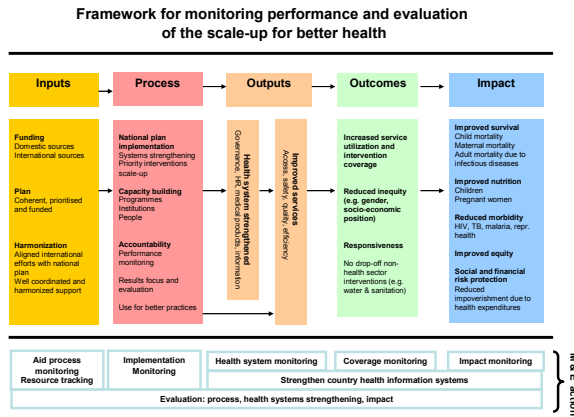


A common framework for monitoring performance and evaluation of the scale-up for better health

Monitoring & Evaluation Working Group International Health Partnership+ 1

February 2008



¹ This paper has been prepared by an IHP+ working group led by WHO and the World Bank. An earlier version was discussed at a meeting of technical experts in Geneva, December 11 and a meeting of countries, global health initiatives, bilateral donors, and international agencies in Geneva, 10-11 January 2008 and comments are taken into account. The principal authors of the paper are Ties Boerma, Ed Bos, Veronica Walford, Jennifer Bryce and Carla Abou-Zahr.

Summary

The scale-up to achieve the health MDGs is unprecedented in both potential resources and the number of involved initiatives. Such a grand experiment requires a harmonized monitoring and evaluation effort that reinforces both country and global needs to demonstrate results, secure future funding, and enhance the evidence base for intervention packages. Eventually, the scale-up efforts will be judged by country progress towards the health-related MDGs and related health goals, the degree to which major health constraints in countries have been addressed, and adherence to the Paris Declaration on Aid Effectiveness.

This paper provides an overview of the current situation status of country, regional and global efforts to monitor and evaluate progress toward the health MDGs, including country health information systems, monitoring performance, evaluation, and aid effectiveness. A common framework for evaluation of the scale-up for better health is proposed in line with the Paris Declaration. It is intended to provide a basis for developing an investment case for sound evaluations of cost-effectiveness and impact across the major initiatives in a way that it improves harmonization and alignment and strengthens country health information systems.

Individual global health initiatives differ in terms of intervention focus and expected results, but share the current emphasis on strengthen health systems as part of the scale-up. The overall framework can be used as a basis for the development of initiative-specific frameworks, focusing on specific pathways from increased resources to expected results. The final section described how the framework can be operationalized at the country level.

Contents

- 1 Background**
- 2 Current situation**
 - Health information systems
 - Monitoring and evaluation
 - Barriers to evaluation
 - Aid effectiveness
 - Resource tracking
- 3 Framework**
 - Purpose
 - Framework
 - Principles
- 4 Implications for monitoring and evaluation**
 - Monitoring implementation of the framework
 - Indicators and measurement issues
 - Evaluation design
- 5 Applying the framework**
 - Applying to specific initiatives
 - Adapting to countries

1 Background

Countries and their development partners are aiming to make significant increases in the level of investment and activity in health to achieve national or international health-related outcomes, focusing on the health-related MDGs². This scale-up in effort and commitment reflects a new consensus on several fronts: that scaling-up in health requires a far more coherent approach than that of the past decade; that targets related to health outcomes and impact cannot be achieved without adequate investment in the health systems that underpin health service delivery; that investment in health needs to be embedded in broader social and economic development planning; that countries need long-term predictable aid from development partners; that partners need to see a clear link between financing and results; and that mechanisms are needed to hold all partners accountable for their performance against international agreements. Other reflections of this consensus are apparent in follow-up on the High Level Forum on the Health MDGs (HLF), the Paris Declaration on Aid Effectiveness, the development of the GAVI health systems window, the discussion within the Global Fund on modalities for health systems support and conditions for more programmatic funding, the G8 communiqué on scaling up for health in Africa, and most recently the International Health Partnership (IHP).

The scale-up focuses on health-related outcomes. Increasing aid effectiveness, improving policy, strategy and health systems performance, and mobilizing stakeholders (including non-state actors) more efficiently are all means to this end. Actions will be country-focused and country-led, building upon existing mechanisms and health plans in countries and supported by dedicated regional and global activities, if necessary. The work will include the support for the implementation of the IHP and related partnerships and initiatives, such as the Global Fund, GAVI, the Global Campaign for Health MDGs, the Catalytic Initiative to Save a Million Lives, Women and Children First, the Partnership for Maternal, Newborn and Child Health, and the Global Health Workforce Alliance.

The scale-up to achieve the health MDGs is unprecedented in both potential resources and the number of involved initiatives. Such a grand experiment requires a harmonized monitoring and evaluation effort that reinforces both country and global needs to demonstrate results, secure future funding, and enhance the evidence base for intervention packages.^{3 4 5} It also raises questions about how best to scale-up different types of interventions and approaches in widely varying but uniformly poor country contexts, particularly with regard to long-term predictable and short-term results-based funding. Eventually, the scale-up efforts will be judged by country progress towards the health-related MDGs and related health goals, the degree to which major health constraints in countries have been addressed, and adherence to the Paris Declaration on Aid Effectiveness.

This paper provides an overview of the current situation status of country, regional and global efforts to monitor and evaluate progress toward the health MDGs, including country health information systems, monitoring performance, evaluation, and aid effectiveness. A common framework for evaluation of the scale-up for better health is proposed in line with the Paris Declaration. It is intended to provide a basis for developing an investment case for sound evaluations of cost-effectiveness and impact across the major initiatives in a way that it improves harmonization and alignment and strengthens country health information systems. The final section outlines how the framework can be applied to design evaluation and meet the needs of international initiatives and countries.

² Scaling up for better health. IHP+ work plan of the Health 8 agencies. October 2007.

³ Victora CG, Black RE, Bryce. Learning from new initiatives in maternal and child health. *Lancet* 2007, 370: 113-144.

⁴ Murray CJL, Frenk J, Evans T. The Global Campaign for the Health MDGs: challenges, opportunities and the imperative of shared learning. *Lancet* 2007, 370: 1018-1020.

⁵ Bennett S; Boerma JT, Brugha R. Scaling up HIV/AIDS evaluation. *Lancet* 2006, 367: 79-82.

2 Current situation

This section describes the current situation regarding country health information systems, monitoring performance, evaluation and its barriers, aid effectiveness and resource tracking.

Health Information Systems

Evaluation is rarely an integral part of country health information systems. Conversely, evaluation studies are rarely conducted in a way that they strengthen health information systems. In large measure, this is because health information systems are simply too weak to generate the data required for monitoring performance and evaluation. The information system is fragmented by disease-specific programme demands which often are related to donor requirements and international initiatives directed to specific areas such as malaria, HIV/AIDS, or tuberculosis. Reporting on progress towards achieving specified targets for specific indicators has become more important with the introduction of performance-based disbursement in several international initiatives, such as GAVI and the Global Fund to fight AIDS, TB and Malaria (GFATM), and the President's Emergency Plan for AIDS Relief, and has in several countries further aggravated the gaps between demand and supply.

The main data sources of a country health information system are either based on administrative or clinical information or derived from household visits. The current situation in most low and lower middle income countries shows that there are large differences between the strength of the different sources which results in significant gaps in data availability and quality. In general, for most administrative and clinical data it is often difficult to ascertain quality. There are major issues with timeliness, completeness, accuracy, diagnostic tools used, etc. Various efforts aim to improve the quality of such data, by setting standards for reporting, training of health workers, reconciliation of administrative and population-based data sources, etc.

The situation concerning data sources has its impact on the availability of data on key components of the scale-up. First, data on health systems inputs tends to be of poor quality and has suffered from considerable underinvestment in the past decade (e.g. human resources, health infrastructure, financing). A recent review of health systems performance indicators showed that there is no lack of indicators, but that data availability, quality and relevance are major issues. Several attempts to develop a set of indicators and measurement strategies are underway.⁶

Second, accurate data on access to services is lacking in most countries. This is because there is no good data on service delivery (availability and quality) and there are too few systematic efforts to link service delivery with population access. Third, data availability on coverage of interventions is better than for other areas through frequent population-based surveys and fairly accurate clinical reporting systems for some interventions (e.g. vaccinations, TB treatment). Lastly, data on health outcomes are still lacking for many indicators. It is easier to obtain sound estimates of child mortality through household surveys than estimates for adult mortality or other health outcomes such as social and financial risk protection.

Monitoring and evaluation

There are multiple dimensions to monitoring and evaluation (see Box 1 for descriptions), ranging from the monitoring of global goals and initiatives, programme and project monitoring focused on processes and immediate outputs, and impact evaluation of large programmes. These dimensions are overlapping and require an integrated approach at the country level. Ideally, sound monitoring provides much of the data

⁶ Walford V. Indicators of health system performance. Report of a review for DFID. DFID Health Resource Centre report. October 2007.

required for an evaluation, including baseline data. Unfortunately, due to the weakness of information systems in many countries, data are often incomplete and of poor quality, which hampers their utility for evaluation studies.

Additional data collection is usually required for evaluation purposes because routine data collection must be limited to avoid overburdening programme implementers and because contextual factors - usually not part of monitoring activities - need to be taken into account.

Common terms

Evaluation is the rigorous, science-based collection of information about program activities, characteristics, outcomes and impact that determines the merit or worth of a specific program or intervention.

Monitoring can be defined as the routine tracking and reporting of priority information about a program and its intended outputs and outcomes.⁷ This primarily includes monitoring of program inputs and outputs through record-keeping and regular reporting systems, which is sometimes referred to as process evaluation. Monitoring is a basic component of all programs to assess whether resources are spent according to plan and whether the program is resulting in the expected outputs.

Monitoring performance is a form of project or programme monitoring which aims to provide feedback for improving performance and implementation. Ideally, there are well-defined benchmarks which are used to measure progress in relation to inputs, mostly financing.

Performance-based funding was developed in the education sector and is currently used by several development initiatives. The Global Fund reported that performance-based funding appears to provide powerful incentives to scale up the fight against AIDS, TB, and malaria by linking finance to the delivery of health services.⁸ Targets are set and owned by countries to ensure that poorer countries are not penalized for lower performance.

Data quality audits are used by the GAVI Alliance to assess recording and reporting systems in the context of performance-based disbursements to reward for achieving increased immunization coverage. GAVI provides grants to improve the quality of data, but reward payments are ultimately affected if data quality does not meet a certain standard.⁹

Obstacles to evaluation

Current practices in evaluation are still sub-optimal, particularly if public health programmes in developing countries are to be driven by sound evidence¹⁰. The importance of evaluation has clearly been illustrated in a compilation of 15 major large-scale public health successes ranging from prevention of HIV/AIDS and

⁷ UNAIDS. National AIDS programmes. A guide to monitoring and evaluation. Geneva: UNAIDS/00.17E. 2000. Rugg D, Peersman G, Carael M (eds.) Global advances in HIV/AIDS monitoring and evaluation. New Directions for Evaluation 103, 2004. Rehle T, Saidel T, Mills S, Magnani R. Evaluating programs for HIV/AIDS prevention and care in developing countries. A handbook for program managers and decision makers. Family Health International, Arlington, VA: 2001. UNAIDS and World Bank. National AIDS councils: Monitoring and evaluation operations manual. UNAIDS/02.74E. 2002: Geneva.

⁸ Low-Beer D, Afkhami H, Komatsu R et al. Making performance-based funding work for health. PLoS Medicine 2007, 4: 1308-1311.

⁹ Ronveaux O, Rickert D, Hadler S et al. The immunization data quality audit verifying the quality and consistency of immunization monitoring systems. Bull WHO 2005, 83: 503-510.

¹⁰ Evaluation of the effects of the scale up. Health Metrics Network Working Group on Evaluation, Informal Planning Meeting. Johns Hopkins School of Public Health, Baltimore, Maryland, 23 June 2004.

sexually transmitted infections in Thailand, to improving the health of the poor in Mexico¹¹ and multi-country evaluations of major international programmes such as Integrated Management of Childhood Illness (IMCI).¹²

Country evaluations are rarely conducted as there are several obstacles to evaluation.^{2,3,4} First, evaluations require collective action. Neither country-level stakeholders nor any single global level stakeholder appears to have sufficient incentive to invest adequately in evaluation, and agreement to jointly invest in an evaluation has rarely occurred.

Political and technical issues related to attribution pose a second challenge. Donor countries and global initiatives often have considerable pressure from their own constituencies to demonstrate achievements in terms of lives saved or other health outcomes. From the country perspective, such preoccupation makes less sense as the primary concern is the measurement of overall progress in the context of a scale-up involving multiple actors.

Attribution is also a challenge from the technical perspective. Causal chains in public health are complex. Victora et al.¹³ distinguished three types of scientific inference that are often used for policymaking in the field of health and nutrition, each based on different approaches in evaluation. Probability statements are based on randomized clinical trial results, and in a few cases on randomized community trials. Plausibility statements are derived from non-randomized evaluations aimed at making causal statements using observational designs with a comparison group. Adequacy statements result from demonstrations that trends in process indicators, impact indicators, or both show substantial progress, suggesting that the intervention has an important effect. More emphasis is needed to develop evaluation standards and protocols for use in circumstances where randomized controlled trials are not appropriate. In fact, the pre-occupation with RCT which is less suitable for evaluation of a broad-based scale-up of multiple health interventions has hampered investments in other types of evaluation. It is often impossible to establish comparison groups that resemble the intervention groups and because it is often very difficult to minimize exposure to the interventions in the comparison group¹⁴.

Third, there are concerns about unfavorable results of an evaluation. Large recipients of funding may understandably be worried about evaluation findings, which they feel might affect the way their country is perceived and jeopardize future funding. Moreover, evaluators are not always sufficiently independent from country or international pressures which may adversely affect the quality and/or credibility of the results.

Fourth, evaluation is often an afterthought and too little investment has been made in baseline data collection and systematic monitoring of various components of implementation of the programmes and interventions. Country health information systems are often too weak to provide the kind of quality data required for a robust evaluation and few efforts are made to strengthen such systems, except for short-term quick fixes focusing on monitoring data. There has been a disproportionate focus on the creation of large numbers of indicators and reporting requirements, which has led to fragmentation and duplication of data collection and poor quality of statistics.¹⁵

¹¹ Levine R and the What works working group. Center for Global Development. Washington DC. November 2004.

¹² Bryce J, Victora CG, Habicht JP, Black RE, Scherpbier RW. Programmatic pathways to child survival: results of a multi-country evaluation of Integrated Management of Childhood Illness. *Health Pol Plann* 2005, 20 (Suppl 1): i6-i17.

¹³ Victora CG, Habicht JP, Bryce J. Evidence based public health: moving beyond randomized trials. *AJPH* 2004, 94: 400-405.

¹⁴ Figueroa ME, Bertrand JE, Kincaid L. Evaluating the impact of communication programs. MEASURE Evaluation and the Population Communication Services project. October 4-5 2001, Belmont Conference Center, Elkridge, MD, USA.

¹⁵ Boerma JT, Stansfield S. Health statistics now: are we making the right investments? *Lancet* 2007, 369: 779-786.

Fifth, countries may be hesitant to undertake rigorous evaluation studies due to weak capacity in research and evaluation. Finally, while it is often recommended that as much as 5% to 10% of the scale-up funds need to be set aside for monitoring performance and evaluation, this rarely occurs. In addition, resources are very limited for operational research to address complex questions such as finding the most effective mix of results-based and long term predictable funding and the system-wide effects of the scale-up.

Aid effectiveness

The Paris agenda on aid effectiveness and the principles for Global Health Partnerships agreed at the High Level Forum on the Health MDGs (HLF) identified the importance of improved coordination and harmonization of partner support to countries. More recently, the OECD/DAC has adopted health as a "tracer sector"¹⁶ in the lead up to the Accra High-Level Forum on Aid Effectiveness, to be held in September 2008. Support in this area will aim to link country, regional, and global processes.

The Paris Declaration is based on five key assumptions about the ability of aid to promote country development: country ownership, alignment with country development strategies and systems, harmonization among donor activities, managing for results, and mutual accountability. Regarding the latter, the Paris Declaration states that "partner countries and donors commit to jointly assessing through existing and increasingly objective country level mechanisms mutual progress in implementing agreed commitments on aid effectiveness, including the partnership commitments".

This requires systems to monitor the progress and performance of all stakeholders. An accountability framework outlining milestones and processes for this work will be needed to measure the performance of development partners working at country level, as well as the performance of countries in meeting their national commitments. Even though the focus is on results and accountability, evaluation is not highlighted in the Paris Declaration. The focus is on performance monitoring, including the provision of quality of information, stakeholder access to information (transparency), and coordination of monitoring and evaluation in countries.

In 2006, OECD published the first report on monitoring the Paris Declaration¹⁷, using 12 indicators developed to monitor the process of aid effectiveness. Most data are collected from a wide range of key informants among donors and partner countries. A subsequent assessment will be conducted in 2008 and targets have been established for 2010. The health sector has been identified as a tracer sector for aid alignment. The 2006 assessment concluded that in half of the countries signing on to the Paris Declaration, partners and donors have a long way to go to meet commitments they have undertaken. Several sector-specific studies are conducted on resource flows and aid effectiveness, mostly focusing on specific diseases such as HIV/AIDS (see Box). Building upon the OECD and disease-specific experiences, further work is needed to determine a set of evaluation indicators and monitoring methods.

¹⁶ A WHO-World Bank programme of work has been agreed with the OECD/DAC secretariat which has four elements: developing the evidence-base on aid effectiveness; identifying donor constraints for providing long-term sustainable financing in health; strengthening the link between aid effectiveness and health systems development; and strengthening global accountability mechanisms.

¹⁷ OECD. Survey on monitoring the Paris declaration. OECD Journal on Development, vol 8, no.2, 2006.

The three ones for HIV/AIDS

The Global Task Team on Improving AIDS Coordination among Multilateral Institutions and International Donors built upon the Paris Declaration on Aid Effectiveness and the "Three Ones" principles and aimed to address the realities of donor practices at country level¹⁸. Greater harmonization and alignment are prerequisites for effective use of increased funding. In spite of the efforts however, national M&E plans only exist in half of the reporting countries - too little is being done to build capacity to enable national ownership. Working methods of development partners have to be changed. Procurement is a well known stumbling block, and attribution is another. Half of the countries report low to moderate sharing of M&E results by international partners.

The essence of mutual accountability is answerability - providing information and explanations for action and inaction and being liable to sanctions for failure to deliver (reactive form). Another view focuses less on control, but more on transforming relationships between those making decisions and those affected by them. This proactive approach, involving stakeholders in decision-making draws attention to the potential of accountability to improve performance and therefore as a process to be embraced rather than feared. There may be tension between national ownership and donor concerns about accountability.

Resource tracking

Global level resource tracking relies to a large extent on the Creditor Reporting System (CRS), maintained by the Development Assistance Committee of OECD.¹⁹ In addition, the DAC database allows an evaluation of completeness of reporting to the CRS, which is high for the health sector. There are several mechanisms and studies in place which seek to supplement and validate CRS data, often focusing on specific health and disease programmes. Donors have a clear role in providing recipient governments with timely data on aid resource flows. They also have a global role of tracking aid flows by (sub)sector. Accountability to domestic electorates for bilateral donors may take precedence over the need to provide detailed information to recipient countries and the international community, and selective reporting of success may occur. The principles laid out in the Paris Aid Declaration aim to minimize such potential conflicts of interest.

At the country level, National Health Accounts (NHA) exercises are required to obtain comprehensive information on expenditure within a health system. Hitherto, only about 50 developing countries have conducted a NHA. Ideally, NHA relies heavily on public expenditure management systems for data on government expenditures. Multiple data sources are used to collect data on private expenditures, such as surveys on public expenditures.

¹⁸ Buse K, Sidibe M, Whymys D, Huijts I, Jensen S. Scaling up the HIV/AIDS response: from alignment and harmonization to mutual accountability. ODI Briefing Paper No. 9, August 2006.

¹⁹ Powell- Jackson T, Mills A. A review of health resource tracking in developing countries. Health Policy and Planning 2007, 22:353-362.

3 Framework

Purpose

A common framework for monitoring performance and evaluation for the scale-up for better health, sometimes also referred to as the campaign for the health MDGs, aims to ensure that the demand for accountability and results from single donor and joint initiatives is translated in well-coordinated efforts to monitor performance and conduct evaluation in a way that it also strengthens country health information systems. A common framework also aims to align the evaluation work with the tenets of the Paris Declaration:

- Country ownership implies involvement national stakeholders in the whole process
- Alignment is needed with national health information systems
- Harmonization means that the focus should be on joint attribution to scale up, rather than individual attribution
- Managing for results focuses on health impact and a need for rapid feedback on outputs
- Mutual accountability implies transparency, independence, sharing and learning.

Various global health initiatives focus on different interventions, although there is now agreement that health systems strengthening should be part and parcel of scale-up efforts. For instance, the Global Fund focuses on interventions against AIDS, TB and malaria, while the Catalytic Initiative aims at interventions to improve maternal, neonatal and child health. For each initiative specific pathways of how the activities and processes are expected to lead to improved health need to be developed, as interventions and health outcomes differ. The common or overall framework presented here is intended as an overall guide for such results frameworks. It is important that initiative-specific evaluations harmonize with other evaluation efforts in the same country and explore ways in which monitoring and evaluation can strengthen country health information systems in a structural manner.

The framework aims to apply to all international initiatives and partnerships involved in the scaling-up for better health, while maintaining immediate country relevance. Scaling-up for better health is defined as all international and domestic efforts to accelerate progress towards achieving the Millennium Development Goals. This may extend to related health priorities in countries such as health security or the double burden of disease. The goal is that all key partners will adhere to the common framework to maximize the health impact of the efforts.

Framework

Figure 1 presents the proposed framework for evaluation of the scale-up for better health in the context of the IHP+. The top section shows the general sequence used in monitoring and evaluation frameworks from inputs and processes to outputs, outcomes and impact. At the bottom of the figure, the proposed actions for monitoring performance and evaluation are shown. The next chapter discusses the implications of the framework for monitoring and evaluation.

Inputs and processes Domestic and international inputs and processes are listed in the leftmost side of the diagram. These include increased funding, improved planning and harmonization practices. Each of the initiatives emphasizes a subset of these inputs, and there is welcome redundancy in areas such as resources and capacity building. Increased funding, better planning and harmonization should lead to better national plan implementation, including health systems strengthening and acceleration of interventions against priority diseases, and to increased efficiency through harmonization and coordination. Capacity

building at institutional and individual levels are an essential element for progress and sustainability. Improved accountability has multiple dimension, such as better performance monitoring, focus on results and evaluation, and enhanced use of information for better practices. The inputs and processes are tracked through aid process monitoring and resource tracking, which are described in greater detail in the next chapter.

Outputs The middle box of the framework shows the expected outputs of improved implementation of the national health sector plans, achieved either through general health system strengthening or through specific intervention programmes. Outputs can include both personal and non-personal health interventions. Increased financing (level and equitable distribution) and better practices should lead to strengthening key building blocks in a generic and programme specific way. Improvements in governance, health workforce, medical products and technologies (supply chain, procurement) and information systems are part and parcel of health systems strengthening. This should lead to improved service delivery, including better access, safety, quality and efficiency, with special attention to equitable distribution of services. This would focus on the key interventions, but should not lead to a drop-off in the delivery of interventions that are not targeted.

Outcomes The increased outputs of the national health plans and processes are expected to increase health service utilization and coverage of key interventions. Coverage is defined as the proportion of the population who receive an intervention among those who need it. Coverage is influenced by supply (provision of services), demand and uptake by people in need of services.

Impact Ultimately, the increased coverage is expected to improve health, modulated by the efficacy of the interventions. For maternal, neonatal and child health interventions (including malaria) the primary impact should be reductions in child and maternal mortality and morbidity, and improvements in nutritional status. If financial and risk protection mechanisms are successful fewer families will become impoverished because of health expenditures, which is measured through household surveys.

Principles

The framework for evaluation of the scale-up, barriers to evaluation discussed in the previous section and the spirit of the Paris declaration can be translated into a set of principles.²⁰

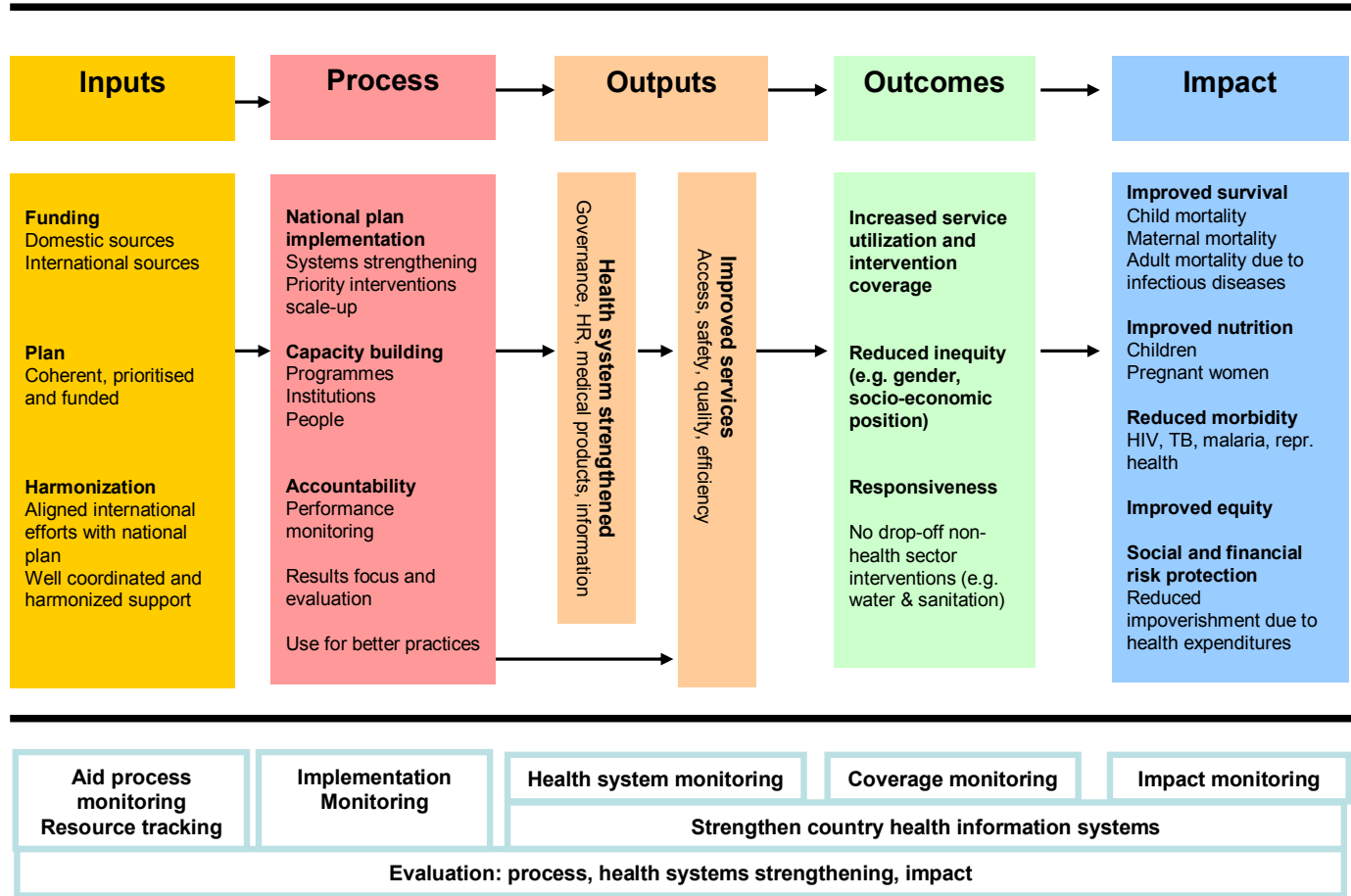
1. **Collective action:** the primary focus should be on the contribution of the collective efforts to scale-up the health sector response. This will imply convincing Boards of the Global Fund, GAVI, World Bank, bilateral donors, international agencies and others, that a jointly supported evaluation effort will be results-focused, and will be better, or at least match, what is already in place for them. Comparability of results and capacity building in monitoring and evaluation can best be developed through collaborative action. The various initiatives will also have individual evaluation needs because of differences in implementation (e.g. localized versus national support), but all evaluation should build from a common base and set of principles. In complex large-scale multi-intervention situations modelling with complete data sets is often the most cost effective way to obtain an idea about attribution in health outcomes to specific interventions or resources.
2. **Alignment with country processes:** monitoring performance and evaluation should build upon national processes that countries have established to evaluate progress in the implementation of national health sector plans. These include assessments and analysis conducted in the context of annual health sector

²⁰ Adapted from Savedoff WD, Levine R, Birdsall N. When will we ever learn? Improving lives through impact evaluation. Report of the Evaluation Gap Working Group. Center for Global Development, Washington DC. 2006.

reviews, mid term reviews of the implementation of the health sector strategic plans and planning of health sector strategic plans.

3. *Balance between country participation and independence*: evaluations should be driven by country needs and ensure active country participation, but conducted in a manner which maintains their independence. Evaluation may be very sensitive to countries, especially large recipients of funding. In particular, there needs to be a clear picture of the link between evaluation and funding eligibility. The latter is more linked to monitoring through performance based disbursement than evaluation is. Countries need to be protected from potentially adverse consequences of participating in evaluations, through making participation a positive condition for future programme funding; acknowledgment of poor results and adverse outcomes can be a powerful incentive to make change. Evaluators also need protection to maintain independence in the face of strong political and technical preferences on the part of the initiatives, but it is important to avoid turning the evaluation into a type of external audit. The evaluation process needs to ensure collaboration between policy makers, project managers, and evaluation experts: this can improve the content and use of impact evaluation, ensuring relevant questions and reliable evidence.
4. *Harmonized approaches to evaluation and performance assessment*: evaluations should use common protocols and standardized outcome indicators and measurement tools, with appropriate country adaptations and leadership, minimizing the implementation of separate evaluation efforts of in individual initiatives, grants and loans. This also means adoption of general principles of data sharing and transparency for all involved. All impact evaluations require advance planning, careful attention to confounding factors and systematic bias, and adequate time and money.¹² It is particularly important to highlight the need to build impact evaluations in at the design stage, when the expected impact of the program is being defined and appropriate comparisons can be established. High standards can be established and upheld by submitting impact evaluation proposals for review by independent experts to increase transparency and assist the research community in identifying publication bias.
5. *Capacity building and health information system strengthening*: systematic involvement of country institutions in the evaluation is necessary to strengthen health information systems and includes data collection, data verification, and analysis of information and evidence. The Health Metrics Network framework and assessments conducted in dozens of countries provide a general guide to strengthening of health information systems in a sustainable manner. The country health information system should include evaluation as an integral part. Strengthening data collection as part of monitoring performance should be an important input into evaluation. Evaluation should be built upon information generated by a health information system, based on regular well-planned data collection and analysis activities (including surveys, surveillance, service statistics etc.) that form the basic information for the evaluation. In addition, special evaluation research activities will be conducted for the purpose of the evaluation but the more integrated these activities are into the health information system the better. As such, evaluation activities can catalyze the development of sound health information systems.
6. *Adequate funding*: as a general guide between 5% and 10% of the overall scale-up funds need to be set aside for monitoring and evaluation, with a large proportion of this going to the in-country efforts and building capacity of national and regional institutions. There should be dedicated resources for the three components: monitoring performance, health information system building by addressing major data gaps, and impact evaluation. Roughly, one-third should go to each component. The health information system building component is information equivalent to the health systems strengthening component as part of the initiative.

Framework for monitoring performance and evaluation of the scale-up for better health



4 Implications for monitoring and evaluation

Monitoring implementation of the framework

Figure 1 shows the different aspects of monitoring the implementation of the framework at different levels. A illustrative set of possible indicators is shown in Annex A.

Aid process monitoring

Monitoring aid effectiveness needs to be based on analysis of aid flows, country health information (from inputs to outcomes), take into account efficiency, and compare with other investments. Aid effectiveness requires evaluation which can be based on analysis of secondary data for a single or multiple countries²¹ or involves significant new data collection using standardized protocols and data collection methods in a single or multiple countries.

Global and country inputs and processes can be monitored through tracking of international resource flows and regular assessment of aid processes. This includes monitoring the level and trends in resource flows by disease and country (National Health Accounts). The extent to which international resources are supporting and integrated into national health plans is monitored as part of the OECD/DAC biannual survey which will need to be adapted to become more specific for the health sector. In addition, the approach developed by UNAIDS for HIV/AIDS in the Country Harmonization and Alignment Tool (CHAT) can be used for an annual assessment of how well international and domestic partners are working together at country level.²²

Strengthening health information systems

The extent to which the scale-up leads to changes in outputs, outcomes and impact needs to be measured at the country level by strengthening data generation and analysis. Health systems strengthening is an essential part of the scale-up. Such interventions are defined as those that address barriers and constraints at different levels of the health system with the overall goal of improving health outcomes. Investing in large scale health system strengthening interventions requires concomitant investments in generating evidence to understand what works.

The success of efforts to strengthen health information systems can be monitored by its ability to produce data and statistics for the key levels of the framework: outputs, outcomes and impact. In the long run, strengthening of health information systems should lead to better availability and quality of health statistics for these key indicators, accompanied by greater use of data for decision-making. This would lead to availability of accurate baselines and sound monitoring of programme implementation to support evaluation work. In addition to meeting the immediate data needs in the context of the scale-up for better health strengthening of health information systems should lead to investments to improve neglected data sources, such as death registration (with cause of death) and health systems and services information.

The harmonization and efficiency of the process of data collection and analysis (e.g. a well-coordinated survey plan to meet all information needs) and the support to country capacity building in this area are

²¹ An example is a recent analysis of immunization coverage trends in relation to funding. Lu C, Michaud CM, Gakidou E, Khan K, Murray CJL. Effect of the Global Alliance for Vaccines and Immunisation on diphtheria, tetanus, and pertussis vaccine coverage: an independent assessment. *Lancet* 2007, 368: 1088-1095.

²² http://data.unaids.org/pub/Report/2007/jc1321_chat_en.pdf

essential. Furthermore, the use of the results for better management and policies is a crucial element, though much hard to measure.

Integrated evaluation plans

Finally, intermediate success is measured by the existence of evaluation plans, investments in such plans and ultimately proper implementation of evaluation studies. The latter refers to conducting the evaluations in line with the general principles of the framework. In addition to evaluation of country progress, special research studies are required to address complex questions such as finding the most effective mix of results-based and long term predictable funding.

Indicators and measurement issues

The development of indicators for different levels of indicators is usually given much attention. For monitoring performance the indicator selection is the basis for setting baselines and targets and for reporting. For evaluation the indicators - often the same as for monitoring, but with more emphasis on health outcomes and non-health related factors - are intended to carefully document trends with special reference to attribution and equity. This selection of a core set of indicators helps harmonization and currently there is general agreement about most key measures of coverage, risk factors, morbidity and mortality.

In the context of evaluation decisions about investments in data collection (or in some cases secondary analyses of existing data) are more critical. All indicators for health impact evaluations roughly fall into four domains and each domain is linked to specific data collection sources. Table 1 shows the primary and secondary sources of data for these indicators. It shows that household surveys are the preferred mode of data collection for coverage, morbidity (often only if a good biological test exists), and mortality, notably child mortality. For mortality and causes of death in general vital registration would be the preferred source. Clinical data are generally less useful to estimate coverage, except for some interventions that cannot be collected effectively in household surveys (e.g. ARV therapy or TB treatment). Data on health system inputs and processes are derived from facility assessments and administrative sources.

Table 1: Primary and secondary sources of data for health indicators by domain.

	Health system inputs/ processes	Coverage	Equity, determinants, risk factors	Morbidity	Mortality
Household survey		++	+++	+	++
Facility assessments	++				
Clinic data		+		+	
Administrative records	++				
Death registration			++		+++

Performance based disbursement mostly relies on clinical data or administrative sources. Disbursing funds based on data for indicators generated through population-based health surveys would only be possible if there is a long reporting/disbursement period. This could involve coverage and there is also some discussion about using health outcome data. Hitherto, no initiative has experience with such performance monitoring.

In evaluation design, major decisions will need to be taken about the size and frequency of data collection. It will be important to consider existing and planned data collection surveys and facility assessments, and avoid a tendency to implement data collection for the sake of the evaluation only. This way the evaluation contributes to building a country health information system. Another way would be to ensure that some of

the scale-up funding is set aside for the less immediate but most neglected data sources, notably death registration with proper coding of causes of death. Such mid to long term investments need to be part and parcel of health system strengthening.

Outputs: regarding outputs service delivery is a major focus of the scale-up and needs considerable investment to improve monitoring, building upon ongoing work to harmonize indicators and data collection approaches. Service delivery is one of the six building blocks of health systems, according to the WHO framework for health systems²³. The other five are finances, health workforce, information, governance, medical products and technologies. Monitoring health systems strengthening is still in its early stages. A WHO/World Bank working group is working with IHP+ and other countries to develop guidance and tools for methods and indicators for monitoring the strengthening of health systems and assessment of health systems performance. In general, there is reasonable agreement about the indicators for most building blocks, but the measurement investments have been inadequate. For instance, most countries do not have a system of regular monitoring of the access to interventions²⁴. Such a system needs to capture the availability at the service delivery point, and the distribution of interventions within the country; have a direct link with management of the scale-up of services; and contribute to health system strengthening. This generally implies a focus on districts or equivalent administrative structures. A regular system of monitoring all public and private facilities would be required, including assessment of the availability of trained health workers, infrastructure and equipment and drugs and commodities. Such data provide information on geographic or physical access to services that meet a minimum standard, initially based on facility density per district population. More advanced monitoring of physical access, such as populations living with a specific radius of a service delivery point or traveling time, are still in the realm of special studies, although advances in the use of geographic information systems may help.

Outcomes: The main data source for coverage rates are population-based surveys, which can provide more accurate data for several interventions, broken down by key determinants such as gender, wealth and social class. For some interventions, clinical data are the sole source of information, e.g. ARV therapy coverage to TB treatment success rates. For a few interventions both clinical and population-based sources are used, e.g. immunization coverage. Sentinel clinics, intensified supervision or a regular facility census may help obtain better data on numbers of people using a specific service.

Impact: Regular high-quality household surveys are the primary source of information, if death registration systems are not functioning. Changes in the incidence or prevalence of disease are difficult to document with a few exceptions, and have to rely on a mix of data from surveillance and surveys and modelling to fill data gaps. Investments in ways to obtain information on child and adult mortality and causes of death are needed, disaggregated by key equity related factors. The extent to which the health system is responsive to the users' expectations and to which it provides social and financial risk protection are also important outcomes that require more systematic investment in measurement. This includes surveys, improvement of financial record systems, and regular reviews, such as national health accounts, to bring together data from multiple sources

²³ WHO. Everybody's business. Strengthening health systems to improve health outcomes. WHO's Framework for Action. Geneva. 2007.

²⁴ Access to services has multiple dimensions. It can be defined in terms of reach-ability (physical access), affordability (economic access), and acceptability (socio-cultural access) of services that meet a minimum standard of quality. See e.g. Tanahashi T. Health services coverage and its evaluation. *Bulletin of the World Health Organization*, 1978, 56:295–303.

Evaluation design issues²⁵

Evaluation requires planning, investments and implementation. It should build upon monitoring data, providing information on baselines and trends in key health indicators and exposure to interventions, but generally requires additional data collection. The monitoring system will undoubtedly have gaps requiring more in-depth information, and periodic assessments of data quality and comprehensiveness. There are a number of design issues that must be addressed in collaboration with those planning for the development and implementation of accelerated strategies at country level. Examples of these issues include the following:

- *Definition of results:* how will “results” be defined for performance-based disbursement? Numbers of recipients of interventions? Quantities of commodities disbursed? Numbers of lives saved?
- *Periodicity and timing:* how often is information needed on outcomes and impact, in terms of performance based disbursement?; when should an evaluation take place?
- *Comparison groups:* while experimental designs are usually not possible with large-scale intervention programmes, there may be opportunities to classify populations by level of intervention intensity and quality, which strengthen the ability of the evaluation to attribute changes to the interventions. Will it be possible to define comparison groups for the evaluation? Will interventions be scaled up rapidly at national scale or deployed in selected districts only? If implementation is national, will be phased in gradually in different areas? If implementation is staggered in time, or interventions are introduced at different times, what is the time line for each area?
- *Placement of interventions:* often, the allocation of the intervention to populations is non-random, even in national programmes (placement bias). For instance, if the interventions are placed in more easily accessible better off populations there benefits may be greater than would be the case if the interventions had been placed elsewhere. On the other hand, if the intervention takes place in the poorest and least served areas, the scope for improvement may be greater, due to higher mortality associated with more easily preventable causes.
- *Existing data sources:* the integration of additional data collection into the existing health information system is important. For instance, aligning with the timing of a Demographic and Health Survey would have great cost-saving and help integration with the country data collection plans. When were the last population-based surveys carried out (e.g., DHS/MICS)? When are they planned for the future? What is the coverage and quality of the existing health information system? Is it likely to improve during the timeline of the evaluation?
- *Contextual factors, including health systems strength:* non-intervention changes -changes in socio-political, economic and environmental situation - can have large implications on the outcome of a health programme. In some cases, this involves macro level data (e.g. on political stability), in others it implies incorporating data from non-health sources such as economic survey or meteorological data. How much information is needed on socioeconomic changes and on separate health programs that can influence coverage and impact (that must be taken into account when attributing changes to the initiative)? How will health systems strengthening activities be incorporated into the conceptual model within each country? How can elements of the system, such as human resources, be monitored in comparable ways across very different contexts?

²⁵ Expanded from Institute for International Programs, Bloomberg School of Public Health, Johns Hopkins University. Thoughts on the common evaluation framework for the CATALYTIC INITIATIVE TO SAVE A MILLION LIVES. OCT 2007.

- *Economic component of the evaluation:* What are the aims of the economic component of the evaluation? How does the scope of the initiative affect the design of economic evaluation (e.g. catalytic or leveraging function and/or direct disbursement of funds)?
- *Synthesis and modeling:* evaluations bring together data from multiple sources and rely on modelling to fill data gaps. Modelling is also relevant to examine attribution of specific effects to specific interventions or initiatives.

When evaluating programmes and projects it is useful to consider the following DAC Criteria, as laid out in the DAC Principles for Evaluation of Development Assistance²⁶. These are summarized in Annex B.

²⁶The DAC Principles for the Evaluation of Development Assistance, OECD (1991), Glossary of Terms Used in Evaluation, in 'Methods and Procedures in Aid Evaluation', OECD (1986), and the Glossary of Evaluation and Results Based Management (RBM) Terms, OECD (2000).

5 Applying the framework

Applying to specific initiatives

The overall framework describes the general aspects of performance monitoring and evaluation of the scale-up for the health MDGs. It can be used as a basis for the development of initiative-specific frameworks. In particular, these adapted frameworks need to be more specific about interventions and expected results. A few examples are given below.

In the *Catalytic Initiative to Save a Million Lives* (and related initiatives such as Accelerating Child Survival and Development) the focus is on the scale-up of a selected number of proven interventions to improve maternal, newborn and child survival.²⁷ The *Catalytic Initiative* focuses on specific elements of the overall framework. For instance, at the level of country inputs and processes the emphasis is on development of a coherent, prioritized and funded plan for maternal, neonatal and child survival, ability to monitor progress and using information to improve practices and results. The main outputs is a scale up of priority interventions and services for MDG4 and MDG5, leading to higher coverage among the target population. The results model further defines the pathways through which activities and processes are expected to result in reductions in maternal, newborn and child mortality, as well as elements within the broader context that may affect these pathways.

At country level, the conceptual model can be understood as having two major parts: strategy development and implementation, which will vary widely by country; and results, which should have a common core of outcome and impact indicators across countries, with some variation depending upon their epidemiological situation (e.g., the prevalence of malaria and HIV). The results to be achieved by countries working in the *Catalytic Initiative* and other efforts to scale up maternal, neonatal and child health activities will be similar, because all are working toward the improved nutrition and survival defined by the MDGs, and there are a short list of effective interventions to achieve these improvements. The results framework will need to be adapted to different country and geographic settings based on epidemiological and socio-economic features, but all will share a similar set of outcomes and impact.

The emphasis of the *International Health Partnership (IHP+)* is somewhat different from the Catalytic Initiative. At the global level it focuses on mutual accountability processes, coordination and harmonization among funders, and addressing agency barriers. At the country level, a broad set of process should lead to a strong and funded health sector plan, well-harmonized and coordinated support from donors and initiatives, improved financial and technical capacity to implement plans, including ability to monitor progress and evaluate. Improved accountability to inform domestic and global stakeholders and better use of information to improve practices and results are key elements. The major output is a strengthened health system that can increase access to quality services, resulting in improved coverage, with appropriate social and financial risk protection, and ultimately improved health.

The *Global Fund Five year evaluation study* is an example of an ongoing multi-country evaluation of a scale-up of interventions against specific diseases. The evaluation study was launched early 2007 and involves three study components. The first two components aim to assess the functioning of the secretariat and the partnerships in countries. In terms of the overall framework this involves an assessment of global and country inputs and processes. The third component is a study to assess the outputs, outcomes and health

²⁷ From: Institute for International Programs, Bloomberg School of Public Health, Johns Hopkins University. Thoughts on the common evaluation framework for the CATALYTIC INITIATIVE TO SAVE A MILLION LIVES. OCT 2007.

impact achieved in the last five years. It focuses on documenting trends in health outcomes, service coverage and health system inputs for AIDS, malaria and tuberculosis. Eight countries were selected for in-depth evaluation and 12 for secondary data analysis. A common framework was developed including data collection protocols. The individual country focus is not primarily on the Global Fund contribution but on whether the scale up is making a difference for the course of the three diseases. It is also assessed whether there are adverse effects on the health systems and the delivery of other interventions. Detailed results frameworks are used for each of the three diseases and for health system effects in general, building upon the System-Wide Effects study. Modelling will be used to fill data gaps (there is lack of baseline information on e.g. mortality by cause of death which hampers the ability to draw strong conclusions) and address issues of attribution.

What are the priority actions at the global level?

- *Harmonization*: efforts need to be made by H8 and major initiatives to harmonize efforts to evaluate and monitor performance in line with the Paris declaration
- *Joint planning*: while the different initiatives have different needs there is considerable overlap and the potential benefits from close collaboration on evaluation outweigh the costs, especially if one consider the ultimate goal of country progress
- *Breadth*: the framework should cover all countries in which activities are scaled up, it should also include all initiatives that are active in countries; while the focus is on national scope initiatives, this should also include multi-country large non-national efforts to scale up.
- *Establish monitoring process*: the implementation of these priority actions can be done through the aid effectiveness monitoring system, outlined in section 3.

What are the priority actions for health systems strengthening?

- *Evaluation*: there is a need to better specify the interventions for health systems strengthening, otherwise evaluation will not be possible. There may be considerable variation between countries in what the priority interventions are.
- *Health system metrics*: a common set of indicators and measurement tools will be essential, especially for service delivery. The indicators should be used to assess the strength of the health system, the extent to which changes occur as part of health system strengthening, and should be extended to include assessment of the performance of the health system using standardized approaches which can be applied at the country level.
- *Data quality*: there is a need to invest in the data sources for health systems monitoring, such as administrative records and databases (finances, human resources) and integrated facility surveys.

What are the priority actions for evaluation

- *Develop evaluation-specific frameworks*: this implies defining pathways through which the scaling up can have results. For outcomes (notably intervention coverage) and impact (mortality etc.) there is good agreement on the indicators, while data availability and quality has been improving. The general framework can be used to develop more specific evaluation frameworks, that provide clear pathways through which country interventions can lead to results. This also involves the clear definition and possible control of confounders.
- *Harmonization*: maximizing integration of the different framework is important as the instruments and vehicles for data collection methods will often be the same, e.g. a household survey.

Adapting to countries

Different kind of monitoring and evaluation activities may take place in countries (Table 2). The first two are related to monitoring activities in country, the third compiles data from countries, but may involve reporting requirements. The last two activities focus on evaluation and often require additional data collection and

analyses. Because the new and existing initiatives are so diverse, specific evaluation needs and designs will also vary. Yet, they have much more in common than is often thought, as they all rely on the same small set of data sources and involve the same institutions and actors in countries.

Multi-country evaluations using a standard framework can be helpful. Countries can learn from comparisons, especially with countries in regional groupings.

Table 2
Summary of proposed approach for monitoring and evaluation of scaling-up for better health and IHP+

M&E activity	Approach	Timing
Monitoring country progress in implementing national health plans that scale up services	Prioritised investments in data collection and analysis to improve quality and availability of core data – e.g. DHS, add new facility surveys, health workforce data	Annual, part of existing country annual review process where this exists
Monitoring implementation of specific initiatives <ul style="list-style-type: none"> - IHP - Catalytic initiative - Results based financing 	Look at process of implementation in selected countries, possibly with some ‘controls’, learn lessons and share experience	Annual, separate or joint across initiatives (depending on what is in each country)
Global level monitoring of scale up	Collate data from countries, plus DAC reporting. Include standard health system metrics and coverage data e.g. Countdown.	Annual
Evaluation of country health progress	Joint evaluation at country level, based on national plan and global framework.	To fit country health planning timetables and needs, agreed mid term reviews etc and link DHS etc to this
Global level evaluation of scale up	Multi-country evaluation studies: additional data collection as necessary; compilation of findings from country monitoring and evaluations	After 3 years, 6 years, post 2015

What are the priority actions at the country level?

- *Monitoring practices*: this includes monitoring of the buy-in into the national plan and monitoring the extent to which partners are adhering to global agreements at the country level, building upon the UNAIDS approach to monitor the three ones
- *Align evaluation with country planning cycles and mechanisms*: a major use of monitoring performance and evaluation results is for the annual health sector reviews, mid term reviews and 5 year health sector plan reviews.
- *Linking evaluation and monitoring performance to the national health information plans and system strengthening*: the evaluation can become a catalyst for improvement of the system, especially neglected areas such as causes of death and health systems information.

The availability of a country health sector strategic plan is a prerequisite for M&E. Most countries have a clear policy framework and national 5-year strategic plan for health. This includes annual operational plans and priority indicators. Many countries have developed the M&E framework as part of the SWAP but these plans rarely address data collection issues and there are too little resources.

The strengthening of country health information systems to improve monitoring performance and evaluation has the following priority areas:

- *Donor behaviour*: donor coordination mechanisms essential to harmonize funding streams and major disease programmes, including agreement on a minimum set of indicators in line with national strategic plan and sharing of evaluation plans
- *Health systems monitoring*: investments to improve the availability and quality of health systems and services data, with special attention for the district level
- *Survey harmonization*: common resource basket for health surveys with no strings attached, including mortality and causes of death
- *Institutional capacity building*: address weakness of institutional and analytic capacities. Evaluation must address issues of sustainability, country ownership, and learning component while ascertaining objectivity and independence.
- *Address orphan data sources*: investments in some data sources such as death registration may not pay off immediately, but are essential to ensure better future evaluation.

Annex A

Possible indicators to focus for evaluation of IHP+

FWC = First Wave Countries for IHP.

Input	Effects	Outputs	Outcomes	Impact
<p>Improved coordination among agencies</p> <p>Changes introduced by H8 agencies to streamline procedures and operations, harmonise, align.</p> <p>GFATM funds country health plans.</p> <p>Increasing efficiency of aid use; shown by more flexible funding for health systems and more predictable funding</p> <p>Strong mutual accountability process that holds agencies to account for their in-country behaviour and funding (incl. whether have kept to compacts in FWC)</p> <p>Volumes of aid disbursed - for TA, AIDS, other.</p>	<p>Aid effectiveness measures for health in FWC compared to other countries</p> <p>Appraisal of health plan – does it build systems? Does it have good M&E plan? Is the costing and budget realistic?</p> <p>Proportion of donor funds that are on budget and support the health plan; proportion that use national procurement systems;</p>	<p>Indicators & reviews of plan implementation – activities carried out, targets reached, expenditure levels met, services expanded, financing reforms developed</p> <p>Engagement of civil society and Parliament in planning and progress reviews</p> <p>Availability of data for M&E, quality of data</p> <p>Use of M&E findings in revising plans, strategies and budgets</p>	<p>Improved indicators of health system performance, in FWC and others – use Health System Metrics working group indicators or country indicators</p> <p>Service access and availability indicators, e.g. facilities equipped, staffed & with supplies to deliver ART</p> <p>Coverage of maternal and child health, AIDS, TB and malaria interventions including skilled birth attendance; PMTCT uptake; numbers on ART; immunisation rates; TB cure rates; IPT and ACT use etc</p> <p>Coverage analysed by socio-economic groups and vulnerable groups</p>	<p>Child and maternal mortality rates</p> <p>HIV, TB, malaria prevalence</p> <p>Change in numbers of families pushed into poverty by health spending</p>

Plus need to consider key assumptions and possible side effects, which would also need to be monitored.

Annex B

DAC criteria for evaluation

Relevance: the extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor. In evaluating the relevance of a programme or a project, it is useful to consider the following questions:

- To what extent are the objectives of the programme still valid?
- Are the activities and outputs of the programme consistent with the overall goal and the attainment of its objectives?
- Are the activities and outputs of the programme consistent with the intended impacts and effects?

Effectiveness: A measure of the extent to which an aid activity attains its objectives. In evaluating the effectiveness of a programme or a project, it is useful to consider the following questions:

- To what extent were the objectives achieved / are likely to be achieved?
- What were the major factors influencing the achievement or non-achievement of the objectives?

Efficiency: Efficiency measures the outputs -- qualitative and quantitative -- in relation to the inputs. It is an economic term which signifies that the aid uses the least costly resources possible in order to achieve the desired results. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been adopted. When evaluating the efficiency of a programme or a project, it is useful to consider the following questions:

- Were activities cost-efficient?
- Were objectives achieved on time?
- Was the programme or project implemented in the most efficient way compared to alternatives?

Impact: The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators. The examination should be concerned with both intended and unintended results and must also include the positive and negative impact of external factors, such as changes in terms of trade and financial conditions. When evaluating the impact of a programme or a project, it is useful to consider the following questions:

- What has happened as a result of the programme or project?
- What real difference has the activity made to the beneficiaries?
- How many people have been affected?

Sustainability: Sustainability is concerned with measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn. Projects need to be environmentally as well as financially sustainable. When evaluating the sustainability of a programme or a project, it is useful to consider the following questions:

- To what extent did the benefits of a programme or project continue after donor funding ceased?
- What were the major factors which influenced the achievement or non-achievement of sustainability of the programme or project?