For example, in Ethiopia, a 2014 report found that a lack of basic infrastructure presents challenges to implementing an IT system (Alebachew, et al., 2014).

Health M&E systems tend to rate lower on analysis of M&E information, as shown in *Figure 5* in section 4.3.3. This finding is consistent with Holvoet & Inberg’s (2013) observation that for the Rwanda and Uganda health systems, the monitoring components have been much more thoroughly developed than the evaluation components. No health systems use cost-effectiveness evaluations, though four systems have plans to introduce them. Uganda is the only health system with “medium” attention to impact evaluation, and six systems receive a “low” rating, suggesting that evaluating the effectiveness of programs is a low priority. Four countries specify challenges related to the capacity for evaluation, including understaffed M&E departments (e.g. Kenya Health), inadequate skills and analytical capacity (e.g. Kenya Health, Uganda Health, Malawi Health, Malawi WASH), and high staff turnover and defection to donor agencies (e.g. Uganda Health).

Seven of the eight health M&E systems (with the exception of Malawi WASH) have a strategic framework specific to the health sector. Four of these are expressed as a theory of change, results chain, or logframe, which help to connect activities and outputs to the desired outcomes and impact. As indicated in section 4.3.1, a frequent challenge in the M&E systems we reviewed is an overemphasis on outputs rather than outcomes. This issue appears particularly often in health sector systems, with many systems tracking a wide variety of output indicators but few outcome indicators. Four of the health systems do not appear to distinguish outputs from outcomes, and only two specify separate short-term, intermediate, and long-term outcomes.

Five health systems have an established process for incorporating M&E information, but only Malawi Health has a high rating for use of results information to influence policy and strategic planning, and four have low ratings (Rwanda Health, Kenya Health, Uganda Health, Ethiopia Health). Just two systems (Rwanda, Health, Uganda Health) indicate that budget allocations are connected to results, though three systems describe plans for the same. Periodic joint progress reviews are one area where Health M&E systems are rated more highly than HIV/AIDS M&E systems, as five health M&E systems conduct these reviews, typically on an annual basis.

While nearly all the health systems we review receive donor support for M&E (seven of eight), we find lower levels of mutual accountability and harmonization of government and donor M&E demands compared to national and HIV/AIDS systems, as shown in *Figure 8* in section 4.4. Mozambique has the only health system that receives a “high” rating for mutual accountability, but two countries, Ethiopia and Nepal, show progress in the frequency and quality of assessments so are rated “medium.” Five of the eight health systems disseminate M&E information externally, via Joint Health Sector Reviews (e.g. Rwanda) or the Ministry of Health website (e.g. Uganda). Only two countries, Ethiopia and Nepal, receive “high” ratings for harmonization between donors and government. In Rwanda, many facilities create separate reports for donors and government, and in Malawi, donors still primarily use their own M&E systems (OECD, 2012). One positive sign of harmonization is the widespread use of external frameworks in health sector M&E systems, shown on *Table 1* in section 4.4.3. Seven of the eight systems incorporate the Millennium Development Goals (MDGs), two systems incorporate the African Health Policy/Strategy, and four include elements of other external frameworks.

5.3 HIV/AIDS M&E Systems

Our review includes 14 HIV/AIDS-specific M&E systems. Ten of these are in sub-Saharan Africa, but one is in Latin America (Belize) and three are in South or Southeast Asia (Bangladesh, Nepal, Vietnam). The strength of the body of evidence for these systems varies. Four are rated “high” or “medium-high” (Kenya, Malawi, Nigeria, Tanzania), seven are rated “medium” or “medium-low” (Cote d’Ivoire, Ethiopia, Bangladesh, Rwanda, South Africa, Uganda), and three are rated “low” (Belize, Ghana, Nepal, Vietnam).

For HIV/AIDS M&E systems internal demand for M&E is “low” in five systems, “medium” in four systems, and “high” in four systems (Cote d’Ivoire HIV/AIDS, Ghana HIV/AIDS, Nepal HIV/AIDS, Nigeria HIV/AIDS). Demand generally comes from multi-sectoral agencies (five out of 14) established to monitor progress in achieving HIV/AIDS outcomes. Examples of multi-sectoral agencies include Belize’s National AIDS Commission, Ethiopia’s HIV/AIDS Prevention and Control Offices (HAPCOs), Nepal’s National Centre for AIDS and STD control, Malawi’s National AIDS Commission, and Nigeria’s National Agency for the Control of HIV/AIDS. National health departments or ministries are other sources of internal demand (two systems). In our rankings, M&E systems with a high level of internal demand for M&E are characterized by engagement of a wide range of stakeholders in strategic planning processes (e.g. Nepal HIV/AIDS, Nigeria HIV/AIDS) and a focus on reviewing results and performance to inform planning (e.g. Nepal HIV/AIDS).

M&E roles and responsibilities in HIV/AIDS systems tend to be well defined. We categorize 10 out of 14 systems as having “high” clarity around M&E roles and responsibilities, and just one as “low” (Cote d’Ivoire HIV/AIDS). The office primarily responsible for coordinating M&E is generally an HIV/AIDS-specific multi-sectoral agency (nine systems), though in three cases it is the Ministry of Health (MoH) and in one case a MoH sub-committee. HIV/AIDS systems are distinct in their use of multi-sectoral agencies for coordinating M&E.

Five systems receive “high” ratings for harmonization of M&E activities. Each of these systems are coordinated by a multi-sectoral government agency and have clear responsibilities for M&E activities. We rate two systems as having “low” harmonization (Kenya HIV/AIDS, South Africa HIV/AIDS). For Kenya HIV/AIDS, M&E activities are decentralized but poorly coordinated, leading to issues with parallel M&E systems, inadequate data quality assurance systems, gaps in data collection, and delays in reporting (Karani, Bichanga, & Kamau, 2014). In South Africa, the HIV M&E system was established as external to the Health system, which has led to some overlap and confusion with M&E activities (Kawonga, Blaawu, & Fonn, 2012).

Each HIV/AIDS system is based on a specific HIV/AIDS strategic framework, and 11 of the systems clearly distinguish between outputs and outcomes. South Africa HIV/AIDS, however, serves as an example of a system that focuses primarily on outputs and does not generate information on outcome indicators (Kawonga, Blaawu, & Fonn, 2012). In eight cases the selection of M&E indicators posed significant challenges. For example, for Ethiopia HIV/AIDS choosing indicators posed a challenge since the intent was not to overburden the Ministry of Health (Ethiopian Ministry of Health, 2009). In Uganda, Cote d’Ivoire, and Malawi the HIV/AIDS systems are challenged by the fact that there are too many indicators.

HIV/AIDS systems generally rate well on indicators for data collection and monitoring. Ten out of 14 HIV/AIDS M&E systems indicate the presence of documented data collection rules. In the Rwanda HIV/AIDS system, separate rules for data collection for health care facilities and for community-based organizations are established (Rwanda Biomedical Center, 2013). The Malawi HIV/AIDS system specifies acceptable data sources in addition to specifying responsibility for and frequency of data collection. Four of the systems receive “high” ratings for the level of local data collection coverage, and only for Belize HIV/AIDS is coverage rated “low.” Ten of 14 HIV/AIDS systems document rules for data aggregation. Most systems are hierarchical, as in the Rwanda, Nigeria, and Nepal HIV/AIDS systems, where data is collected at the facilities level and aggregated at the district, state and national levels for reporting purposes. However, the Tanzania HIV/AIDS system includes a single national database instead of a hierarchical aggregation model (Tanzania Commission for AIDS, 2011).

Four systems include established rules for data verification, though two others have “somewhat” developed rules. The Nigeria HIV/AIDS system, for example, uses a tool developed by the Global Fund at least biannually for assessing data quality. The National Agency for the Control of HIV/AIDS also reviews the quality of data submitted from local agencies on a monthly and quarterly basis (MEASURE Evaluation, 2014). The Kenya HIV/AIDS system specifies the groups responsible for ensuring data quality from government sources and from CBOs and NGOs (Kenyan National AIDS Control Council, 2005). Five HIV/AIDS systems have functional IT M&E systems used as tools for data collection, aggregation, and verification. Four of these systems are specific software suites designed for the purpose of managing HIV/AIDS data. One software suite, the

DHIS 2.0, used by Nigeria, was developed by PEPFAR for this purpose (MEASURE Evaluation, 2014). However, some IT systems appear to be functional without being exceedingly complex. The Belize HIV/AIDS system uses standardized MS Excel forms as the basis for their IT systems (Belize National AIDS Commission, 2008).

Three systems (Nepal HIV/AIDS, Malawi HIV/AIDS, Belize HIV/AIDS) explicitly describe the use of results information over time to evaluate effectiveness, and six track results over time but do not specify whether they attempt to evaluate effectiveness. Eight of the HIV/AIDS M&E systems use or plan to use targets for outcomes. However, only Nepal HIV/AIDS clearly describes a formalized framework for using these targets to evaluate performance. Nepal HIV/AIDS is also the only HIV/AIDS system that commissions external independent evaluations.

In general (eight of 14), HIV/AIDS M&E systems outline plans to use results information in their strategy and planning activities, but we did not find evidence that these plans had been implemented. Only Malawi HIV/AIDS demonstrates a medium level of use of results information in decision-making for strategy and planning. Challenges for the use of results information include lack of capacity for analysis (e.g. Rwanda HIV/AIDS), data quality issues (e.g. Rwanda HIV/AIDS), and weak links between the M&E system and decision-makers (e.g. Malawi HIV/AIDS, South Africa HIV/AIDS). For example, the Malawi HIV/AIDS system was described as underutilizing M&E information and that this information is not effectively disseminated for policy analysis purposes (Malawi National AIDS Commission, 2009).

HIV/AIDS sector systems tend to have more international demand than indigenous demand for M&E. We find evidence of international demand for M&E information in all 14 systems. Strong international demand for HIV/AIDS M&E is also reflected in the prevalence of indicators from external frameworks, such as the Millennium Development Goals and UNGASS, in the HIV/AIDS strategic frameworks of these systems (see *Table 1* in section 4.4.3). However, many HIV/AIDS systems are characterized by low levels of harmonization between donor and government M&E demands. Seven of the 14 systems we review are rated “low” for harmonization between governments and development partners, indicating that there is no, very little, or only planned harmonization of M&E activities.

Three HIV/AIDS systems (Belize HIV/AIDS, Nepal HIV/AIDS, Ethiopia HIV/AIDS) receive “high” ratings for harmonization with development partners, indicating a high level of integration in M&E activities. Evidence of M&E harmonization between governments and development partners includes the use of common M&E frameworks, indicators, and reporting systems between countries and donors. Of all the systems we reviewed, HIV/AIDS systems have the greatest proportion of “low” ratings (seven out of 14 systems). Six out of the seven systems with “low” harmonization (Kenya HIV/AIDS, Ghana HIV/AIDS, Cote d’Ivoire HIV/AIDS, Bangladesh HIV/AIDS, South Africa HIV/AIDS, Uganda HIV/AIDS) point to donor’s demands and separate reporting systems as a challenge for harmonization. Separate development partner Program Implementation Units (PIUs) also reduce harmonization (e.g. Cote d’Ivoire HIV/AIDS, Uganda HIV/AIDS). A 2012 study on South Africa’s HIV/AIDS system indicates, “the desire for control over their funded programme often drives donors to continue supporting even inappropriate vertical systems” (Kawonga, Blaauw, & Fonn, 2012).

5.4 Agriculture M&E Systems

Our evidence base for agriculture M&E systems is limited. We reviewed one document each for agriculture M&E systems in Ghana and Malawi, and two for Rwanda’s system. We judge the evidence for Malawi to be slightly stronger than for Ghana and Rwanda, as the document is a specific report on Malawi’s agriculture M&E system. Drawing on such a small number of sources, our ability to report on trends, challenges, and best practices in agriculture M&E systems is constrained.

When information is available, donor-government alignment appears generally weak in agriculture M&E systems. In Malawi, demand for M&E information comes from a wide variety of bilateral and multilateral institutions, including the IMF, World Bank, USAID, DFID, and others. Harmonization of demands from these donors with Malawi’s own M&E demands is low. Donors use multiple M&E frameworks and reporting requirements not clearly connected to Malawi’s agricultural strategic framework. The misalignment of M&E systems causes serious problems for line ministries in Malawi, which struggle to fulfil the large number of separate data requests from both central government offices and donors (Phiri, 2014). In Ghana, harmonization of M&E demands is somewhat better, but the Canadian International Development Agency continues to use a mixture of its own and Ghana’s M&E frameworks (Senadza & Laryea, 2012).

Except for Malawi, reporting on internal coordination of agriculture M&E systems in our evidence base is lacking. In Malawi, internal demand for M&E information is strong, with demand coming from the Ministry of Agriculture and from academic institutions. However, roles and responsibilities for M&E are not very clear, since there is a current transition in leadership, and harmonization of M&E activities is low (Phiri, 2014).

Reporting on data collection and monitoring is also generally limited. All three countries base agriculture M&E on an agriculture-specific strategic plan, with a list of indicators based on the strategic framework. Malawi's and Rwanda's strategic frameworks incorporate elements of the Millennium Development Goals. Although there appears to be agreement on what type of data to collect, there are some challenges with basic monitoring capacities. In Rwanda, there are established rules for data collection, but there are only plans for an IT M&E system and for rules on data aggregation (Kanyarukiga, 2004). A 2014 report on the Malawi Agriculture system notes the importance of further clarifying rules for data collection. The report also noted that the number and skill level of staff at the Ministry of Agriculture are significant constraints on the improvement of data collection processes (Phiri, 2014).

Based on the available evidence, analysis and use of M&E data also appears weak in agriculture M&E systems. For Malawi Agriculture, there is a low level of use of results information in policy and strategy planning, but our sources did not report on the use of results information in Rwanda Agriculture and Ghana Agriculture. Both of these latter systems have regular joint sector reviews with development partners. It is not clear if these reviews affect government decision-making in Rwanda, but in Ghana's joint sector reviews, a 2012 report found that the reviewers' recommendations tend to be the same every year, suggesting that changes are not implemented (Senadza & Laryea, 2012). In Malawi, a lack of analytical skills among Ministry of Agriculture staff is also cited as a challenge to improvement in the use of M&E information (Phiri, 2014).

6. Conclusion

Porter & Goldman (2013) and Segone (2010) highlight the influence of the Paris Declaration on Aid Effectiveness in 2005 and the subsequent Accra Agenda for Action in 2008 on efforts by donor organizations and recipient country governments to improve the accessibility of results information and align their M&E systems. One aim of these efforts is to develop and strengthen government M&E systems and reduce the number of parallel donor M&E systems (OECD, 2011). In this report we review case studies and government documentation describing 42 country-level M&E systems in 23 developing countries, including 17 general national M&E systems, 14 systems focused on HIV/AIDS, eight focused on health, and three focused on agriculture. We examine four main aspects of government M&E systems to assess their level of development: coordination of government M&E activities, quality of data collection systems, use of the results information generated by the M&E system, and harmonization of government and development partner M&E activities.

The government systems we reviewed have improved the coordination of M&E activities, often following increases in demand for results information from government offices, development partners, and civil society in the country. Ongoing challenges to coordination include formalizing roles and leadership within M&E systems, aligning and coordinating M&E activities across sectors, and building internal government capacity to manage and implement M&E activities. We find significant gaps in the evidence for government data collection and monitoring processes: 17 systems do not specify rules or standards for data collection, 18 do not specify rules for aggregation of data, and 22 do not specify a process for data verification. While many systems have a list of indicators based on a strategic framework, few provide information on how those indicators integrate into their data collection and monitoring process and fewer still offered details on prioritization of outcomes. Most government M&E systems incorporate indicators from external framework, such as the MDGs, but in many countries donors continue to oversee separate M&E systems. Many systems track their indicators over time or plan to, but few attempt to analyze the impact of government programs and fewer still evaluate the cost-effectiveness of programs.

Data collection and monitoring processes in HIV/AIDS M&E systems are strong compared to other types of systems. HIV/AIDS systems' comparative strength may be related to high levels of external demand for HIV/AIDS information. HIV/AIDS data may also be inherently easier to collect and aggregate than data for agricultural, health, and general systems, since the singular focus on HIV/AIDS requires fewer indicators and measures.

We find that data collection and analysis often do not lead to data use. Fewer than half of the systems we reviewed have an established process for using M&E data to inform decision-making around strategy, planning, or program management. Challenges for data use common to several M&E systems are lack of capacity for analysis, data quality issues, weak links between the M&E systems and decision-makers, and a lack of a culture of performance-based management.

There are gaps in the evidence around several of the questions from our review framework. The documents we review do not go into a great level of detail about the process of selecting M&E indicators, and there is little information on processes for monitoring aid flows and funding for government programs. For many countries, the documentation does not specify

who is responsible for budget monitoring, evaluation, or incorporating M&E data into planning processes, and many do not describe processes for these M&E activities. In some cases the lack of evidence may indicate that these activities are not included in the M&E system, but in others these activities may be described in documents other than those we reviewed.

Our findings are limited by the availability of documentation describing the M&E systems we review. We were not able to retrieve a specific M&E plan for 24 of 42 government M&E systems, and did not identify any review of M&E implementation for 14 systems. Many of the documents we did retrieve are several years old and may no longer accurately reflect the status of the M&E systems. Eight of the 18 M&E plans we reviewed are from before 2010, as are five of the 28 reviews of M&E implementation.

These findings offer a more complete picture of the range of government M&E systems and their priorities, aspirations, and challenges. This report contributes to the identification of issue areas and formation of plans to improve efficiency in individual systems and collaboration across systems. Current reviews of the implementation of government M&E systems would provide information on the extent to which these challenges are being addressed.

Please direct comments or questions about this research to Leigh Anderson, at eparx@u.washington.edu.

References Cited in Text

- Alebachew, A., et al. (2014). Universal Health Coverage Measurement in a Low-Income Context: An Ethiopian Case Study. <https://www.hfgproject.org/wp-content/uploads/2014/04/Ethiopia_Case-Study-on-Measuring-Progress-towards-UHC_Final.pdf>.
- Bangladesh NASP. (2007). National AIDS Monitoring & Evaluation Framework and Operational Plan: 2006-2010. <http://hivhealthclearinghouse.unesco.org/sites/default/files/resources/AIDS_plan_Bangladesh2006_2010.pdf>.
- Bedi, T., et al. (2006). *Beyond the Numbers: Understanding The Institutions For Monitoring Poverty Reduction Strategies*. Washington, D.C.: The World Bank.
- Belize National AIDS Commission. (2008). National Monitoring and Evaluation Plan of the National Strategic Plan 2006-2011. <http://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---ilo_aids/documents/legaldocument/wcms_127750.pdf>.
- Binnendijk, A. (2000). Results Based Management in the Development Co-operation Agencies: A Review of Experience - Background Report. DAC Working Party on Aid Evaluation. <<http://www.oecd.org/development/evaluation/1886527.pdf>>.
- Blumhagen, D., et al. (2010). USAID/Kenya: Assessment of National Monitoring And Evaluation And Health Management Information Systems. USAID.
- Boerma, T. & Gore, F. (2010). Strengthening monitoring and evaluation practices in the context of scaling-up the IHP+ compact and Country Health Systems Surveillance: Uganda. <http://who.int/healthinfo/country_monitoring_evaluation/Uganda_CHeSS_report_final.pdf>.
- Castro, M. F., et al. (2009). Mexico's M&E System: Scaling Up from the Sectoral to the National Level. *World Bank ECD Working Paper Series* (20).
- CLEAR. (2012). *African Monitoring and Evaluation Systems: Exploratory Case Studies*. Johannesburg: Graduate School of Public and Development Management, University of the Witwatersrand.
- CLEAR-AA. (2013). *Study on the Demand for and Supply of Evaluation in Malawi*. Johannesburg: Graduate School of Public and Development Management, University of the Witwatersrand.
- Ekouevi, D. K., Karcher, S., & Coffie, P. A. (2011). Strengthening health systems through HIV monitoring and evaluation in Sub-Saharan Africa. *Current Opinion in HIV and AIDS*, 6:245-250.
- Engela, R., & Ajam, T. (2010). Implementing a Government-wide Monitoring and Evaluation System in South Africa. *World Bank ECD Working Paper Series* (21).
- Ethiopian Central Statistics Agency. (2011). Ethiopian Data Quality Assessment Framework (EDQAF). <http://www.poledakar.com/dqaf/images/3/3a/Ethiopie_EDQAF.pdf>.
- Ethiopian Ministry of Health. (2007). Assessment of the Ethiopian National Health Information System. <http://www.who.int/healthmetrics/library/countries/HMN_ETH_Assess_Final_2007_10_en.pdf>.
- Ethiopian Ministry of Health. (2009). STRATEGIC PLAN FOR INTENSIFYING MULTISECTORAL HIV AND AIDS RESPONSE IN ETHIOPIA II (SPM II): 2009-2014. <http://www.aidstar-one.com/sites/default/files/prevention/resources/national_strategic_plans/Ethiopia_2009-2014.pdf>.
- Ethiopian Ministry of Health. (2010). Health Sector Development Programme IV: 2010/11 - 2014/15. <http://www.nationalplanningcycles.org/sites/default/files/country_docs/Ethiopia/ethiopia_hsdp_iv_final_draft_2010_-2015.pdf>.

- Ethiopian Ministry of Health & MEASURE Evaluation. (2012). USAID HMIS Scale-up Project: Ethiopia. USAID.
- Goldman, I., et al. (2012). Establishing a National M&E System in South Africa. World Bank, PREMnotes September 2012 (21).
- Görgens, M., et al. (2005). Malawi - Developing a National Multisector HIV/AIDS Monitoring and Evaluation System. World Bank Global HIV/AIDS Program.
- Government of Kenya. (2011). National M&E Framework: HIV/AIDS.
<http://www.nacc.or.ke/index.php?view=article&catid=101%3Anational-mae-framework&id=171%3Amae&format=pdf&option=com_content&Itemid=138>.
- Guzman, M., Irarrazaval, I., & de los Rios, B. (2014). Monitoring and Evaluation System: the Case of Chile 1990-2014. *World Bank ECD Working Paper Series* (29).
- Holvoet, N., & Inberg, L. (2010). Sector monitoring and evaluation systems in the context of changing aid modalities: the case of Uganda's health sector. Universiteit Antwerpen, Institute of Development Policy and Management (IOB)
- Holvoet, N., & Inberg, L. (2013). Taking stock of monitoring and evaluation systems in the health sector: Findings from Rwanda and Uganda. *Health Policy and Planning* 29(4), 506-516.
- Holvoet, N., Gildemyn, M., & Inberg, L. (2012). Taking Stock of Monitoring and Evaluation Arrangements in the Context of Poverty Reduction Strategy Papers: Evidence from 20 Aid-Dependent Countries in Sub-Saharan Africa. *Development Policy Review*, 30(6), 749-772.
- IMF. (2010). *Uganda: Poverty Reduction Strategy Paper*. Washington, D.C.: International Monetary Fund.
- IMF. (2011a). *The Federal Democratic Republic of Ethiopia: Poverty Reduction Strategy Paper: Growth and Transformation Plan 2010/ 11-2014/ 15 - Volume I*. Washington, D.C.: International Monetary Fund.
- IMF. (2012). *Burkina Faso: Strategy for Accelerated Growth and Sustainable Development 2011 - 2015*. Washington, D.C.: International Monetary Fund.
- Independent Evaluation Group. (2007). A Diagnosis of Colombia's National M&E System, SINERGIA. *World Bank ECD Working Paper Series* (17).
- Kanyarukiga, S. G. (2004). Strategic Plan of Agriculture Transformation: Financing, Coordination and Monitoring and Evaluation in the Agricultural Sector.
<<http://www.ifad.org/english/operations/pf/rwa/i671rw/web/theme/coordination.pdf>>.
- Karani, F. N., Bichanga, W. O., & Kamau, C. G. (2014). Effective Use of Monitoring and Evaluation Systems in Managing HIV/AIDS Related Projects: A Case Study of Local NGOS in Kenya. *Science Journal of Business and Management*, 2(2): 67-76.
- Kawonga, M., Blaauw, D., & Fonn, S. (2012). Aligning vertical interventions to health systems: a case study of the HIV monitoring and evaluation system in South Africa. *Health Research Policy and Systems*, 10:2.
- Kenyan National AIDS Control Council. (2005). National HIV/AIDS Monitoring and Evaluation Framework.
<http://gametlibrary.worldbank.org/FILES/810_HIV%20M&E%20Framework%20NACC%20Kenya.pdf>.
- Lowrance, D., et al. (2007). Assessment of a national monitoring and evaluation system for rapid expansion of antiretroviral treatment in Malawi. *Tropical Medicine and international Health*, 12(3).
- Malawi National AIDS Commission. (2009). Malawi HIV and AIDS Extended National Action Framework (NAF), 2010-2012: DRAFT. <http://www.aidstar-one.com/sites/default/files/prevention/resources/national_strategic_plans/Malawi_10-12.pdf>.

- Malawi National Statistical System. (2013). Malawi 2013-2017 National Statistical System Strategic Plan. <<http://www.nsomalawi.mw/images/stories/NSSACTIVITIES/NSS%20Strategic%20Plan%202013-2017/National%20Statistical%20System%20Strategic%20Plan%202013-2017.pdf>>.
- MEASURE Evaluation. (2014a). A Case Study to Measure National HIV M&E System Strengthening: Nigeria. USAID.
- Meek, A. & Samanyika, Y. (2013). Development of a National Monitoring and Evaluation Framework for WASH Activities in Malawi. Monitoring Sustainable WASH Service Delivery Symposium, April 9-11, 2013.
- Murray-Zmijewski, A., & Gasana, C. (2010). *Monitoring and Evaluation in Rwanda Country Case Study Final Report*. London: Overseas Development Institute.
- National Planning Commission of Nepal. (2013). National Monitoring and Evaluation Guidelines. <http://www.npc.gov.np/web/new/uploadedFiles/allFiles/M&E_GuidelineEng.pdf>.
- Nepal Ministry of Health and Population. (2012). M&E Framework: Nepal Health Sector Programme II: 2010 - 2015. <http://www.internationalhealthpartnership.net/fileadmin/uploads/ihp/Documents/Country_Pages/Nepal/ME-framework-NHSP-II-Final-08-May-2012.pdf>.
- Nepal Ministry of Health and Population. (2013). National Guidelines on Monitoring and Evaluation of HIV Response in Nepal, 2012. <http://hivhealthclearinghouse.unesco.org/sites/default/files/resources/National_MandE_Guidelines_HIV_Response_Nepal_2012_Updated_Nov2013.pdf>.
- Newfarmer, R., Michele, Z. S., & Vijili, M. (2013). Managing Aid for Trade and Development Results: Rwanda Case Study. OECD Policy Dialogue on Aid for Trade.
- OECD. (2011). Aid Effectiveness 2011: Progress in Implementing the Paris Declaration. *OECD Publishing*.
- OECD & UNDP. (2014). Making Development Co operation More Effective 2014. <http://doi.org/10.1787/9789264209305-en>
- Paul, E., & Dossouvi, C. Y. D. (2011). Documenting the results of implementing aid effectiveness principles in the health sector: Benin case study. International Health Partnership.
- Peersman, G., et al. (2009). Are the investments in national HIV monitoring and evaluation systems paying off? *Journal of Acquired Immune Deficiency Syndromes* 52(2).
- Phiri, M. A. R. (2014). Capacity Strengthening Strategy through Capacity Needs Assessment for Country Level Strategic Analysis and Knowledge Support System (SAKSS). ReSAKSS CNA Report #1.
- Porter, S., & Goldman, I. (2013). A Growing Demand for Monitoring and Evaluation in Africa. *African Evaluation Journal*, 1(1).
- Rwanda Biomedical Center. (2013). Rwanda HIV and AIDS National Strategic Plan: June 2013 - June 2018. <http://www.nationalplanningcycles.org/sites/default/files/country_docs/Rwanda/final_nsp_2013-2018.pdf>.
- Segone, M. (2010). From policies to results: Developing capacities for country monitoring and evaluation systems. UNICEF Evaluation Working Papers.
- Senadza, B., & Laryea, A. D. A. (2012). Managing Aid for Trade And Development Results: Ghana Case Study. OECD Policy Dialogue on Aid for Trade.
- Shorten, T., et al. (2012). The International Health Partnership Plus: rhetoric or real change? Results of a self-reported survey in the context of the 4th high level forum on aid effectiveness in Busan. *Globalization and Health*, 8(13). Retrieved from <http://www.epigeneticsandchromatin.com/content/8/1/13>

Tanzania Commission for AIDS. (2011). Tanzania National Multisectoral HIV and AIDS Monitoring and Evaluation Plan: 2010-2012. <http://www.tacaids.go.tz/index.php?option=com_docman&task=doc_download&gid=71&Itemid=142>.

Ugandan Ministry of Health. (2011). Monitoring & Evaluation Plan for Health Sector Strategic & Investment Plan 2010/11-2014/15. <http://health.go.ug/docs/HSSIP_2010_15.pdf>.

WHO. (2011). Country Accountability Framework: Assessment, Mali. World Health Organization.

WHO. (2012). Country Accountability Framework: Assessment, Malawi. World Health Organization.

References Not Cited in Text

- Agence Francaise de Development. (2008). Aid effectiveness in the health sector in Mozambique: challenges to be met. <http://www.afd.fr/webdav/shared/ELEMENTS_COMMUNS/AFD/Efficacite%20Aide/Mozambique%20-%20Health%20.pdf>.
- Anie, S. J., & Larbi, E. T. (2004). Planning and Implementing a National Monitoring and Evaluation System in Ghana: A Participatory and Decentralized Approach Collaborative Process for Developing a National M&E Plan. *NEW DIRECTIONS FOR EVALUATION* (103).
- Atun, R., et al. (2011). Critical Interactions Between the Global Fund-Supported HIV Programs and the Health System in Ghana. *Journal of Acquired Immune Deficiency Syndromes* (1999).
- Bedi, T. et al. (2006). *Beyond the Numbers: Understanding the Institutions for Monitoring Poverty Reduction Strategies - Chapter 12: Uganda*. Washington, D.C.: The World Bank.
- Bernardi, R., & De Chiara, F. (2011). ICTs and monitoring of MDGs : a case study of Kenya HIV/AIDS monitoring and evaluation in a donor multi-agency context. *Information Technology for Development*, 17(1), 24-41.
- CoE-Nepal. (2012). An Effort in the Process of Promoting Quality Evaluation: A Case of CoE-Nepal. <http://www.ioce.net/download/national/Nepal_CommunityofEvaluators%28CoE-Nepal%29_CaseStudyCleaned.pdf>.
- Dery, B. (2013). Building M&E Capacities to Enhance the National M&E System in Ghana - The Way Forward. 3rd International Conference on National Evaluation Capacities (NEC).
- Development Gateway. (2007). Aid Management Platform in Malawi: In Brief. <http://www.developmentgateway.org/dg_uploads/pdfs/amp_malawi_eng_sept_2011.pdf>.
- Ethiopian Central Statistical Agency. (2009). National Statistical Development Strategy: 2009/10 - 2013/14. <<http://paris21.org/sites/default/files/Ethiopia-NSDS-2009-14.pdf>>.
- Ethiopian HIV/AIDS Prevention and Control Office. (2003). National Monitoring and Evaluation Framework for the Multi-Sectorial Response to HIV/AIDS in Ethiopia. <http://hivhealthclearinghouse.unesco.org/sites/default/files/resources/iiep_ethiopianationalmande2003.pdf>.
- Gasore, B., & Kirui, M. (n.d.). A Work Well Done and a Happy Ending: Recounting gains and milestones in strengthening national statistical information systems and strategic plans. <http://www.futuresgroup.com/files/documents/FUTURES_GROUP_CLOUSEOUT_SUPPLEMENT_1.pdf>.
- Ghana National Development Planning Commission. (2009). Guidelines for the Preparation of The District Monitoring and Evaluation Plan. <http://www.amcod.info/en/publications/monitoring/ghana/DISTRICT_MandE_GUIDELINES_2010_2013.pdf>.
- Global AIDS Monitoring & Evaluation Team. (2007). Decentralising HIV M&E in Africa. <<http://siteresources.worldbank.org/INTHIVAIDS/Resources/375798-1132695455908/GRDecentralisationAug2207.pdf>>.
- Gorgens, M., et al. (n.d.). Malawi: A National HIV/AIDS Monitoring & Evaluation System. MfDR Principles in Action: Sourcebook on Emerging Good Practices.
- Hauge, A. (2001). Strengthening Capacity for Monitoring and Evaluation In Uganda: A Results Based Management Perspective. *World Bank ECD Working Paper Series* (8).
- Holvoet, N., & Rombouts, H. (2008). *The Denial of Politics in PRSP's Monitoring and Evaluation Experiences From Rwanda*. IOB, University of Antwerp.

- IMF. (2011b). *Tanzania: Poverty Reduction Strategy Paper*. Washington, D.C.: International Monetary Fund.
- IMF. (2013). *Mali: Poverty Reduction Strategy Paper*. Washington, D.C.: International Monetary Fund.
- Kenyan National AIDS Control Council. (2011). NATIONAL MONITORING AND EVALUATION (M&E) GUIDELINES AND STANDARD OPERATING PROCEDURES : Pillar 1.
<<http://nascop.or.ke/library/3d/National%20M&E%20Guidelines%20and%20SOPs.pdf>>.
- Kiberu, V. M., et al. (2014). Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. *BMC Medical Informatics and Decision Making*, 14(40).
- Kiregyera, B. P., et al. (2002). Poverty Monitoring System in Malawi: Paris21 Consultants' Report.
<www.paris21.org/sites/default/files/711.doc>.
- Lipenga, K. D. (2012). Malawi 2012 Growth and Development Strategy.
<<http://www.gafspfund.org/sites/gafspfund.org/files/MGDS%20II%20final%20document%20january%202012.pdf>>.
- Malawi Ministry of Health. (2011). Malawi Health Sector Strategic Plan 2011-2016.
<<http://www.medcol.mw/commhealth/publications/3%20Malawi%20HSSP%20Final%20Document%20%283%29.pdf>>.
- Malawi National AIDS Commission. (2003). *National HIV/AIDS Monitoring and Evaluation Plan*.
<http://siteresources.worldbank.org/INTHIVAIDS/Resources/375798-1132695455908/M&E-GR-1_MalawiPlanPartB_July03.pdf>.
- Malawi National Statistical System. (2008). *Republic of Malawi National Statistical System Strategic Plan*.
<<http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=2548>>.
- May, E., et al. (2006). *Towards the Institutionalization of Monitoring and Evaluation Systems in Latin America and the Caribbean*. Washington, D.C.: The World Bank.
- Mbondo, M., et al. (2013). Organizational HIV monitoring and evaluation capacity rapid needs assessment: the case of Kenya. *The Pan African Medical Journal*, 14l.
- MEASURE Evaluation. (2010). Report on the Status of the Nigerian National HIV Monitoring and Evaluation System. USAID.
- MEASURE Evaluation. (2014b). A Case Study to Measure National HIV M&E System Strengthening: Côte d'Ivoire. USAID.
- Mehrotra, S. (2013). The Government Monitoring and Evaluation System in India: A Work in Progress. *World Bank ECD Working Paper Series* (28).
- Millennium Challenge Account - Burkina Faso. (2013). Burkina Faso Monitoring and Evaluation Plan. Millennium Challenge Corporation.
- National Planning Commission of Nepal. (2004). Poverty Monitoring And Analysis System (Pmas) Framework Document.
<http://www.academia.edu/1415153/National_Poverty_Monitoring_and_Assessment_System_in_Nepal>.
- Nigeria National Agency for the Control of AIDS. (2011). The National HIV and AIDS Monitoring and Evaluation Plan: The Nigeria National Response Information Management System (NNRIMS) Operational Plan II.
<http://hivhealthclearinghouse.unesco.org/sites/default/files/resources/NNRIMS_Operational_Plan_2011-2016.pdf>.
- OECD. (2007). 2006 Survey on Monitoring the Paris Declaration. *OECD Journal on Development* 8(2).
- OECD. (2008). 2008 Survey on Monitoring the Paris Declaration: Effective Aid by 2010 - What it will take?. 3rd High Level Forum on Aid Effectiveness, September 2-4, 2008, Accra, Ghana.

- Phillips, S., et al. (2014). A focus on M&E of results: an example from the Presidency, South Africa. *Journal of Development Effectiveness*, 64, 392-406.
- Rwandan Ministry of Agriculture and Animal Resources. (2009). Strategic Plan for the Transformation of Agriculture in Rwanda - Phase II (PSTA II).
<https://www.gafspfund.org/sites/gafspfund.org/files/Documents/Rwanda_StrategicPlan.pdf>.
- Rwandan Ministry of Finance and Economic Planning. (2012). Strengthening Government of Rwanda's National Evidence-Based Policy Planning, Analysis, and M&E.
<<http://www.undp.org/content/dam/undp/documents/projects/RWA/75361%20-%20M%20and%20E.pdf>>.
- Rwandan National AIDS Control Commission. (2006). Rwanda National HIV and AIDS Monitoring and Evaluation Plan.
<http://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---ilo_aids/documents/legaldocument/wcms_172588.pdf>.
- Samjong KPMG ERI Inc. (2012). Strengthening the Capacity of the Monitoring and Evaluation (M&E) System of Ghana at All Levels of National Administration. <https://www.kdevelopedia.org/Resources/government-law/strengthening-capacity-monitoring-evaluation-me-system-ghanat-all-levels-national-administration--04201210090121999.do?fldlds=TP_GOV%20TP_GOV_PA>.
- Sinha, P. K. (n.d.) Can a Useful Aid Effectiveness Index Be Developed using the Paris Declaration Framework?
<http://s3.amazonaws.com/aiddata/Sinha_aiddata.pdf>.
- Tanzania Commission for AIDS. (2004). National Monitoring and Evaluation Framework for HIV/AIDS.
<http://gametlibrary.worldbank.org/FILES/1282_M&E%20Framework-Tanzania.pdf>.
- Tanzania Commission for AIDS. (2008). National HIV and AIDS Response: M&E Indicators Status.
<www.tzdpd.or.tz/.../dpg.../M_and_E_Indicators_Status.ppt>.
- Tanzanian Ministry of Finance. (2011). MKUKUTA II Monitoring Master Plan.
<<http://www.mof.go.tz/mofdocs/DPE/MKUKUTA%20II%20MONITORING%20MASTER%20PLAN.pdf>>.
- Tanzanian National Bureau of Statistics. (2010). Tanzania Statistical Master Plan 2009/10 - 2013/14.
<http://www.nbs.go.tz/takwimu/tsmp/TSMP_Doc_Aug2010.pdf>.
- Ugandan AIDS Commission. (2011). National HIV and AIDS Monitoring and Evaluation Plan for the National Strategic Plan for HIV and AIDS in Uganda.
<http://hivhealthclearinghouse.unesco.org/sites/default/files/resources/Uganda_M_E_Plan2011_2015.pdf>.
- Ugandan Office of the Prime Minister. (2011). National Policy on Public Sector Monitoring and Evaluation.
<http://usaidlearninglab.org/sites/default/files/resource/files/Attachment_J.15-_M%26E_Policy_Final_Draft.pdf>.
- United Nations Economic and Social Council. (2012). Mutual Accountability For Development Cooperation Results: Where Next? <[http://www.un.org/en/ecosoc/newfunct/pdf/dcf_mutual_accountability_busan_study\(29jun\).pdf](http://www.un.org/en/ecosoc/newfunct/pdf/dcf_mutual_accountability_busan_study(29jun).pdf)>.
- Vietnamese Ministry of Health. (2007). National Monitoring and Evaluation Framework for HIV Prevention and Control Programs.
<<https://www.globalhivmeinfo.org/CapacityBuilding/Documents/Vietnam%20National%20MandE%20Framework%20for%20HIV%20Prevention%20and%20Control%20Programs%20%28English%29.pdf>>.
- Wachira, C., & Ruger, J. P. (2011). National poverty reduction strategies and HIV/AIDS governance in Malawi: A preliminary study of shared health governance. *Social Science and Medicine*, 72(12).
- WHO. (2009). Strengthening monitoring and evaluation practices in the context of scaling-up - Kenya. The World Health Organization.

World Bank. (2000). Monitoring and Evaluation Capacities in Ghana - A Diagnosis and Proposed Action Plan. The World Bank.

World Bank. (2007). Results-Based National Development: Strategies Assessment and Challenges Ahead.

<<http://siteresources.worldbank.org/CDFINTRANET/Overview/21790579/Results-basedNationalDevelopmentStrategies-December122007.pdf>>.

Appendix A. Literature Search Process

Our search aimed to identify national government M&E systems in developing countries, such as those tracking progress on Poverty Reduction Strategies or in the HIV/AIDS sector. We conducted one general Google search for government results monitoring & evaluation (M&E) systems¹⁶. This search captured multiple types of documents referring to government systems designed to monitor and evaluate results in domains relevant to international aid flows. We also performed individual Google searches for specific countries' government M&E systems¹⁷, targeting the following Foundation focus geographies: Bangladesh, Burkina Faso, Ethiopia, Ghana, India - Bihar, India - Odisha, Kenya, Malawi, Mali, Nepal, Nigeria, Rwanda, Tanzania, and Uganda.

We elected to use Google rather than academic databases for our searches with the understanding that much of the literature we were looking for would be grey, or unpublished, including reports and evaluations of governments, donors, and non-profit organizations. Google generally returns a large number of search results, and we usually reviewed the first 50 search results, after which results were generally decreasingly relevant. We used Google results as a starting point for additional searches, looking for relevant documents on the sites that came up in the results lists, especially when these were government sites.

We conducted supplemental searches to ensure that we captured prominent academic literature. We searched two relevant academic databases: PAIS International and Scopus¹⁸. On PAIS International, we reviewed all 77 results for the abbreviated search string and both results for the longer string. On Scopus, we reviewed the first 300¹⁹ of the 480 results for the abbreviated search string and the only result for the longer string. Few of the search results from these databases were relevant, and of those that were, most were HIV/AIDS-focused.

We screened the titles and summaries of the results for all of our searches and retrieved all documents that met the following screening criteria:

- Describes an M&E system that is administered by a national government or ministry;
- Describes either a national M&E system or an overall health/agricultural sector M&E system;
- Describes framework OR monitoring/evaluation results for a specific M&E system; and
- Does not describe general principles or guidelines for M&E that are not applied to a specific M&E system.

In some cases evaluations or reports of government M&E systems were not available and we elected to include national M&E frameworks and strategic plans in our body of evidence. These documents provide insight into demand for M&E, challenges, and planned M&E activities, although they could not be treated as evidence that these activities were ever carried out. As the scope of this request is country-level, we excluded documents describing sub-national M&E systems (e.g. region-, program-, and project-level M&E systems) and M&E systems administered by non-government development partners. We included documents describing national-level M&E systems that were specific to the health or agricultural sector, as some countries lacking comprehensive national M&E systems have been supported by development partners to establish systems in these sectors. Finally, we also included documents describing country M&E systems not on our list of target countries, and documents that evaluated multiple countries.

Table A.1 summarizes the results of our country-level searches. We retrieved 138 documents relating to country-level M&E systems that met our screening criteria.

¹⁶ We used the following search string: ("monitoring and evaluation" OR "performance management" OR "performance monitoring" OR "results management") AND national AND (system OR framework) AND ("economic development" OR "international development" OR aid) AND "case study" We chose to append "case study" to the search string as the most relevant results from our initial exploratory searches were framed as case studies. However, we did not find many "case studies" of government M&E systems, so the documents we retrieved include a variety of reports and evaluations of these systems, often commissioned by national governments or donor organizations.

¹⁷ To conduct these searches, we found that using an abbreviated search string, (("monitoring and evaluation" OR "M&E") AND national AND (system OR framework)) AND "Country Name", yielded more relevant results than the longer, more general search string.

¹⁸ On both of these databases, we conducted searches using both of the search strings specified above, but omitting any specific country names.

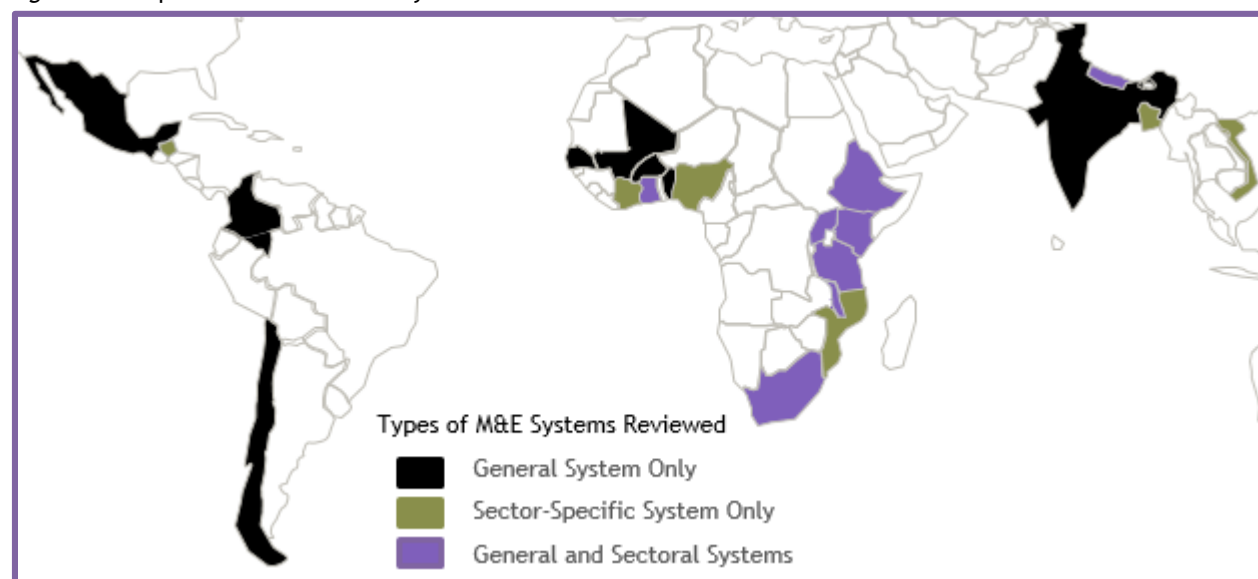
¹⁹ Results after 300 no longer appeared to be relevant.

Table A.1 Summary of Country-Level Searches

Country	Number of Results	Results Reviewed	Documents Retrieved
General	7,810,000	150	50
Bangladesh	379,000	50	1
Burkina Faso	353,000	50	4
Ethiopia	387,000	50	8
Ghana	422,000	50	8
India	21,800,000	100 ²⁰	1
India - Bihar OR Odisha	490,000	50	0
Kenya	427,000	50	6
Malawi	386,000	50	14
Mali	348,000	50	3
Nepal	390,000	50	8
Nigeria	395,000	50	3
Rwanda	370,000	50	9
Tanzania	400,000	50	6
Uganda	405,000	50	6
PAIS International	79	79	2
Scopus	401	301	9
Total		1280	138

13 of the 138 documents retrieved were duplicates leaving 125 unique documents. We then reviewed these documents according to the screening criteria a second time, eliminating 26. Documents were excluded if they described program- or project-level M&E or results measurement, outlined general guidelines that did not apply to a specific M&E system, or described M&E systems in sectors other than health or agriculture. After this stage of screening, our body of evidence consisted of 99 relevant documents. An additional nine relevant documents reporting on multiple country M&E systems were identified during a separate search of donor-level results measurement systems. These nine documents were added to the country-level body of evidence, giving us a total of 108 relevant documents. *Figure A.1* is a map illustrating the countries and types of M&E systems covered in this review.

Figure A.1 Map of Government M&E Systems Reviewed

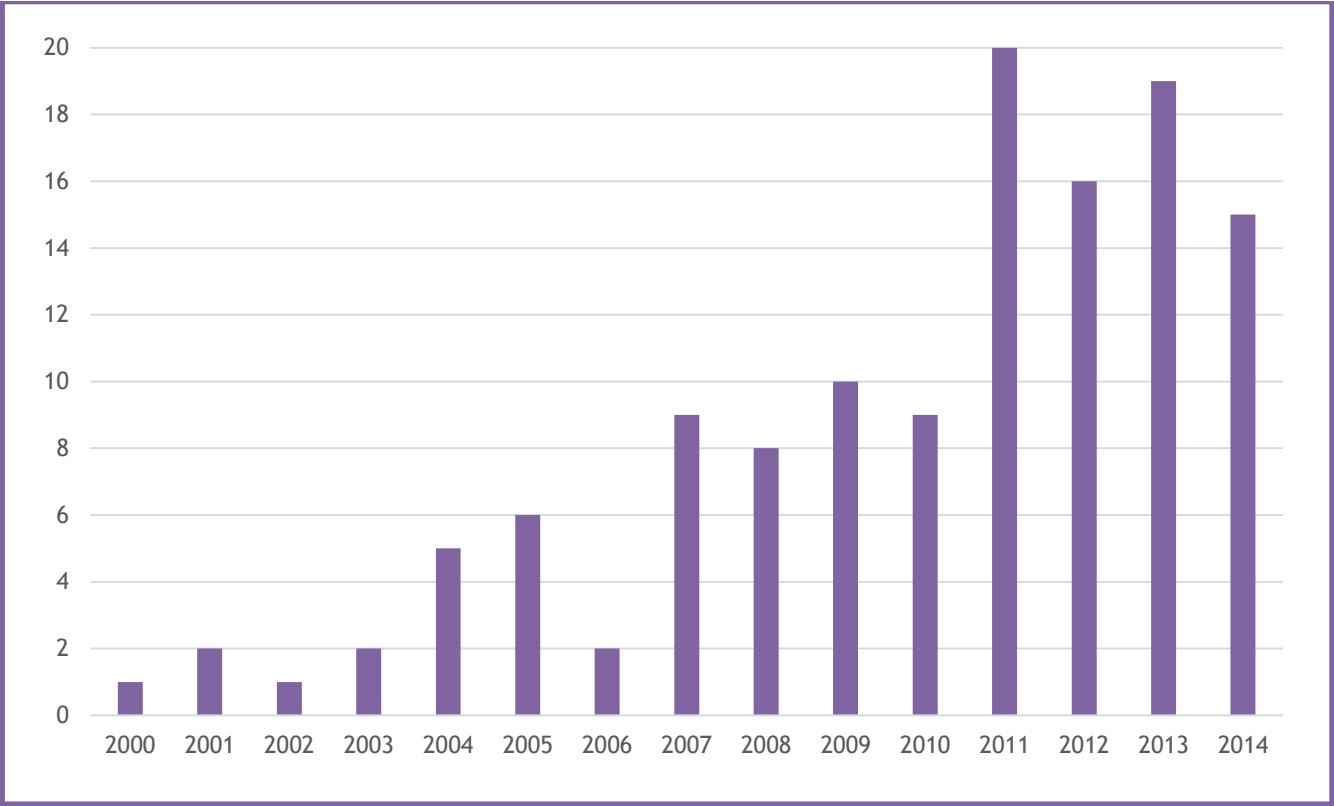


The year of publication of the documents identified as relevant for this review ranges from 2000 to 2014, as illustrated in *Figure A.2*. The majority of the evidence we review dates from the last four years (2011-2014), but for many M&E systems, no recent documentation was available. Our coding and findings are based on the most recent relevant information across

²⁰ We reviewed the first 100 results for India in order to try and find additional relevant documents, but did not identify any.

systems, though in some cases older documents are referenced when they contained the best information. Appendix B presents our level of confidence in the evidence for particular systems, which is affected by how recent the evidence is, including the year of the most recent M&E plan we were able to access and review, and the year of the most recent M&E report or case study. The review spreadsheet that accompanies this report includes the publication dates for all the evidence reviewed for each system.

Figure A.2. Number of Documents Reviewed by Year of Publication



Appendix B. Summary of Body of Evidence

Table B.1 summarizes the body of evidence for each of the 42 government M&E systems we reviewed. The table outlines the year each M&E system was initiated²¹ and the number of documents specific to that particular M&E system reviewed, plus documents that include that system as part of a larger review of government M&E systems. For many countries, we review specific M&E plans (beyond just M&E components of strategic plans) outlining proposed M&E arrangements and activities. *Table B.1* notes the year of the most recent M&E plan reviewed for each M&E system, when one was available²². The table also indicates whether the documents include reports on M&E implementation, as opposed to just M&E plans. In some cases, the only information on M&E implementation is in multi-country reviews that do not provide as much detail on individual government M&E systems. For each M&E system, we note the year of the most recent M&E report or case study that specifically reviews that system.

We also rate the strength of the body of evidence for each M&E system based on the relevance and quality of the information included in the documents. A “low” rating indicates that fewer than two documents are available and that these provide limited detail on recent M&E planning or implementation. A “medium” rating indicates that two to six documents are available, but there are gaps in the information provided. These gaps include lacking either a specific M&E plan or specific M&E evaluation. A “high” rating indicates that at least three documents are available that also provide a high level of relevant and recent information on M&E planning and implementation.

Table B.1 Summary of Body of Evidence for Each Government M&E System

Country	Sector	Year M&E System Initiated	Number of Documents Reviewed	Year of Most Recent M&E Plan	Do Documents Report on M&E Implementation?	Year of Most Recent M&E Report or Case Study	Strength of Evidence
Bangladesh	HIV/AIDS	2007	1	2007	Yes	N/A	Medium-Low
Belize	HIV/AIDS	2008	1	2008	No	N/A	Low
Benin	General	2009	7	N/A	Yes	2012	Medium
Burkina Faso	General	2011	6	N/A	Yes	N/A	Medium
Chile	General	2010	2	N/A	Yes	2014	Medium
Colombia	General	2005	3	2005	Yes	2007	Medium-High
Cote d'Ivoire	HIV/AIDS	2010	1	N/A	Yes	2014	Medium
Ethiopia	General	1996	8	N/A	Yes	2009	Medium-High
Ethiopia	Health	2007	5	N/A	Yes	2014	Medium-High
Ethiopia	HIV/AIDS	2003	3	2003	Yes	N/A	Medium
Ghana	General	2009	11	2009	Yes	2013	High
Ghana	Agriculture	2010	1	N/A	Yes	N/A	Low
Ghana	HIV/AIDS	2001	2	N/A	Yes	2011	Medium-Low
India	General	Not specified	1	N/A	Yes	2013	Low
Kenya	General	2008	6	N/A	Yes	2013	Medium-High
Kenya	Health	Not specified	2	N/A	Yes	2010	Medium-Low
Kenya	HIV/AIDS	2009	6	2011	Yes	2014	High
Malawi	General	2012	11	N/A	Yes	2013	High
Malawi	Agriculture	2008	1	N/A	Yes	2014	Medium-Low
Malawi	Health	2011	3	N/A	Yes	2012	Medium-Low
Malawi	HIV/AIDS	2010	7	2003	Yes	2010	High

²¹ Either the first year that a specific M&E system is described in a strategic plan, or the year specified in an M&E review or plan.

²² In our review framework, M&E activities documented in M&E or strategic plans that could not be confirmed in other documents are classified as “planned.”

Malawi	WASH	2011	1	N/A	Yes	2013	Low
Mali	General	2003	8	N/A	Yes	2009	Medium-High
Mexico	General	2004	2	N/A	Yes	2009	Medium
Mozambique	Health	2006	3	N/A	Yes	N/A	Medium-Low
Nepal	General	2013	6	2013	Yes	N/A	Medium
Nepal	Health	2012	2	2012	Yes	N/A	Medium-Low
Nepal	HIV/AIDS	2012	1	2013	No	N/A	Low
Nigeria	HIV/AIDS	2007	3	2011	Yes	2014	High
Rwanda	General	2002	9	N/A	Yes	2012	High
Rwanda	Agriculture	2009	2	N/A	No	N/A	Low
Rwanda	Health	2009	2	N/A	Yes	2013	Medium-Low
Rwanda	HIV/AIDS	2013	3	2009	No	N/A	Medium-Low
Senegal	General	2003	6	N/A	Yes	2012	Medium-High
South Africa	General	2007	7	N/A	Yes	2014	Medium-High
South Africa	HIV/AIDS	2007	1	N/A	Yes	2012	Medium-Low
Tanzania	General	2011	7	2011	Yes	N/A	Medium
Tanzania	HIV/AIDS	2011	3	2011	Yes	2008	Medium-High
Uganda	General	2011	7	2011	Yes	2012	High
Uganda	Health	2011	4	2010	Yes	2014	Medium-High
Uganda	HIV/AIDS	2011	2	2011	Yes	N/A	Medium-Low
Vietnam	HIV/AIDS	2007	1	2007	Yes	N/A	Low

Appendix C. Review Framework Questions

- Basic Descriptive Questions
 - Number of documents reviewed
 - Do the documents include a specific M&E plan (not just a strategic plan with an M&E component)?
 - Year of the most recent specific M&E plan
 - Do the documents report specifically on M&E implementation (not just reporting on overall strategy)?
 - Do the documents include specific reports on M&E implementation in this country (not just evaluations of M&E across multiple countries)?
 - Year of the most recent specific M&E report, case study, or evaluation
 - Strength of body of evidence
 - In what year was the latest M&E system initiated?
 - Is the M&E system a general national system, or a sector-specific system?
 - Which sector does the M&E system cover?
- Demand for M&E
 - What is the level of internal demand for M&E? (low to high)
 - What government office(s) drives M&E demand (e.g. Ministry of Planning; Line Ministry; Parliament; Office of President, etc.)?
 - Is national civil society involved in demanding M&E information?
 - Is the domestic private sector involved in demanding M&E information?
 - What is the level of internal demand for evaluation and use of monitoring information? (low to high)
 - Is there external demand for M&E?
 - Which organizations within the international donor community are demanding M&E information?
 - Which international civil society organizations are demanding M&E information?
 - What is the level of harmonization of donor- and country-led M&E demands (e.g. combined systems, agreements over M&E frameworks, etc.)? (low to high)
 - Are there challenges with or recommendations for increasing demand for M&E information?
 - Are there challenges with or recommendations for budgeting for M&E activities?
- Institutional Design
 - Are the roles and responsibilities for M&E clearly defined (e.g. which offices have which roles)? (low to high)
 - Who/what office is primarily responsible for coordinating M&E?
 - Is there a national statistics or M&E office involved with M&E?
 - How is the national statistics office involved in M&E (coordination support, data collection, surveys, evaluation, other)?
 - Does the system distinguish between monitoring and evaluation?
 - Who/what office is primarily responsible for overseeing monitoring?
 - Who/what office is primarily responsible for budget monitoring?
 - Who/what office is primarily responsible for evaluation?
 - Who/what office is primarily responsible for reporting?
 - Who/what office is primarily responsible for planning?
 - Is there an institutional framework for combining and coordinating these functions?
 - What is the level of harmonization between these activities? (low to high)
 - Are there challenges with or recommendations for coordination of the different components of M&E?
 - Do coordination challenges include lack of clarity or mandate for coordination body?
 - Do coordination challenges include lack of resources for M&E coordination?
 - Do coordination challenges include lack of capacity for M&E coordination?
 - Do donors provide technical cooperation/support for M&E?
 - Are there challenges with or recommendations for the participation of non-government actors?
- Strategic Framework
 - Is there a unified strategic framework (e.g. Poverty Reduction Strategy Paper)?
 - Is the strategic framework expressed as a theory of change, results chain, or logframe?
 - Is the strategic framework aligned with external frameworks (e.g. Millennium Development Goals)?

- Are there challenges with or recommendations for harmonization of M&E or strategic frameworks across government offices or agencies?
 - Is there a list of indicators based on the strategic framework?
 - Who/what office is primarily responsible for selecting indicators and measures?
 - Are there challenges with or recommendations for choosing between different indicators/measures?
 - Are outputs distinguished from outcomes?
 - Are there separate short-term/intermediate/long-term outcomes?
 - Are certain indicators/outcomes prioritized over others?
 - Are routine monitoring indicators aligned with performance indicators?
 - Does the strategy include capacity building?
 - Do results/performance evaluations influence strategy, planning, or program management? (low to high)
- Budgeting
 - Are budget allocations connected to results data and evaluations?
 - Does monitoring and reporting include disaggregated budget figures?
 - Does monitoring and reporting include disaggregated expenditure figures?
 - Are aid flows recorded?
 - Are donors and funding specified in the budget/expenditures for different activities/outputs/outcomes?
- Monitoring and Data Collection
 - Is there a framework for monitoring results based on the strategic framework?
 - Are there documented rules, standards, or processes for data collection?
 - Are there documented rules, standards, or processes for data aggregation (e.g. across sectors or projects)?
 - Who is responsible for data collection (e.g. line ministries, local government offices, service providers, etc.)?
 - Are there strong links with line ministries/sectors?
 - Is monitoring and data collection decentralized?
 - What is the level of data collection coverage at local level (i.e. are there certain populations, areas, or sectors where data collection is not happening)?
 - Is there a functioning IT M&E system (e.g. data collection, aggregation, and verification software tools are incorporated into the monitoring system)?
 - Is data stored and managed electronically?
 - Is data collected through routine information systems?
 - Are surveys used to collect data?
 - Is data collected in collaboration with NGOs and/or civil society, or is data from these sources used in M&E?
 - Who/what office is primarily responsible for data aggregation and management?
 - Is there a process for data verification?
 - Are there challenges with or recommendations for ensuring data quality?
 - Is data open and public?
 - Are there challenges with or recommendations for the capacity for monitoring?
- Analysis
 - Is there a system for evaluation of government results/performance (At minimum, this must include tracking and comparing outcomes over time)?
 - Does the evaluation system involve comparing outcomes against targets?
 - Is a monitorable performance assessment framework used for evaluation?
 - Is the national statistics agency involved in evaluation?
 - Is there external funding for evaluations (e.g. from development partners, donors, civil society)?
 - Are there economic/cost evaluations (i.e. that evaluate the cost of supporting positive change)?
 - Are there audit evaluations (i.e. that evaluate expenditures against the budget)?
 - How does the evaluation system deal with impact evaluation?
 - Does the government commission external evaluation of its results/performance?
 - Does the M&E system acknowledge or incorporate non-government evaluations?
 - Are there challenges with or recommendations for the capacity for evaluation?
- Reporting and Use of M&E Data
 - Are there documented rules for reporting on M&E information?
 - What is the frequency of principal national level M&E reporting?

- Is M&E information disseminated internally within the government?
- Is M&E information disseminated externally (e.g. to CSOs, donors, etc.)?
- Are there established processes for incorporating M&E information?
- Are there periodic joint/coordinated progress and performance reviews?
- Is there mutual accountability between countries and donors? (low to high)
- Are there challenges with or recommendations for the use of M&E information?
- Are there challenges with or recommendations for conflicting information demands or uses?

Appendix D. Questions Reported on in Each Section

Table D.1 Relevant Institutionalization and Coordination Questions from the Review Framework (Section 4.1)

Questions addressed in review framework	Also reported on in this section?
What is the level of internal demand for results-oriented M&E?	Yes
What government office(s) drives M&E demand (e.g. Ministry of Planning, Finance, Audit; Line Ministry; Parliament; etc.)?	Yes
Is national civil society involved in demanding M&E information?	Yes
Is the domestic private sector involved in demanding M&E information?	No
What is the level of internal demand for evaluation and use of monitoring information?	No
Are there challenges with or recommendations for increasing demand for M&E information?	Yes
Are there challenges with or recommendations for budgeting for M&E activities?	Yes
Are the roles and responsibilities for M&E clearly defined?	Yes
What office is primarily responsible for coordinating M&E?	Yes
What office is primarily responsible for overseeing monitoring?	Yes
What office is primarily responsible for budget monitoring?	Yes
What office is primarily responsible for evaluation?	Yes
What office is primarily responsible for reporting?	Yes
What office is primarily responsible for planning?	Yes
Is there a national statistics or M&E office involved with M&E?	Yes
How is the national statistics office involved in M&E (coordination support, data collection, surveys, evaluation, other)?	Yes
Does the system distinguish between monitoring and evaluation?	No
Is there an institutional framework for combining and coordinating these functions?	Yes
What is the level of harmonization between these activities?	Yes
Are there challenges with or recommendations for coordination of the different components of M&E?	Yes
Do coordination challenges include lack of clarity or mandate for coordination body?	Yes
Do coordination challenges include lack of resources for M&E coordination?	Yes
Do coordination challenges include lack of capacity for M&E coordination?	Yes
Are there challenges with or recommendations for harmonization of M&E or strategic frameworks across government offices or agencies?	Yes
Does the strategy include capacity building?	Yes
Is data open and public?	Yes

Table D.2 Relevant Data Collection Questions from the Review Framework (Section 4.2)

Questions addressed in review framework	Also reported on in this section?
Is there a list of indicators based on the strategic framework?	Yes
Who/what office is primarily responsible for selecting indicators and measures?	Yes
Are there challenges with or recommendations for choosing between different indicators/measures?	No
Are routine monitoring indicators aligned with performance indicators?	Yes
Is there a framework for monitoring results based on the strategic framework?	No
Are there documented rules, standards, or processes for data collection?	Yes
Are there documented rules, standards, or processes for data aggregation (e.g. across sectors or projects)?	Yes
Who/what office is responsible for data collection (e.g. line ministries, local government offices, service providers, etc.)?	Yes
Is monitoring and data collection decentralized?	No
What is the level of data collection coverage at local level (i.e. are there certain populations, areas, or sectors where data collection is not happening)?	Yes
Is there a functioning IT M&E system?	Yes
Is data stored and managed electronically?	Yes

Are data collected through local/routine M&E systems?	Yes
Are surveys used to collect data?	No
Are data collected in collaboration with NGOs and/or civil society, or are data from these sources used in M&E?	Yes
Who/what office is primarily responsible for data aggregation and management?	No
Is there a process for data verification?	Yes
Are there challenges with or recommendations for ensuring data quality?	Yes
Are there challenges with or recommendations for the capacity for monitoring?	Yes

Table D.3 Relevant Analysis and Use Questions from the Review Framework (Section 4.3)

Questions addressed in review framework	Also reported on in this section?
Is there a unified strategic framework (e.g. Poverty Reduction Strategy Paper)?	Yes
Is the strategic framework expressed as a theory of change, results chain, or logframe?	Yes
Are outputs distinguished from outcomes?	Yes
Are there separate short-term/intermediate/long-term outcomes?	No
Are certain indicators/outcomes prioritized over others?	No
Do results/performance evaluations influence strategy, planning, or program management?	Yes
Are budget allocations connected to results data and evaluations?	Yes
Does monitoring and reporting include disaggregated budget figures?	Yes
Does monitoring and reporting include disaggregated figures for expenditures?	Yes
Are aid flows recorded?	Yes
Are donors and funding specified in the budget/expenditures for different activities/outputs/outcomes?	Yes
Is there a system for evaluation of government results/performance?	Yes
Does the evaluation system involve comparing outcomes against targets?	Yes
Is a monitorable performance assessment framework used for evaluation?	Yes
Is the national statistics agency involved in evaluation?	No
Are there economic/cost evaluations (i.e. that evaluate the cost of supporting positive change)?	Yes
Are there audit evaluations (i.e. that evaluate expenditures against the budget)?	Yes
How does the evaluation system deal with impact evaluation?	Yes
Does the government commission external evaluation of its results/performance?	Yes
Are there challenges with or recommendations for the capacity for evaluation?	Yes
Are there documented rules for reporting on M&E information?	No
What is the frequency of principal national level M&E reporting?	Yes
Is M&E information disseminated internally within the government?	Yes
Are there established processes for incorporating M&E information?	Yes
Are there periodic joint/coordinated progress and performance reviews?	Yes
Are there challenges with or recommendations for the use of M&E information?	Yes
Are there challenges with or recommendations for conflicting information demands or uses?	Yes

Table D.4 Relevant Harmonization Questions from the Review Framework (Section 4.4)

Questions addressed in review framework	Also reported on in this section?
Is there external demand for results-oriented M&E?	No
Which organizations within the international donor community are demanding M&E information?	No
Which international civil society organizations are demanding M&E Information?	No
What is the level of harmonization of donor- and country-led M&E demands (e.g. combined systems, agreements over M&E frameworks, etc.)?	Yes
Do donors provide technical cooperation/support for M&E?	Yes
Are there challenges with or recommendations for the participation of non-government actors?	Yes
Does the strategic framework incorporate elements of external frameworks (e.g. Millennium Development Goals)?	Yes
Is there external funding for evaluations (e.g. from development partners, donors, civil society)?	Yes
Does the M&E system acknowledge or incorporate non-government evaluations?	Yes

Is M&E information disseminated externally (e.g. to CSOs, donors, etc.)?	Yes
Is there mutual accountability between countries and donors?	Yes