

Health sector monitoring

Approaches, issues and lessons from a review of eleven countries

Javier Martínez, Mark Pearson and Dan Wilde

Executive Summary	2
1. Introduction	3
2. Measuring progress: the indicators	4
2.1 What do we want to measure?	4
2.2 The types of indicators: what are we actually measuring?	4
2.3 The number and quality of indicators	6
2.4 Is there an ideal mix of indicators?	9
2.5 Choosing indicators: summary of lessons	9
3. The annual review	10
3.1 What to review?	11
3.2 The annual review process	11
4. Follow up mechanisms	14
4.1 Follow up of annual reviews: lessons learnt	14
5. Conclusions and key messages	16

Executive summary

Fifteen years after the launch of sector wide approaches, health sector monitoring is receiving increasing attention, linked with broader international efforts to improve aid effectiveness. However, the development of effective performance monitoring frameworks continues to face challenges. This document reviews the main approaches used and emerging lessons learned, drawing on experience in the health sectors of eleven countries.

The **purpose of sector or programme monitoring** is not to measure everything, but to get a snapshot of whether things are going as expected, and to identify factors that affect progress. One overarching monitoring framework should help to reduce transaction costs and put government in the driving seat as the policy-making and implementing agency, but this balance is difficult to achieve.

This overview paper describes the three main components of a sector monitoring framework:

- a) a list of **indicators**;
- b) a review mechanism, usually taking the form of **annual reviews**;
- c) arrangements for government and development partners to jointly look at the results of annual reviews and to take necessary action (**follow up mechanisms**).

The first service a good **monitoring framework** can provide is to create demand for good information. The “perfect” set of indicators is often the enemy of an effective framework, and while there is no ideal number or mix, it should be possible to operate with twenty indicators or fewer.

Good **indicators** should be relevant, easy to interpret and understand, reliable, based on accessible data, and affordable in terms of data collection. In general, the quality of the information collected matters more than the quantity.

The indicators selected can be driven by the nature of government-development partners relationships and expectations, rather than a considered technical assessment. A large number of outcome, impact and financial input indicators suggest a “hands-off” approach. A predominance of output and process indicators suggests a “hands-on”/interventionist approach. Each has its pros and cons.

Annual reviews provide an opportunity to assess progress against the strategic plan on which the SWAp or sector programme is based, and should be an opportunity for collective reflection and revitalisation of the partnership. An over-ambitious review often results in too many poorly prioritised recommendations, and an unrealistic action plan.

The **follow-up** process determines the success of annual reviews as a planning tool. Unfortunately, the track record for realistic, manageable expectations from annual reviews is not good. Unreasonable recommendations and action plans erode government leadership and affect staff morale. It is better to under-commit and over-perform than the opposite.

In **conclusion**, although sector reviews can generate useful information, the results are often seriously under-utilised. Countries should be supported in developing a monitoring framework that meets their needs and is proportionate to the resources available and local capabilities.

1. Introduction

The issue of health sector monitoring is not new, but it is receiving increasing attention as the number of countries that support either a health sector strategic plan or a health sector programme has grown markedly over the last decade. Broader international efforts to improve aid effectiveness have given these efforts further impetus.

The development of frameworks to assess the performance of national development strategies and sector programmes is a key pillar of the Paris Declaration on Aid Effectiveness. The target is for 75% of partner countries to have developed transparent and monitorable performance assessment frameworks to assess progress against sector or programme plans by 2010¹. However, progress to date has been limited. The recent Survey on Monitoring the Paris Declaration² found that “translating evidence on results into processes of policy improvement remains a major challenge in the large majority of surveyed countries”. The survey suggests that “countries and donors should use performance assessment frameworks and more cost-effective results-oriented reporting” and that this “will require donors to invest in capacity development and rely more on country results reporting systems”.

Effective sector or programme monitoring is necessary for the government and its development partners to measure progress and assess whether resources result in more, and better, services as well as improved health outcomes. Yet, one and a half decades after the launch of sector wide approaches (SWAs), the development of robust and meaningful performance indicators and review mechanisms in sector programmes remains a major challenge³.

What are the key issues facing countries, and what lessons can be learnt from current experiences?

This document reviews the main approaches used and the issues faced in applying performance monitoring frameworks in the health sectors of eleven countries. Most of the countries in the sample (Ghana, Kyrgyzstan, Mozambique, Nicaragua, Tanzania, Uganda and Zambia) have adopted a sector wide approach to manage their health sector. The remaining countries (Cambodia, Ethiopia and the State of West Bengal in India) do not necessarily have a SWA in place – and may not even be in the process of developing one. Even so most external development partners support a single health plan defined and led by government, and government and partners increasingly agree to use common approaches, indicators, review processes and reporting mechanisms.⁴ In short, all the countries in our sample are attempting to apply the principles of harmonisation and alignment to their health sectors.

Most countries implementing sector programmes, whether in the form of a SWA or through “programme-based” approaches (increasingly used in Francophone Africa), have adopted very similar arrangements and tools for sector and programme monitoring. These generally include:

- a) a list of **indicators**, usually combining input, output, process, outcome and impact levels;
- b) a review mechanism, usually taking the form of **annual reviews**;
- c) arrangements for government and development partners to jointly look at the results of annual reviews and to take necessary action based (**follow up mechanisms**).

1. See glossary in the CD-Rom for definition of terms.

2. The 2006 Survey on Monitoring the Paris Declaration: Overview of the results (OECD, 2007) can be read at <http://www.oecd.org/dac/effectiveness/monitoring>

3. Not all health sector programmes use a SWA (where government and development partners agree on the use of common monitoring, review and reporting arrangements under government leadership). However, sector programmes and SWAs face similar issues in terms of adopting effective sector or programme monitoring arrangements.

4. Countries were selected on the basis of the availability of information, geographical representation, and efforts to adopt harmonisation and alignment principles.

All the eleven countries reviewed used such arrangements as part of their *sector monitoring framework*.⁵ This document describes and analyses approaches adopted in these countries. It is structured in three parts that examine: sector monitoring indicators (Section 2), the annual reviews (Section 3) and subsequent follow up mechanisms (Section 4). Country notes accompany this paper⁶ and provide details on these matters for each country examined.

2. Measuring progress: the indicators

2.1 What do we want to measure?

The purpose of sector or programme monitoring is not to measure everything but **to focus attention on a selection of indicators** that jointly depict whether the sector or programme is progressing as expected. The aim is, therefore, to measure what is important (as opposed to whatever is interesting, or easy to measure) and to identify factors affecting progress. Sector and programme monitoring are meant to provide a “snapshot” of progress, not a detailed diagnosis of every single part or component of the sector or programme. In a SWAp or sector programme, a single monitoring framework should be used by both government and development partners to measure progress. This, in turn, is expected to significantly reduce the need for each development partner to request further information to measure the effectiveness of their financial support to the sector.

In short, common monitoring arrangements in SWAs and in sector programmes that follow the principles of harmonisation and alignment aim to reduce transaction costs to the government, to free time for government staff and to deter development partners from becoming too involved in programme implementation, on the understanding that programme implementation is the responsibility of the government, not of development partners⁷.

Setting up sector and programme monitoring frameworks does not imply replacing or abolishing other existing monitoring frameworks, for example those used for disease control programmes like malaria, avian flu or HIV/AIDS, or for national programmes like the Expanded Programme for Immunisation (EPI) or the National Maternal and Child Health programme. Typically, these programmes have their own set of indicators, and while sector monitoring frameworks may include some indicators from these programmes, they should not attempt to incorporate too many of them. However as we will see later, achieving an overall sector view with an appropriate level of detail is difficult to achieve in practice.

2.2 The types of indicators: what are we actually measuring?

Table 1 shows the number and of type of indicators used in the countries covered in this paper.

The table clearly shows the wide variation in the numbers, types and mix of indicators used. This is perhaps not surprising given the different challenges that countries are facing.⁸ In practice, however, the choice of indicators is rarely based on a considered technical assessment. Rather they tend to reflect and respond to the nature of the partnerships and interactions (and the levels of trust) between governments and their development partners in each sector programme. For example:

5. Defined as the indicators, processes and results that jointly enable the monitoring function in the context of a programme or sector plan.

6. Available separately in the CD-Rom.

7. This does not apply to projects directly managed by development partners, a practice that is (allegedly) being phased out by most development partners, except in emergency and post-conflict situations.

8. Having said that, global efforts are under way to identify a possible “menu” of appropriate indicators from which partners could choose. The Health Metrics Network (HMN), a global partnership established to address the problems of weaknesses in health information, has developed a “Framework and Standards for the Development of Country Health Information Systems,” which includes a “parsimonious” set of core indicators. The Working Group on Health System Metrics, which involves HMN, WHO, the World Bank, the Global Fund to Fight AIDS, TB and Malaria and the GAVI Alliance, is developing a “dashboard” of indicators that can be used as a health system profile for countries, and guidance to countries on how to monitor health systems. For more information visit the HMN website: <http://www.who.int/healthmetrics/en/>

Table 1: Examples of indicators used for sector or programme monitoring

Country	Number of Indicators	Type of Indicators*				Comments
		Input	Output	Process	Outcome & Impact	
Bangladesh	62	7	12	32	11	
Cambodia	63 or 131	n/a	n/a	n/a	n/a	The M&E framework contains 63 indicators, but 131 were used for the Joint Annual Review.
Ethiopia	26 (HSDP II)	2	17	7	0	
	85 (HSDP III)	4	73	0	8	
Ghana	31	10	16	0	5	
Kyrgyzstan	27 (dashboard) 78 (milestones)	10	10	- 78	7	Two different sets of indicators are used.
Mozambique	36	n/a	n/a	n/a	11	The other indicators cover key areas of the Health Strategic Plan (14) or relate to ministry of health core functions (10).
Nicaragua	22	3	9	2	8	
Tanzania	33	7	5	4	17	22 annual, 11 periodic.
Uganda	25	3	15	7	-	
West Bengal	36	-	-	36	-	These are indicators of compliance to a workplan.
Zambia	26	4	13	8	1	26 indicators against 22 benchmarks on key programme areas.

* Note: Except for Ethiopia, Mozambique, Nicaragua and Uganda the indicators were classified by the reviewers.

- A predominance of **impact, outcome** and **financial input** indicators would suggest a sector programme where development partners adopt a “hands off” approach in relation to the implementing agency (usually the ministry of health). The implementing agency enjoys considerable flexibility in setting up its operational priorities – as long as progress can be demonstrated against key indicators linked to health status (e.g. mortality and fertility rates, lower incidence rates of major diseases, etc.). This model was adopted in the first health SWAPs in Africa (the “first generation” SWAPs), on the understanding that it was not the business of development partners to become involved in operational or managerial issues that were the clear responsibility of the ministry of health. This model has the advantage of focussing attention on higher level sector goals, but these indicators are only measured every four or five years (through Demographic and Health Surveys or similar), which can be a problem for development partners who need evidence that “more is being done” in the interim. By the time any lack of progress becomes apparent it

may be too late to do anything about it. This argument may or may not be true, since most ministries of health use intermediate indicators⁹ in their national programmes which allow progress to be measured and remedial actions to be identified where necessary.

- More recent (“second generation”) health SWAps in Africa and all the health SWAps in Asia have adopted tables of indicators that include a larger proportion of (service) **output** and **process** indicators (and targets) than of outcome and impact indicators. This approach enables greater and more regular attention to the performance of the implementing agency (usually the ministry of health) in transforming policy into practice. It denotes an arrangement where development partners have a more “hands on” or “interventionist” approach. The disadvantage of this approach is that it can easily become too rigid for government – since programme priorities do change, thus requiring frequent updates. An additional drawback is that the approach can become donor driven, reflecting the individual preferences of donors rather than those of government. Pressure from development partners can lead to the adoption of an increasing number of indicators and targets, which in turn makes regular (annual) measurement more difficult and costly.
- In a more recent approach, **process** indicators (often called milestones¹⁰) predominate. This approach is increasingly found in sector programmes in Asia, and is highlighted by the case of West Bengal. Cambodia and Bangladesh have also adopted more process indicators and milestones in their latest sector monitoring frameworks (see Table 1). In the context of sector monitoring, milestones refer to achievements agreed between the government and its development partners which somewhat demonstrate the government’s commitment to the sector programme. For example, the establishment of a monitoring unit within the directorate of planning may be a key milestone of the sector programme. The risk is that this approach might end up assessing performance in terms of the relationship between government and donors, which is important but may play a small role in terms of explaining progress achieved at sector level. This may be an issue where government performance and sector performance diverge.¹¹ An additional risk is that governments may perceive milestones as a hidden form of conditionality, particularly when they are linked to financial incentives or to disbursement of external funds (as is the case in Bangladesh under the performance based financing arrangement). Finally, using milestones often reflects a programme that is very donor driven, with less ownership and leadership by the government. This can lead to annual reviews that highlight failures rather than successes, and with greater focus on government rather than development partners performance (e.g. are they meeting funding pledges and disbursing on time?).

While these approaches do not exist in pure form, and tend to simplify a more complex reality, they are useful to highlight some of the issues to take into account when designing a sector programme and agreeing on sector monitoring indicators.

2.3 The number and quality of indicators

One of the most important lessons is that effective sector and programme monitoring does not necessarily require a large number of indicators. In fact, the “perfect” set of indicators is often the enemy of an effective monitoring framework. Countries should develop monitoring frameworks that meet their information needs and that are **proportionate to the resources available and to the capabilities that exist in country**. It is fairly obvious that the larger the number of indicators adopted, the greater the capabilities required for data and information handling. Nevertheless countries with a limited tradition and capacity for sector monitoring

9. Defined as the indicators between input and outcome.

10. Milestones are wrongly called “benchmarks” in some sector programmes. Benchmarks are meant to enable internal comparison of certain variables (for example “hospital quality” benchmarks). Since the purpose of sector monitoring is not to compare different health sectors, the use of the term benchmark seems inappropriate.

11. An example where programme performance and sector performance diverge can be seen in Bangladesh, where reasonable progress has been made with key health indicators despite the fact that Department of Health and Family Welfare performance was consistently rated as poor in annual sector reviews. This has prompted observers to question how sector indicators can improve when programme performance is poor, and illustrates that the link between sector and programme indicators is often elusive.

and with little interest in building such capacity tend to be the very ones opting for the most indicators. This asymmetry may be a result of external pressures (for instance from development partners keen to see their individual preferences reflected among the sector monitoring indicators and who perceive that an extensive framework can provide adequate safeguards where there are concerns about likely performance). Or, this may simply reveal a poor assessment of what can be realistically attempted.

Large numbers of indicators often imply that many of them are not measured at the required times, whether for technical, cost or capacity reasons. This has been reported in Tanzania, Ethiopia, Ghana, Cambodia and Bangladesh. In Bangladesh as many as 35 and 42 out of 62 indicators had not been measured at the time of the 2006 and 2007 annual reviews (see Box 1). If this situation recurs over time, there is a danger that it will impact negatively on the effectiveness of sector reviews, as an incomplete indicators table will become almost irrelevant to the review exercise. On the other hand, failure to report on certain indicators can be used positively for capacity building or investigation purposes. In Kyrgyzstan, for example, studies on informal payments in health facilities are ongoing and the presence of indicators in the performance framework acted as a trigger for commissioning the necessary work.

In general, the quality of the information collected matters more than the quantity. The practice of adopting large numbers of indicators tends to have a negative impact on the technical quality of data collected.

The first service a good monitoring framework can provide is to create demand for reliable data. Demand is unlikely to be generated when monitoring frameworks are used rigidly or have too much focus on government performance. Developing good indicators can be very difficult (for instance to measure variables like quality or equity) but is essential if sector progress is to be linked with implementing agency performance.

Box 1. Experiences from the Bangladesh Annual Programme Reviews

In Bangladesh the 2006 and 2007 Annual Programme Reviews did not manage to analyse progress against the Results Framework. Data on many key indicators was missing, either because it was not available, or because it could not be made available in a timely fashion, or because its validity was questioned; baseline information was not available for 42 out of the 62 indicators. This was partly due to the existence of a large number of indicators, together with the fact that the government had had little incentive to collect the data.

Too many indicators

The Bangladesh results framework had an extraordinarily large number of indicators, many of which were for annual reporting. The levels of effort and resources for data collection and analysis required were probably beyond the capacity of the government health system.

Lack of incentive to collect the data

The Results Framework seems to have been designed to serve the monitoring needs of the International Development Association/World Bank and other development partners. As a result, the government has little incentive to collect and utilise the information. Further, the health management information system cannot routinely provide information of many of the indicators, many of which would require costly (and at times annual) surveys. Ministry staff and development partners often lack qualified staff to interpret and put in context the information from the Results Framework. But if this is not done, the perceived value of the indicators diminishes, and the incentive for continued data collection is lost.

(Adapted from the Bangladesh country note. See the country note in the CD-Rom for sources and further information.)

The essential characteristics of a good indicator include being relevant, easy to interpret and understand, reliable, and based on accessible data. A good indicator should be SMART (meaning Specific, Measurable, Achievable, Realistic and Timely). Affordability should also be a key consideration, since the incremental costs associated with data collection of some indicators can be very substantial, even astronomical in the context of a poor country. Our review suggests that many indicators, especially from the larger monitoring frameworks, fail to meet these criteria.

Table 2 shows examples of indicators that did not meet one or more of these characteristics.

Table 2. Examples of sector or programme indicators that did not meet essential criteria (relevance, clarity, reliability, accessibility and affordability)

Country	Indicator	Comments
Kyrgyzstan	Informal payments as % of household out of pocket expenditure for hospitalisation	Information may be too sensitive and therefore hard to get. No baseline data is available.
	Percentage not-seeking care when needed due to financial and geographical reasons	Difficulties in assessing what constitutes “need”. No baseline data is available.
Bangladesh	Percentage of adults (age 15-55) who use tobacco disaggregated by gender and age	Progress against this indicator is likely to be slow and depending on factors outside what the ministry of health can realistically achieve in a 5-year sector programme. Disaggregation makes matters more complicated. No baseline data is available.
Ethiopia	Percentage of co-operating partners using a single set of procurement procedures	The phrasing is probably mistaken – does it refer to using a single or a common set of procedures? In addition, given the very different size and types of investments by various donors, is it realistic to envisage a single procurement procedure for them all?
Zambia	Proportion of research findings translated into policy and practice	How to count “research findings”? How long do they take to be translated? Is the research any good? Is this a central indicator in a sector programme (which is not a research programme)?
	Percentage of district hospitals with a functioning X-ray unit, and visited	The problematic part is “and visited”, which turns an acceptable input indicator into an impossible output indicator.
Mozambique	Percentage of central and provincial Directorates of Health (DPSs) with trained management staff	This indicator is supposed to measure capacity building efforts, but the definition of trained management staff is unclear. Not surprisingly information on this indicator has been never provided.
Tanzania	Proportion of public health facilities in a good state of repair	What makes “a good state of repair” needs to be clearly defined. There is no regular collection of information that would enable measurement. While one might like to study this as a basis for deciding whether and how to invest in upgrading facilities, this does not seem a useful way of monitoring annual progress.

2.4 Is there an ideal mix of indicators?

The relationships between governments, service providers, professional associations, civil society, community groups and development partners affect how health sector performance is measured, and reaching consensus around a monitoring framework can be lengthy and problematic.

The first countries to develop a SWAp spent a considerable amount of time trying to reach consensus on the numbers and types of indicators used to monitor the sector programme. In some cases the original selection was subsequently changed several times to accommodate new needs for information (with a tendency to *increase* rather than *decrease* the number of indicators over time). In the case of Cambodia, this has resulted in lack of clarity about how many and what indicators make up the indicators table.

The mix of indicators is critical since individual indicators change their meaning depending on context. Suppose an output indicator like “demand for curative services in public hospitals” declines over two consecutive years: is this a failure of hospitals (implying a bad result for this indicator) or is it a consequence of increased demand at the first level of health care, and therefore a good result? For how long should the indicator be allowed to fall? What level of demand for health care is reasonable? How much would imply overuse? Therefore, in order to measure the impact of health policies, modes of delivery and intervention strategies we need a mix of indicators, assessed over time.

Finally, the variety in number and type of indicators in our sample of countries suggests that there is no “ideal number” or “ideal mix”. However, although context will determine the specific mix, experience indicates that it should be possible to operate an effective sector monitoring framework with 20 indicators or fewer.

2.5 Choosing indicators: summary of lessons

The following issues and lessons are important for designing sector programmes and their monitoring tools.

- a) The “political climate” and trust between government and development partners plays a key part in shaping the numbers and mix of indicators that will be used.
- b) Some monitoring frameworks place greater focus on output and outcome indicators, while others emphasise the performance of the implementation agency – usually the ministry of health - through a greater emphasis on milestones. Each approach has its own limitations and risks: the milestones approach may become donor driven and remove focus from national programmes and health outcomes. Excessive focus on sector level outcomes may miss the importance of intermediate indicators such as service and programme coverage outputs. A balance in the mix of indicators seems the obvious solution.
- c) More indicators do not necessarily mean better monitoring – in fact, the opposite is often the case. A maximum of 25 to 30 indicators would seem to be enough for effective sector monitoring in the context of a sector programme. In selecting indicators there needs to be a balance between scope and depth, as often more information can lead to rather superficial analysis, which affects the effectiveness of sector reviews.
- d) Sector level indicators are not meant to replace other indicators used by implementing agencies to monitor national programmes for effective programme or disease control monitoring. On the other hand, sector monitoring is meant to reduce transaction costs to the government and to limit the information “needs” of development partners.
- e) There also needs to be a balance between ambition (how much information and how many indicators to measure) and capacity for information management and use within the implementing agency. All indicators must meet the essential (SMART) characteristics: be relevant, easy to interpret and understand, reliable, based on accessible data, and affordable.

- f) Capacity for use and management of information is always an issue in developing countries, so the more ambitious the monitoring indicators, the greater the investments that will be needed. This fundamental point is too often neglected and it is often the case that the countries with less institutional capacity for data handling are the ones using more indicators and the ones investing less in health information systems.

Box 2. Long term capacity development

Sector programmes can provide a strong incentive for improving the capacity of countries to handle and use health information. Data management capabilities have certainly improved in countries like Bangladesh, Ghana and Mozambique¹². In Nicaragua, after the launch of the SWAp development partners realised that there was a real need to strengthen the National Institute for Statistics, which, as a low priority for the government, had suffered from lack of funding and long term planning.¹³

3. The annual review

All sector programmes covered in our sample have some form of annual sector or programme review in place. Annual reviews provide an opportunity to assess progress against the strategic plan or programme on which the sector programme or SWAp is based. They are intended as an opportunity for collective reflection to keep the plan on course, and to rebuild momentum and inject energy into the sector partnership. Annual reviews usually complement other more in-depth reviews built into the sector programme, such as mid-term reviews or final evaluations. When properly planned and implemented, annual reviews represent a key opportunity for joint learning about the sector, about working together and about simplifying and unifying sector monitoring and reporting mechanisms. In fact, when annual reviews are part of a SWAp they are expected to replace parallel monitoring and reporting processes by development partners.

A common problem experienced by most countries in our sample is that sector reviews can become too broad and too ambitious in terms of the number of issues covered. When the scope is too wide the depth and quality of the analysis may suffer. This is certainly the outcome in countries where annual reviews have become a sort of annual in-depth evaluation, where too much is expected to be reviewed in too little time. In addition to their impact on the quality of information, “heavy” annual reviews have an impact on resources, both financial and in the form of transaction costs to government and sector partners.

However, the main casualty of overly ambitious annual reviews is in their follow up, since they tend to deliver long lists of issues and recommendations, often poorly prioritised, which result in unrealistic agendas for action. And the lack of proper follow up on these agendas may lead to the perception (often unwarranted) that the programme is not performing well. All this has a detrimental effect on morale and on the relationships between the implementing agency and its development partners.

There are important differences, issues and lessons from the ways in which annual reviews are conducted. They will be briefly reviewed in this section.

12. Veronica Walford, A review of health sector wide approaches in Africa. Paper prepared for DFID, published by the HLSP Institute, July 2007.

13. Elisabet Jané, personal communication.

3.1 What to review?

Annual reviews provide an opportunity for sector partners to look at:

- progress achieved on agreed goals, specific objectives and milestones contained in the sector programme;
- performance against the indicators that are part of the monitoring or results framework;
- progress against the recommendations made in earlier annual reviews;
- sometimes, a selected number of topics (usually one or two) agreed by sector partners in greater detail¹⁴.

While all sector reviews look at similar issues, they differ considerably in the approaches to data collection and in the amount of time and effort devoted to analysing each of the content areas. For example, some sector reviews focus primarily on sector or programme indicators, while others emphasise programme objectives or the follow up of the previous annual review recommendations. It all depends on how the process of the annual review is organised, as will be explained later.

Careful preparation of the contents of annual reviews is very important, and partners should refrain from adding new topics, particularly if these have received insufficient attention during the previous year. The rule of thumb should be that if a topic has not been effectively analysed or even discussed between government and development partners in the year preceding the annual review, then it should not be included in the review at all. In some cases issues which could be resolved during the year are unnecessarily saved up for annual or mid term reviews.

3.2 The annual review process¹⁵

Three main stages can be identified in the process of annual reviews: (a) the preparation stage; (b) the data collection and analysis stage; (c) the prioritisation of findings and recommendations leading to a plan of action.

a) Preparation

Some countries set up specific task groups combining government and development partners to prepare for and coordinate the annual review process.¹⁶ This seems a sensible approach as annual reviews tend to involve many stakeholders and can be complicated to handle. The main responsibilities of the task group usually involve:

- agreeing on dates and key milestones of the annual review (including, for instance, if the review will involve field visits);
- approving terms of reference for the review and for individual consultants who will make up the review team (when applicable);
- agreeing on skills mix and team composition, particularly if the review is to incorporate topics requiring more in-depth analysis;

14. This approach is common in countries like Bangladesh and Mozambique. Topics selected for in-depth coverage include human resources, compliance with the Paris Declaration, specific national programmes, emerging epidemics like avian flu, etc.

15. This section is tailored for the benefit of readers who are not very familiar with SWAPs and sector reviews.

16. Even when the management of the annual review is sub-contracted to a consulting firm, the task group acts as commissioning agency, clears TORs and confirms suitability of consultants.

- briefing and de-briefing review consultants, including making key documents available to the review team;
- acting as bridge between the broader constituency of the ministry of health, development partners and the review team on operational matters.

Access to key documents well before the review process is even started is essential for reviewers, and efforts to make key documents easily accessible are important. Several countries have chosen to make these documents accessible via the internet¹⁷ - again, a sensible and transparent approach.

b) Data collection and analysis

In most countries data collection is undertaken by external consultants. They produce the review report, which is then discussed by government and its partners. There are however slight variations of practice around the world.

- In Mozambique data collection is undertaken mainly by external consultants under the team leadership of a senior officer from the ministry of health. The advantage of this approach is that documentation and interviews can be more readily available to consultants. The main disadvantage is that the ultimate decision on what goes in the report is made by the team leader, a government official, which may introduce an element of bias.
- In most countries the review team is made up entirely by external consultants, including the team leader. The assumption is that this independent team has greater objectivity, but this objectivity must be matched by a good understanding of the country and sector by the team leader and key team members – this phase is often too short for complete outsiders to familiarise with the sector. Facilitation by the ministry of health can help the review team in this phase.

The number of consultants involved in annual reviews varies, from three to four in most African countries to seven to ten in larger Asian countries like Cambodia and Bangladesh. A rule of thumb is that more consultants do not necessarily guarantee a more in-depth analysis, unless they are able to integrate their findings before they leave the country. The danger is that sector review reports become a long list of findings and recommendations that are then difficult to prioritise and follow up. To avoid this situation it is important that the review team de-brief to a small group (for example the annual review task group mentioned earlier). Some countries also invite the team leader and national consultants to present their findings to a larger audience once the draft report has been cleared for wider dissemination.

c) Prioritising findings and recommendations

Finding “space” for sector stakeholders to look at, assimilate and discuss the information that has been generated is crucial. Two broad approaches can be identified.

- **Analysis in small technical groups followed by plenary.** In this approach several working groups or task groups arranged along thematic areas look at the results of the review and endorse what they see as main findings and recommendations. The report and the views from the groups are then discussed in a plenary. This approach (similar to the one used in Tanzania) has a clear advantage in terms of prioritising findings and recommendations but, critically, the quality of its outputs depends largely on the composition, dedication and competence of members of the working groups.
- **Separate analysis (by government and development partners) followed by plenary,** with prioritisation following in a smaller mixed group of government and development partners. This is by far the most

17. A good example of a well designed website for sector reviews is provided by the World Bank office in Bangladesh, see <http://www.hnpinfobangladesh.com>

common approach, wherein the report is widely shared and analysed from the different viewpoints of government and external partners. This approach (used in Mozambique and Bangladesh) can make prioritisation and preparation of the action plan difficult, particularly if these are expected to take place in the plenary. Plenary sessions are useful means of information sharing or consensus building but are less effective for analysis and prioritisation purposes, since they tend to be too large and too short, and not all participants may be technically competent. Attempting an action plan in a large group also risks polarising the debate (government versus external partners) and focusing attention on individual positions and differences rather than on what is needed for the sector and how to work together to achieve it.

In both approaches the plenary session or group work is usually followed by a meeting of a smaller group of stakeholders who agree on or endorse the recommendations and linked action plan.

It is hard to say which approach works best, since annual reviews are usually embedded in the working arrangements and code of conduct that regulate the sector partnership in individual countries. In fact, they rely heavily on positive working relationships, to the extent that for reviews to be effective they should take place in an environment where partners are willing to engage in open and constructive analysis and dialogue. When strains and tensions remain unresolved the quality of annual reviews is likely to be affected. Annual reviews may become grounds for confrontation, emphasising what does not work and what has not been achieved. In short, the “climate” and the dialogue between government and development partners preceding a sector review has to be positive enough to enable all parties to approach the review in a constructive manner.

Box 3. Who should participate in annual reviews?

Throughout this document we have been referring to “sector stakeholders” or “partners” involved in the sector programme, and hence in the annual reviews. Partners usually include the government (represented by one or more ministries) and representatives from the development partners (multilateral and bilateral) who provide official development assistance to the health sector. However, what about other agencies providing non-official development assistance (like the Bill and Melinda Gates Foundation, the Clinton Foundation, or the GAVI Alliance) and large international NGOs (like Oxfam or Save the Children)? Should they also be part of and influence annual reviews? And what about national civil society groups and national NGOs who operate in the health sector: should they also be part and, if so, in what capacity?

This is not an issue that can receive the attention it deserves here. The eleven countries in our review show very different approaches and policies in respect to who are the “legitimate” members or partners in a sector programme and how they should be involved in it. The practice in most countries is for sector programmes to become increasingly inclusive, enabling all stakeholders to be part of mechanisms for information sharing, including information on annual reviews.

Bangladesh has a specific “stakeholders forum” that meets with government to discuss the outcomes of annual reviews. Medicam, an NGO in Cambodia, has produced a sector position paper to formalise its position on sector priorities. Mozambique (just like other countries with a working SWAp) allows representatives of NGOs and civil society in the SWAp Forum and in the annual Sector Coordination Committee. The trend is certainly one of inclusiveness and openness, but this does not mean that the participation of all stakeholders in sector programmes and sector reviews has been effectively resolved. For example, a few governments are still unwilling to share information from sector reviews with civil society groups for fear that this will be leaked out and used to criticise government. Whether or not this is a real situation or a possibility, it is certainly a legitimate issue for sector programmes to resolve.

4. Follow up mechanisms

While a positive climate and mutual trust are essential for conducting annual reviews, it is the follow up that determines their success as a planning tool. Follow up is closely linked in the first place to whether review recommendations are realistic – in terms of number, scope and scale of the effort and resources that will be needed to see them implemented. Unfortunately, the track record of realistic, manageable recommendations is not good and, as a result, the follow up of recommendations from annual reviews in the eleven countries covered in this paper is often weak (see Box 4).

Box 4. Examples of follow up of recommendations and milestones set in annual reviews.

The Tanzania 2006 Joint Annual Health Sector Review highlighted that the linkage between sector reviews and budget support meant that the health sector would partly be judged according to achievement of milestones, and that performance against milestones to date had been poor. The Milestones Report recorded that of the fourteen milestones for the previous year, four were achieved, five were partially achieved, four were not achieved, and one milestone lacked the necessary information for measurement. Reasons given for this poor performance were that some of the milestones were too ambitious, some were poorly specified, and others did not have strategies and resources to address them.

In Bangladesh, consultants involved in the Annual Programme Review (APR) in 2006 reported that there had been hardly any follow up of the 2005 recommendations. To improve the situation a Coordination Committee was established with a mandate to follow up the APR, and by 2007 the follow up of recommendations had been much improved. Key to the improvement were the efforts made by all parties to reduce the overall number of recommendations from the APR, prioritise them, and make them explicit in a table of key follow up actions and responsibilities.

In Ghana, slight discrepancies among stakeholders have been reported in terms of the reasons for poor follow up of recommendations from annual reviews. It is said that reasons include the difficulties involved in dealing with the problems identified, perhaps suggesting that more realistic timeframes and targets for problem resolution need to be set.

Realistic recommendations may not ensure follow up per se but they are a key factor, much more important in our experience than the “lack of government ownership and leadership” that is often mentioned in many review reports to explain weak follow up. In fact, it could be argued that unreasonable recommendations and action plans emerging from annual reviews erode government leadership and affect staff morale. On the other hand, leadership can be enhanced through realistic recommendations and action plans.

4.1 Follow up of annual reviews: lessons learnt

The following issues have been found to affect follow up of annual reviews in most countries in our sample:

- **Too many recommendations in the report submitted by external consultants that are not sufficiently prioritised by the sector partners.** This is a very common problem, reflecting weaknesses in the discussions of the report delivered by the reviewers. For example, this problem emerged in Mozambique during the Mid-Term Review of the Health Sector Strategic Plan in 2005, where it was suggested that the recommendations of the annual joint review were simply “endorsed” by the Sector Coordination Com-

mittee without any real prioritisation. Lack of prioritisation can send wrong messages, including that the government accepts all the recommendations made by the reviewers, which is not actually the case in most programmes known to the authors. While it may not be possible to restrict consultants from making as many recommendations as they wish, government and sector partners should be more explicit with respect to those that they concur with, and to those which should receive attention during the following year.

- **The recommendations have been prioritised by sector (SWAp) partners, yet these are still either too numerous or unfeasible.** This is also a common issue, suggesting that there is too much pressure on the government to over commit without taking into account either the complexity of the problem or the capacity and track record performance of the implementing agency (ministry of health). Examples of this situation abound and include: over-simplistic recommendations to deal with complex human resource and staffing issues; recommendations that compete for scarce resources, such as expecting increases in coverage without provision being made for more resources or greater efficiency; or recommendations that are pushed by development partners against the will of the government. The rule of thumb is that it is better to under-commit and over-perform than the opposite.
- **The recommendations may be right in number and scope, and realistic, but follow up is weak.** The internal working arrangements within many ministries of health may not be up to speed in terms of enabling regular follow up of recommendations. When such arrangements are known to be weak, mechanisms need to be put in place to enable the ministry (with or without development partners) to review progress with implementation of annual review recommendations. It is also important to consider whether existing mechanisms for dialogue between government and its partners might accommodate a periodic and systematic review of annual review recommendations. For example, Bangladesh has greatly improved its follow up of Annual Programme Review recommendations in 2006 and 2007 through regular (monthly or bi-monthly) meetings of the so called Health, Nutrition and Population Coordination Committee, where follow up of recommendations is a fixed item in the meeting agenda. In Nicaragua, the creation of the Technical Committee linked to the *Mesa Sectorial* (the SWAp roundtable) may achieve similar objectives.
- **The same or very similar recommendations are delivered every year in the review reports, giving the impression that little or no progress is being achieved.** This situation is often a reflection of sector reviews that are too broad and cover the same areas every year, or of the fact that some systemic issues take longer to change than a one year cycle. This problem can be overcome by reducing the scope of the annual reviews (the suggested focus on “special topics” each year) and by improving the feasibility of recommendations.

In all these issues technical agencies (such as WHO, UNICEF, UNAIDS, UNFPA, etc.) can and probably should play a supportive role, ensuring that governments do not over commit and that technical and implementation weaknesses in the ministry of health receive attention through technical assistance or any other means. Bilateral partners may or may not have the technical skills required to fully appreciate the implications of certain recommendations made in annual reviews. They may therefore push for certain recommendations because these coincide with their own priority areas without giving due consideration to technical feasibility, including resource implications.

5. Conclusions and key messages

The existence of arrangements and tools for sector monitoring does not necessarily mean that sector monitoring is taking place as expected or that it is as effective as key stakeholders would wish.

There is often a perception that although sector reviews generate large amounts of information, the results are often seriously underutilised. This means that sector reviews in several countries no longer represent an opportunity for sector stakeholders to jointly and objectively look at progress.

Why is this happening? The reasons vary from country to country. They range from “fatigue” affecting the relationships between government and development partners, to situations where the sector monitoring framework has to compete with an increasing number of parallel indicators and monitoring arrangements. In most cases these “parallel monitoring systems” are linked to the new global health initiatives or to an increasing number of projects and vertical interventions supported directly by development partners.

However, even when their implementation arrangements are in clear breach of harmonisation and alignment principles, it would be wrong to put all the blame for the poor performance of sector monitoring frameworks on the new global and donor-led initiatives. In fact, there are fundamental problems which have affected sector programmes and sector wide approaches from the start. Among these is the gross underestimation of the efforts, time and investments required to establish an effective sector monitoring framework¹⁸, together with lack of capacity, or lack of real commitment (or both) on the part of some government departments (including ministries of health) to take sector monitoring seriously and to respond to the results of sector reviews.

Finally, some agencies that are formal partners in the SWAp have never really endorsed the practice of adopting a single monitoring framework led by government, and seem to have put more effort into highlighting the weaknesses of the “government’s” sector monitoring framework rather than contributing to its improvement. Clearly, this it has not helped the cause of unified sector monitoring.

Key messages

- Don’t underestimate the time and effort involved in establishing a sector monitoring
- Don’t overestimate government capacity to implement it
- Don’t expect all partners to buy into it (including government if indicators are seen as donor led)
- Keep the number of indicators down and don’t let their number creep up over time. Focus on what is important, not what is easily measurable. Don’t try to cover everything – a sector monitoring framework is intended to complement rather than replace existing monitoring systems. It is important to focus on donor as well as government performance.
- There is no ideal review process. The approach will need to reflect country circumstances. There are difficult tradeoffs and balances to be managed. Is it better to keep participation down to allow effective management of the review while also considering how to enable wider participation (involving for example civil society)? How independent should the review be? Country examples provide useful tips on how to go about the review process.
- Good preparation is essential and rapid feedback is helpful. Thought needs to be given to how the approach should be structured – to achieve a robust dialogue without running the risk of undermining or even embarrassing government.
- Annual and mid term reviews need to be kept focused. Broadening the scope results in extensive, unfocused and prioritised recommendations which are unlikely to be followed up.

18. Several countries in our review have reported investments in health and management information systems, and in the national offices for statistics, but these are seen as largely insufficient.