



Results-Based Financing for Health (RBF)

# Verification at a Glance

A Series of Snapshots of Experiences in Verifying Performance  
Linked to Financial Incentives for  
Results-Based Financing (RBF) Programs  
from Selected Countries

Editors: Joseph F. Naimoli and Petra Vergeer

*Results-Based Financing (RBF) Team*

The World Bank

*Afghanistan*

*Argentina*

*Burundi*

*Cambodia*

*Haiti*

*Rwanda*

*U.S.A.*

*GAVI Countries*

*Global Fund Countries*

*Global Program on Output-Based Aid Countries*

# Table of Contents

Background .....	4
Methodology.....	5
Findings .....	6
References .....	7
<i>At-A-Glance: Afghanistan</i> .....	8
<i>At-A-Glance: Argentina (Plan Nacer)</i> .....	9
<i>At-A-Glance: Burundi</i> .....	10
<i>At-A-Glance: Cambodia</i> .....	13
<i>At-A-Glance: Haiti</i> .....	15
<i>At-A-Glance: Rwanda</i> .....	17
<i>At-A-Glance: U.S. Medicare Program</i> .....	19
<i>At-A-Glance: GAVI</i> .....	20
<i>At-A-Glance: Global Fund</i> .....	22
<i>At-A-Glance: Global Program on Output-Based Aid</i> .....	23

## **Acknowledgements**

**The editors wish to thank the following people for their contributions to this effort.**

*World Bank:* György Fritsche, Rafael Cortez, Timothy A Johnston, Tekabe Ayalew Belay, Ghulam Dastagir Sayed, Mohammad Tawab Hashemi , Kavitha Viswanathan and Lars Johannes

*Johns Hopkins University:* Kate Gilroy and Olga Joos

*GAVI:* Peter Hansen

*Cordaid:* Piet Vroeg, Michel Bossuyt

*Ministry of Health, Cambodia:* Vijay Rao

*Management Sciences for Health, Haiti:* Bernateau Desmangles

## Background

The verification of performance linked to financial incentives in the health sector assumes crucial importance as more and more countries shift from financing health system inputs and processes alone to financing service delivery outputs and health outcomes as reflected in performance-based agreements or contracts. Furthermore, there is a concern that payments linked to performance may result in over-reporting of the performance (Lim et al., Lancet, 2008). In Results-Based Financing (RBF) schemes, verification is an essential element of program implementation and to date different schemes have adopted different approaches to verification.

*For some schemes, verification has been limited to ensuring the accuracy and consistency of reporting on the volume and/or quality of services provided.* These schemes usually rely on routine service delivery data, which are generated and recorded at the point of service provision, and eventually entered into the health management information (HMIS)<sup>1</sup> and/or another data management system. The verification process usually includes some assessment of the reliability of reporting by providers and/or their supervisors through some form of repeated measurement or “recount” of the original or source data. The “recount” data, which may be collected by an internal “controller,” a third party, or both, are compared with the source data, and financial incentives are provided when the discrepancy between the two datasets is found to be within an acceptable, pre-determined margin of error.

*Other schemes ensure reliable reporting and confirm that patients who were reported to have received services actually received them.* Confirmation that services have been received has been approached in different ways: for example, a random household spot check of a sample of patients drawn from health facility registers (i.e., patient follow-up) or a systematic, population-based household survey. Independent, external agents, acting alone or in collaboration with internal representatives of the health system, have been called upon to apply these different methods.

*Other schemes choose to rely upon direct observation of the conditions of service delivery and actual care rather than on self-reporting.* In these cases, independent, third party agents may carry out a health facility assessment or survey, which may include interviews with health

---

<sup>1</sup> One advantage to reliance of these schemes on the routine HMIS is that it may help to spur much-needed improvements (Gething et al., 2006) in information systems and data quality, particularly when financial and technical support are provided to obtain these improvements. Consequently, in these schemes an assessment or test of the “goodness of fit” of the HMIS may be carried out at the design stage to ascertain the extent to which current systems are providing relevant and high quality data. These assessments also attempt to ensure that procedures for collecting, processing, analyzing, and reporting data are adequate to minimize risks to data quality. The findings from these assessments may result in remedial actions prior to and possibly during implementation and verification.

staff and patients, direct observation of preventive and/or curative care, and an audit of management practices, equipment, supplies, and information. Incentives are provided when the overall performance of the health facility or providers, often quantified in a composite “score,” is judged to be in accordance with established benchmarks or is equal or superior to the pre-determined performance standard. These schemes may or may not include patient follow-up.

*In some cases, countries have adopted a combination of these different approaches.*

At present, recommending definitive best practice in verifying results linked to financial payments is difficult. A recent data quality audit experience of a performance-based inter-governmental transfer scheme (Ronveaux, 2005) underscores the challenges associated with this task. In support of its performance-based strategy of rewarding governments \$20 for each additional child vaccinated above an agreed-upon target, GAVI instituted the *Data Quality Audit* or “DQA.” The DQA was a tool for independently assessing the proportion of nationally reported 3<sup>rd</sup> dose diphtheria-pertussis-tetanus (DPT3) vaccine administered that could be verified by written documentation at health facilities and districts. Two independent companies conducted twenty-five country audits during 2002-2003. In sixteen of these, the proportion of verified doses administered was less than 85%. Moderate over-reporting (verification rates of 70%-84%) was documented in seven of these sixteen audits, while considerable over-reporting (verification rates of < 70%) was documented in nine.

According to a recent stocktaking of World Bank HNP projects with RBF components implemented between 1995 and 2008, no project adequately documented implementation processes (Brenzel et al., 2009). Consequently, very little is known about the mechanics of performance verification linked to financial incentives. Even in well-documented demand-side schemes, such as Conditional Cash Transfers (CCTs), explicit and detailed information on the verification process has been less available than information on other aspects of CCT design, implementation, and effect (Fiszbein and Shady, 2009). To better understand verification, we undertook a rapid review of selected recent experiences.

## **Methodology**

*Criteria for country selection.* We began with the afore-mentioned Stocktaking Paper on RBF by Brenzel et al., which identified twenty-four Bank-supported projects in seventeen countries with substantial RBF efforts during FY 95-08. We contacted the Bank Task Manager from most of the projects to ascertain whether the scheme:

- had a supply-side component,
- 2) included a verification process,
- 3) could provide written documentation of the process,

- 4) could specify the type of verification method used,
- 5) had completed at least one round of verification in the last twelve months,
- 6) could provide access to the agents who had carried out the verification exercise, if necessary, and
- 7) was suitable to include in our inventory, in the manager's opinion.

*Sample selection.* Based on these criteria, we selected for further investigation four countries—Afghanistan, Argentina, Cambodia, and Rwanda—all of which transferred funds from one level of government to another. In addition to Bank-supported countries, we decided to include the performance-based experiences of two NGOs in low-income countries (Cordaid in Burundi and Management Sciences for Health in Haiti) supported by the Netherlands and the European Commission, and USAID, respectively. To further enhance the diversity in our convenience sample, we included one industrialized country (U.S.A.) with substantial experience in the verification of claims invoices and which adopted a post-payment (ex post) verification approach (Medicare).

We also decided it would be valuable to explore the experience of performance-based Global Health Partnerships (GHPs), in particular GAVI and the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM), both of which provide funds to national governments conditional on performance. These partnerships have developed data quality audit tools that might be relevant to the intra-governmental transfer experience. We also selected the Global Partnership on Output-Based Aid (GPOBA), which aims to improve the delivery of basic infrastructure and social services, including health services, to the poor.

*Data collection and analysis.* We decided to limit description of each of these experiences, at this stage, to a brief “snapshot.” We drafted an initial snapshot that included a rationale for why verification had been introduced and some minimal description of how the process was carried out, where, by whom, using what methods, and, with what effect. After several revisions of this prototype, we used it as a model to write other snapshots. Some snapshots were drafted based on a review of existing documentation and discussion with key informants. For others, we identified collaborators from selected countries and asked them to draft similar snapshots based on the model we provided. We reviewed and revised all drafts we received, returned them to the collaborators for further review and clarification, where necessary, and finalized each snapshot. We also asked collaborators if they would be willing at a future point in time to work with us in developing more robust case studies in accordance with a standardized, structured outline. Most collaborators responded favorably.

## Findings

Together, these snapshots offer a rich, albeit preliminary, collection of experiences with verification, in a diverse range of settings, under different financing mechanisms. The collection is intentionally indicative rather than exhaustive, and raises many questions that we plan to answer in subsequent stages of analysis.

## References

Brenzel L et al. (2009). Taking Stock: World Bank Experience with Results-Based Financing (RBF) For Health. Washington, D.C.: The World Bank.

Fiszbein A and Schady N (2009). Conditional Cash Transfers: Reducing Present and Future Poverty. Washington, D.C.: The World Bank.

Gething PW, Noor AM, Gikandi PW, Ogara EAA, Hay SI, et al. (2006). Improving imperfect data from health management information systems in Africa using space-time geostatistics. *PLoS Med* 3(6): e271 doi:10.1371/journal.pmed.0030271.

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

Ronveaux O, Richert D, Hadler S, Groom H, et al. (2005). The immunization data quality audit: verifying the quality and consistency of immunization monitoring systems. *Bulletin of the World Health Organization*, 83, 503-510.

## *At-A-Glance: Afghanistan*

---

In 2003, the government of Afghanistan received a \$59.6 million grant from the World Bank for a three-year Health Sector Emergency Reconstruction and Development Project. Most of the grant funds were intended to expand rapidly the delivery of a basic package of health services (BPHS) aimed primarily at women and children in seven under-served and primarily rural provinces (population approx. 2.7 million). This was to be accomplished through performance-based partnership agreements (PPAs) between the government and NGOs (approximately 67% of total grant funding), and by strengthening Ministry of Health (MOH) efforts to improve service delivery. The government signed three-year PPAs with seven competitively selected NGOs (chosen through a quality and cost based procedure), who were responsible for implementing the BPHS in the provinces.

Winning bidders received the bid amount and were eligible to receive a performance bonus of up to ten percent of the value of the contract for good performance on a series of service outputs over the life of the contract period. The contract period was initially for three years, which could be extended subject to the availability of funds. Service outputs were assembled in a “Balanced Scorecard” (BSC), which included twenty-nine core indicators and benchmarks representing six domains of health services (patient perspectives, staff perspectives, capacity for service provision, service provision or technical quality, financial systems, and overall vision for the health sector), together with two composite measures of performance (percent of upper and lower benchmarks achieved). The indicators were selected during a series of workshops and discussions with the MOH, NGOs, and other development partners, including front-line health workers and managers. For each indicator, upper and lower benchmarks were set to indicate levels that were achievable in Afghanistan.

Each year an independent 3<sup>rd</sup> party agent (Johns Hopkins University with its partner the Indian Institute of Health Management Research) collected data to calculate these indicators and benchmarks through a health facility assessment. Data collection methods included observation of patient care, exit interviews with patients, interviews with health workers, and observation of structural aspects of facilities (including availability of staff, equipment, drugs, laboratory functionality, etc.). The assessments were conducted under the auspices of the MOH’s Planning, Monitoring and Evaluation and HMIS Unit. Facilities were rated on a scale of 0-100, and further categorized as high, medium or low performers. The annual performance bonuses were paid on the basis of achieving quality of care targets within these six domains, which included both quantitative and qualitative measures of performance. To earn a bonus, an NGO had to demonstrate a ten-percentage point or more improvement over its previous score on the assessment. Results of the assessments were publicly available.

*Source: Tekabe Ayalew Belay, Ghulam Dastagir Sayed, Mohammad Tawab Hashemi and Kavitha Viswanathan (World Bank)*



## *At-A-Glance: Argentina (Plan Nacer)*

---

Plan Nacer, the provincial social insurance program that targets uninsured pregnant women and children under six years of age, is a performance-based financing project funded by the government of Argentina and the World Bank, through an Adaptable Program Lending (APL) Fixed-Spread Loan. The objective of Plan Nacer is to provide an estimated 80 services free of charge to pregnant women and mothers (up to 45 days after delivery), as well as children under the age of six who currently are not covered by health insurance. Phase I was launched in 2004 in the nine most impoverished provinces in the country. The World Bank provided an APL of US\$135.8 million to cover the estimated cost of US\$289.9 over five years. Phase II was launched in 2007 in the remaining 15 provinces with an APL of US\$300 million to cover the estimated cost of the US\$646.3 million over five years.

Of the funds transferred to the Provincial Government, 60% are “Monthly Base Transfers” determined by the number of eligible beneficiaries enrolled in Plan Nacer. Each month, the provincial insurance unit verifies enrollment eligibility, in accordance with established criteria, which are cross checked against other social insurance databases to ensure there are no duplications. This process is repeated at national level. Once completed, payments are transferred from the national Ministry of Health (the National Health Services Purchasing Team= NHSPT) to the provincial health insurance unit (the Provincial Health Services Purchasing Team), which subsequently reimburses the providers based on a fee-for service basis (i.e., for the agreed upon 80 services).

The remaining 40% of the results-based financing funds are “Complementary Transfers”, which are determined by the achievement of stated targets for ten output and outcome health indicators or tracers<sup>1</sup>. Tracer targets are calculated on the basis of the previous year’s results. They are agreed upon annually by the national and provincial levels and described in the Annual Performance document. Tracer results are calculated every quarter by the provincial insurance unit based on clinical information.

Quarterly audits are carried out in each province by independent agents contracted by the NHSPT. The eligibility of the reported population is analyzed by repeating, independently, a cross-check of the enrollee bills against the social security databases. In addition, the auditor takes a random sample of the registrations (enrollment and tracers) to verify, in the field, the existence of enrollees, their voluntary enrollment status, the existence of clinical histories, and other supporting information for the tracers. Finally, the auditors are responsible for verifying the correctness of payments made to providers. Reports of the audit findings are submitted to the NHSPT. Penalties apply for inaccurate reporting: the per capita amount is deducted during the next cash transfer for each ineligible beneficiary, with an additional 20% of the per capita amount deducted as a penalty.

*Source: Kate Gilroy and Olga Joos (Johns Hopkins University); Rafael Cortez (World Bank)*

---

<sup>1</sup> Indicators are: i. Timely inclusion of eligible pregnant women in prenatal care, ii. Effective neonatal/delivery care (Apgar Score), iii. Effective prenatal care and premature birth prevention (weight above 2.5kg), iv. Quality prenatal and delivery care (mothers immunized and tested for STD), v. Medical auditing maternal and infant deaths, vi. Immunization coverage (measles), vii. Sexual and reproductive health care, viii. Well child care (<1 year), ix. Well child care (1-6 years old) x. Inclusion of Indigenous Population

## *At-A-Glance: Burundi*

---

In 2006, the Dutch NGO Cordaid introduced and supported Performance-Based Financing (PBF) pilot projects to improve basic health care in two provinces in Burundi with financing provided by the government of the Netherlands. The national government adopted the approach, which Cordaid is now implementing in 7 provinces<sup>2</sup> with funding from the EU, GAVI and Cordaid. The Burundi PBF approach is based on the separation of three functions: service provision, health system regulation and purchasing (which includes verification); however, Cordaid has not been able to adhere to this separation principle in its strictest sense because of the guidance and training roles that it was forced to assume in introducing PBF.

Cordaid created Local Fund Holding agencies (FHAs), which assumed, once trained, full responsibility for introducing PBF (including training of all actors involved and introducing various PBF tools and data collection instruments) and supporting its implementation. The FHAs contract individual health facilities for a set of well-defined services<sup>3</sup> and verify both service quantity and quality in health centers and hospitals as a condition for releasing performance incentives.

### Verification of service quantity

During the first five days of each month, using data from consultation records, health facilities (health centers and district hospitals) prepare two reports simultaneously: a monthly health management information system (HMIS) activity report, which they forward to the District Health Office (BDS), and a summary report of contracted indicators (an “invoice”), which they forward to the FHA. The FHA auditors, who are independent of the local health system and government, visit each public health facility monthly. They verify the consistency of the data reported on the monthly summary reports by reviewing the records of the health facility (and any sub-contracted facility) and recounting the number of services registered for the specific indicators.

Based on the data approved by the FHA auditor and facility staff, an invoice is prepared<sup>4</sup>. The invoice is countersigned by the person in-charge of the health facility, verified by the FHA financial controller,

---

<sup>2</sup> In two other provinces Health Net International TPO and the Swiss Development Cooperation also support PBF projects, with a similar operational structure to Cordaid’s model. During 2010, a new national PBF model, fundamentally based on Cordaid’s model, will be implemented nationwide.

<sup>3</sup> At health center level, fifteen to twenty services are purchased, whereas twenty services are purchased at hospital level. The contracts are based on a fee-for-service model wherein an additional bonus can be earned based on performance against quality criteria. Contracted health facilities are encouraged to subcontract with the private sector (both for-profit and not-for-profit groups) in their catchment areas. These arrangements can result in additional funds for both parties.

<sup>4</sup> Separate formats exist for the basic package of health services with its corresponding indicators and the comprehensive package of health services.

and approved by the FHA director before funds are transferred to the account of the service provider. This monthly invoice is prepared in triplicate: one copy is kept by the health facility, a second copy is sent to Cordaid's financial department in Bujumbura, and the third copy is kept by the FHA for payment, filing, and reference. The audit of the records may reveal discrepancies with the data reported in both monthly reports, which provides an opportunity to correct the HMIS reported data by triangulating those data with the PBF indicator data.

The FHA contracts one local community organization<sup>5</sup> for each health facility to carry out additional verification. Each quarter, four members of the local organization will trace 60-90 patients registered in the health facility to verify that these patients exist and have actually received the services. Patients are selected randomly by the auditor of the FHA, for each service to which an incentive is attached. The greater the volume of patients registered for a certain service, the greater the number of patients who will be traced for that particular service. The members of the local organization are required to collect certain data from the patients (i.e. number of births), of which FHA auditors are cognizant, as a means of verifying the members have actually visited the household.

#### Verification of service quality

Service volume is weighted by the quality of the services provided; health facilities can receive a maximum bonus of 10% of the service volume. Community perceptions of quality as well as technical quality, which are verified every three months, are taken into consideration for the bonus calculation<sup>6</sup>. The FHA also contracts local community organizations to assess community satisfaction with the services provided by health facilities, and to convey any suggestions for improvement, in addition to patient tracing. The FHA organizes this activity, including data collection and processing. A "perceived quality score", which ranges from 0 to 100 percent, is produced for each facility.

The Provincial Bureau of Health in cooperation with the District Health Office, evaluate the technical quality of services provided by health facilities. They operate under a performance-based contract with the FHA, in which they agree to carry out their work correctly and in a timely manner. A grid of quality components is used, which are divided into various indicators, with values from 1 to 5 points. Many of the indicators are composite indicators (i.e., they combine a number of elements that all must be present or rated positively to obtain the established point or points). This grid is used to rate or classify

---

<sup>5</sup> Eligible organizations are those that have been registered with the Ministry of Internal Affairs or those which have been recognized by the local community and have existed for at least 2 years. Preferably, the organization serves a community purpose. The organization cannot have any links with the health facility so as to prevent any conflicts of interest.

<sup>6</sup> Fifty percent of the bonus is based on the technical quality determined by health authorities, with the precondition that none of the quality elements assessed may score below 60%. The other fifty percent is based on other elements, namely: 10% for quantitative verification of documents, 10% for existence of registered users, 10% for the provided service, 5% for average cost of subsidized services, 5% for cost perception by users and 10% for users' satisfaction with services provided by the facility.

the various activities of the health centers as compared with standards. This evaluation yields a “technical quality score” ranging from 0 to 100 percent.

Hospital technical quality is evaluated through a Peer Review mechanism, which also uses a grid. A small team of senior staff from at least two other hospitals (including the chief medical officer, the administrative and financial director, and the director of nursing) together with representatives of the District Bureau of Health and the FHA evaluate the provision of services by another district hospital. During the evaluation, doctors evaluate clinical aspects. Administrative and financial directors evaluate aspects of management, organization, and finance. Directors of nursing evaluate hygiene and quality of care. The visit is conducted over one day, and carried out in two phases: 1) the evaluation, using the grid and 2) presentation to the providers of the results, with recommendations for improving quality.

This peer evaluation yields a “technical quality score” that can range from 0 to 100 percent. Any score below 50 percent is considered inadequate. In these cases, peer evaluators must identify the factors contributing to this low score and propose urgent actions for improvement, including administrative measures, if necessary. Any component on the grid receiving a score of less than 60 percent also must be analyzed and corrective actions proposed. During an immediate presentation of the results to the providers in the hospital concerned, the team of peers suggests how the quality bonus can be used to achieve rapid improvements.

\*\*\*\*

According to Cordaid, the rigor of the audit system (and its continuous improvement) has contributed to substantial gains in the quality of routine health management information data. Monthly tracking ensures the availability of reliable, auditable data that is collected using transparent methods. Reliable data have facilitated realistic planning of activities by the authorities. It is believed that the audit team’s independence has contributed to greater transparency and neutrality. Cordaid reports that there is almost no political pressure on the team and the turnover rate of team members has been low. Furthermore, Cordaid notes that the threat of severe sanctions (withdrawal of 20% of subsidies provided by Cordaid, possible cancellation of contract, etc.) in the case of misreporting has discouraged such practices. In addition, Cordaid is of the opinion that because the FHA team reports directly to the donor and is always aware that its own work will be audited, it has a vested interest in establishing mechanisms that prevent providers from generating false data.

*Source: Piet Vroeg, Michel Bossuyt (Cordaid)*

## *At-A-Glance: Cambodia*

---

In 1999, Cambodia's Ministry of Health (MOH) first contracted with NGOs to provide health services in selected health Operational Districts (ODs) under the Basic Health Services Project, which was funded by the Asian Development Bank (ADB). The success of this first effort led to the incorporation of NGO contracting into the successor project—the Health Sector Support Project 2003-09—which was jointly funded by ADB, The World Bank, DFID, and UNFPA. A total of eleven ODs were contracted out to seven international NGOs based on competitive bids received from local and international NGOs based in Cambodia from 2004 to early 2009.

Each contract was a lump sum agreement for an initial five-year period, with the possibility of extension. Each contract stipulated OD coverage and utilization rates that were expected to be achieved by the contract's end for a range of reproductive, maternal, newborn and child health indicators. Financial penalties were envisioned for those contractors who failed to achieve their targets. Based on contract targets, each NGO contractor introduced either staff-based or health facility-based performance incentive schemes in their respective ODs. The purpose of the incentives was to motivate low-paid staff to perform key tasks that would contribute toward achievement of stipulated targets. The MOH did not impose a standardized scheme; consequently, each contractor was free to design and operate schemes that suited their local context. Remuneration rates varied from one NGO to another.

Verification arrangements were established at two levels within the MOH: central and provincial. A third level of verification was established by NGO contractors at health facility level, although the schemes they adopted varied across the ODs. At the central level, a high-level Monitoring Group (MG) was set up initially under a Deputy Director General and subsequently under a Secretary of State. The MG consisted of senior staff from the Department of Planning and Health Information, the Department of Budget and Finance, and the Department of Personnel. The provincial health directors of each of the 7 provinces with contracted ODs were also members of the MG. The Project's Contracting Specialist served as secretary to the Group. Responsibilities of the MG included dispute resolution between contractors and ODs, reviewing quarterly reports from and the performance of contractors, and authorizing the release of quarterly payments to the contractors.

Based on the monitoring scheme included in the Project design, quarterly field visits were conducted for verification purposes. In the initial phases of the Project, a team of 2-3 MG members from the central MOH level participated in these visits; however, due to the increasing burden of routine departmental tasks, midway through the Project they could no longer participate. From that point on, only the Contracting Specialist represented the central level during field visits, which he conducted together with provincial-level staff. The visits to each OD involved verification of coverage and utilization figures from the Health Information System (HIS). Special attention was paid to the target population denominators used for calculating indicator values. In addition, household surveys were conducted in two villages each in the catchment areas of two health centers. Household interviews were completed with all mothers of infants under the age of one year resident in the village. Results of the two village surveys were then presented to the MG at the next quarterly meeting. Health centers were also visited

to examine staff schedules, assess attendance, and observe the quality of care provided. Feedback from each field visit was provided to both the OD management team and the contractor.

In addition to the combined MOH central-provincial monitoring team, each contractor also established an M&E unit, whose members conducted intensive monthly verification visits to all health facilities operating in the OD. These visits included verifying staff performance, conducting spot checks of patients drawn from facility registers, and calculating performance scores to determine which staff members were entitled to receive incentive payments. Wherever discrepancies in register were noticed, the contracted NGO applied appropriate penalties to facility staff. Each NGO had strong incentives to keep costs down through intensive monitoring because of the lump sum nature of the contract. This appeared to contribute to the high quality of record keeping found at facility level. A final evaluation of contracting, which relied on individual household surveys in each of the 11 ODs in the first quarter of 2009, found substantial agreement with HIS figures and confirmed that verification arrangements had been successful.

*Source: Vijay Rao (Ministry of Health), Timothy A Johnston (World Bank)*

## *At-A-Glance: Haiti*

---

In 1999, USAID, with the assistance of Management Sciences for Health, piloted a performance-based financing scheme with three non-governmental organizations (NGOs), which provided health services to approximately 534,000 people. Originally, NGOs were reimbursed for documented expenditures. This approach was subsequently changed to one in which NGOs were paid, in part, based on the achievement of performance targets or outputs. The approach has further evolved and has been expanded to include twenty-five NGOs that were considered as of 2005 to be ready to implement a performance-based approach.<sup>1</sup>

At the outset of the pilots, an independent survey research firm collected pre- and follow-up data. Baseline and follow-up data on immunization coverage in each NGO service area was collected through household surveys using a standard cluster sampling methodology. Facility-based exit interviews were conducted with women to determine the proportion that were using Oral Rehydration Salts (ORS) to treat childhood diarrhea. Household interviews were carried out and administrative records were reviewed to determine pre- and follow-up coverage of pregnant women having completed three or more prenatal visits. Rates on discontinuation of oral or injectable contraceptives were ascertained by reviewing health facility registers. Other variables, such as “quality of care” and “degree of collaboration with local health authorities” proved more difficult to measure accurately and were omitted in the next phase of the program.

Certain methodological constraints made it difficult to determine whether follow-up results in the service areas were different in a statistically significant way from the baseline. When results were below the pre-determined target, but within one confidence interval, the NGO received the bonus. As a result of these constraints, the verification of performance was modified in subsequent phases of the program, as described below.

Population-based surveys, which were expensive and methodologically problematical, were replaced by a scheme that relied on ensuring both the reliability and validity of routinely reporting service delivery data by NGOs. To ensure accurate reporting, the project team together with an independent firm compared facility registers with the summary data reported to the project management team (reliability check). To ensure that women and children actually received services that they were reported to have received, a sample of clients were selected from the registers and visited in their households (validity check). The findings from these two methods were subsequently discussed with the NGOs and health facility staff. The main cause of discrepancies was reported to be a difference in the dates of service delivery. Moreover, under-reporting proved to be more prevalent than anticipated over-reporting.

---

<sup>1</sup> A determination of “readiness” was made through an assessment of the quality of services that NGOs provided, and organizational capacity, such as the quality of financial management and the functioning of the health information system.

In the next phase of the program, an expanded list of service delivery indicators<sup>2</sup> was adopted with the aim of ensuring NGOs would improve results for multiple interventions, while a random selection of indicators was verified to ensure verification costs were kept to a minimum. Management indicators, such as drug management or the timely and correct submission of reports, were also included to strengthen these functions. An independent local firm together with the project team identified the indicators to be used and evaluated performance throughout the year. Bonuses were made available immediately following performance review rather than waiting for the annual service assessment. In subsequent phases, the project team alone verified the management indicators.

*Source: "Performance-Based Incentives for Health: Six years of Results from Supply-Side Programs in Haiti" (2007) by Eichler, Auxila, Antoine and Desmangles; Bernateau Desmangles*

---

<sup>2</sup> Uniform targets were set for the package of indicators for all NGOs contracted, which was revised in 2006 with the use of customized performance targets.



## *At-A-Glance: Rwanda*

---

In 2006, the Rwandan government adopted Performance-Based Financing (PBF) for use throughout the country. The design and implementation of the national program was shaped by the lessons drawn from 3 pilot projects, which were supported by three different actors (Cordaid and Health Net International TPO since 2002 and the Belgian Technical Cooperation since 2005). The national program was carried out in the context of a concomitant and far-reaching decentralization, which included the creation of a new district health department, which was responsible to the local government administration (not to the Ministry of Health). The Government is the most important purchaser of health services in the national program, which is a departure from the pilot schemes, in which NGOs or bilateral actors were purchasing agents. There are four PBF models in Rwanda: the health center, the district hospital, the community and the central MOH models. This snapshot describes the health center model, which is based on a fee-for-service system that is conditional on quality.

Verification of results is an integral part of the national program. Its principal aim is to ensure that results are accurate prior to the release of incentive payments to health facilities (ex ante verification). It also aims to enhance transparency and accountability at all levels. There are two types of verification processes or controls: internal (ex ante) and external (ex post). A brief description of each follows.

*Internal.* The health facility management reports the quantity or volume of services (primarily preventive) provided in a month in a PBF invoice. There are about twenty-four services that are purchased: 14 from the basic package of health services, and ten HIV services. The facility in-charge and the President of the Health Management Committee, a community representative, together confirm the accuracy of the invoice and sign it before sending it to district level. At district level, a health “controller” from the local government office visits the health facility to ascertain the accuracy of the invoice by comparing it to the data in the registers. Each purchased service has its own primary register, such as a Voluntary Counseling and Testing (VCT) register. Secondary registers, such as a laboratory register for VCT testing), can be used when problems arise with the primary register.

Once a quarter a member of the District Hospital supervisory team, using a checklist<sup>1</sup>, assesses and scores the performance related to the quality of the conditions to provide care. The reason for separating the internal verification of quantity and quality of services is to ensure the involvement of both local government authorities and the district health management team, with the aim of lessening the potential for conflict of interest and to ensure a balance of power in the district health system. The clinic invoice data and summary results from the quality checklists are then entered at district level into a web-based, real time PBF management information system to calculate entitlements. The system also

---

<sup>1</sup> The checklist was adapted from a tool developed by an NGO during a pilot experience. It was based, to a large extent, on existing supervisory forms in use by various vertical disease control programs. The checklist currently consists of a series of composite indicators and measures of quality across 14 services. The checklist evolves each year in response to new norms and standards as they become available, lessons learned from using the checklist, the perception of end-users and technical assistants in the PBF program, and on-going improvements in quality of care.

provides a set of tools that allow comparisons of district performance. Each district can view the performance of other districts.

The District PBF Steering Committee (made up of representatives of civil society, technical assistants from NGO/fund holders, MOH district and local government, the district AIDS commission, and delegations of public and faith-based managed health centers) meets quarterly to reconcile the electronic and paper versions of the data. The Steering Committee also discusses numerous issues related to the performance of the health facilities as well as other health-related matters. Upon a satisfactory reconciliation, the Committee sends a request for payment to the relevant fund holders.

Following a rapid confirmation that all procedures have been followed appropriately, the fund holders pay the invoice. Upon receipt of payment, the Committee deposits the funds into the bank accounts of the individual health facilities. The payment cycle is quarterly. The health facilities follow standard rules and regulations that help them convert these earnings into performance bonuses, which they distribute monthly. The Committee is held accountable for its actions through a multi-lateral contract with the district mayor. The Committee has become the most important decentralized district planning platform for health in Rwanda.

*External.* Every quarter, a third-party agent, contracted by one fund holder<sup>2</sup>, validates that services reported to have been delivered were actually received by patients (ex post verification). The agent applies a standard protocol that incorporates a multi-stage, random sampling methodology. Districts and facilities are chosen, using a random number generator, during a plenary meeting with representatives of the MOH and civil society. Four (of 30) districts are randomly selected and 25% of health facilities in these districts are chosen.

At the health center level, three services from the basic health package and three HIV services are chosen randomly (from a total of 24 purchased services). Six services from the basic packages are chosen when no HIV services are provided. Using the primary patient registers, six month's worth of services are selected, and 15 clients are randomly selected. The agent compares "reported" services (drawn from the registers) with "paid" services (drawn from the electronic invoice system). A grassroots organization, preferably consisting of people living with HIV, is selected from the catchment area of the health center (according to a set of objective criteria) to follow patients.

For each client traced and interviewed in the community, the organization receives \$2. Data are compiled and entered in a database (EPIINFO). Feedback is provided at community, district and central levels. The overall level of misreporting has been low: less than 5% of 900 clients visited each quarter cannot be traced in the community. Corrective actions, such as the firing of a health center in-charge in cases of misreporting, have been taken. Semi-annually, the degree of accuracy of the quality checklist is also verified (ex post). The evaluation is conducted by a group of technical assistants from a national coordinating body, which is predominately staffed by non-state actors.

*Source: György Fritsche, World Bank*

---

<sup>2</sup> The fund holders are mandated by the MOH to contract a third party and the MOH is involved in the negotiations of the output-based contract with the third party.

## *At-A-Glance: U.S. Medicare Program*

---

The Social Security Act established the Medicare program in the U.S. in 1965. Medicare covers the health care needs of people aged 65 and over, the disabled, people with End Stage Renal Disease, and others who elect to purchase Medicare coverage. Medicare costs and the number of beneficiaries have increased dramatically since 1965. Benefit outlays through both the Fee-for-Service (FFS) and managed care programs were approximately \$381.8 billion in 2006, representing almost 15% of the total federal budget at that time. The FFS program uses a network of contractors to process more than a billion claims each year, make payments in accordance with Medicare regulations to some one million health care providers, and educate providers about how to submit accurate claims. The sheer volume of claims submitted (approximately 4.5 million/work day) and budget constraints that limit claim reviews require contractors to pay most claims without examining the medical records associated with the services listed in the claims. Despite various actions over the years by the Centers for Medicare and Medicaid Services (CMS) to protect the integrity of the program and to limit improper payments, problems persist. The Office of Management and Budget reported improper payments totaling \$10.8 billion in 2007. The primary reasons for improper payments include paying for services that are not medically necessary, incorrectly coded, or insufficiently supported by appropriate documentation. Other errors include applying outdated fee schedules, paying for duplicate claims, or paying claims that should have been paid by a different health insurance company.

To safeguard the use of federal funds and limit improper payments, the CMS has established a multi-faceted, post-payment system of verification, which includes diagnostic, remedial, and recovery programs. The *Comprehensive Error Rate Testing Program* produces error rates and estimates of improper payments to evaluate contractor and program performance retrospectively. Using the error rates to identify where problems exist, the *Error Rate Reduction Plan* (ERRP) attempts to target problem providers and improve their efforts, primarily through further clarification of CMS policies, research, and education-related initiatives. The *Program Safeguard Contractors Program*, established in 2006, appointed specialized fraud fighters who investigate all provider and supplier types nationwide and may refer cases to law enforcement, among other actions to control waste and abuse. In response to growing concern that these efforts may have been necessary but not sufficient to protect adequately the Medicare system against improper payments, Congress mandated the *Recovery Audit Contractor Program* (RACs), which was introduced in a limited number of states in 2008 after a successful 3-year demonstration project. The major finding of the demonstration project was that independent auditors could detect and correct (i.e., recover) past improper payments. As of March 2008, RACs had succeeded in correcting more than \$1.03 billion in improper payments, 96% of which were overpayments. Congress has required CMS to make the program permanent and nationwide by no later than January 1, 2010.

*Source: The Medicare Recovery Audit Contractor (RAC) Program: An Evaluation of the 3-Year Demonstration (2008), Center for Medicare and Medicaid Services*

## *At-A-Glance: GAVI*

---

GAVI's "Immunization Services Support" window, launched in 2000, is a performance-based reward program in which eligible countries (annual GNI per capita < \$1,000 in 2003) can receive US \$20 for each additional child vaccinated with the third dose of the diphtheria, pertussis, tetanus vaccine (DPT3) above an agreed upon target. The reward is calculated annually, based on the number of children vaccinated with DPT3 as reported in the WHO Joint Reporting Form (JRF), which is reviewed by WHO and UNICEF. Before a country can receive a reward, it must "pass" a Data Quality Audit (DQA).

The DQA assesses the accuracy, quality, timeliness and completeness of the routine immunization reporting system, and audits the reported DPT3 coverage in a specific calendar year. The standard two-week audit is carried out by two external auditors and national/district counterparts, who visit the national level, four districts, and six health centers per district, which are randomly chosen. A representative sample of the tally sheets and registers of the previous year are reviewed and the results discussed with health staff. A preliminary report is made at the end of the audit and a country plan of action in response to the recommendations is developed. The DQA costs approximately \$50,000 to conduct.

The key summary measure produced by the DQA—the "verification factor (VF)"—summarizes the level of consistency between DPT3 records and reports audited at facility and district level. A VF of 1.0 represents a perfect match between records and reports. Countries that obtain a VF greater than or equal to 0.8 are eligible to receive an annual reward if they meet the other criteria; a VF < than 0.8 requires a country to improve its reporting system before it receives its reward. In a recent review of 42 countries where DQAs were implemented, approximately 60% passed their first audit, 26% their second, 7% failed their first and second, and 7% failed their first and were under review for their second.

The DQA tool has served an important purpose in helping countries to identify the strengths and weaknesses within their administrative data systems. It has also enabled the GAVI Alliance to assess the robustness of country data and make appropriate decisions about the release of reward money. The DQA tool is currently in the process of being revised; it is envisioned that a revised tool will continue to play a central role in GAVI's performance-based systems in the future.

One option being explored is a step-wise approach that begins with an assessment of the soundness of the routine monitoring of the immunization system. This would include an assessment of how well the quality of administrative data is monitored on a routine basis at different levels of the system. Based on the findings of the first phase of the assessment, a module would be implemented in the second phase to assess the consistency and accuracy of administrative data. This assessment would cover the full "paper trail" from when a child enters the system, to the determination of immunization status, entry in register, administration of vaccine, recording of tallies and summaries from the register, the reporting of data up the system and feedback and use of data. The assessment would include a systematic examination of the level of agreement between coverage estimates from the administrative system and those from household surveys and other sources, and tracking of trends over time in discrepancies among sources. Options will be explored regarding the most appropriate means of

institutionalizing this step-wise assessment and linking it with the annual health sector review and the country's monitoring and evaluation plan.

*Source: Peter Hansen, GAVI*

## *At-A-Glance: Global Fund*

---

The *Global Fund to Fight AIDS, Tuberculosis and Malaria (the Global Fund)*, which was launched in January 2002, has adopted a performance-based funding (PBF) mechanism similar to GAVI's. The Global Fund has disbursed over \$3 billion to HIV, TB, and Malaria programs in 130 countries. Funding is released incrementally, "based on demonstrated results against agreed country-owned targets and indicators that are set out in the initial grant agreement" (Low-Beer et al., 2007). The Fund uses explicit performance ratings when deciding whether to terminate, reduce, or accelerate funding to countries. Many concerns about data quality have been raised at different points in time by observers and evaluators of Global Fund performance. The lack of local capacity among principal and sub-recipients in M&E, grant management and oversight, and verification, and a tendency of countries to focus on process indicators and to change their targets after grant signing, are a few examples of such concerns (Macro International et al., 2008). The 2008 formal evaluation of the Global Fund furthermore asserts that the PBF model is a "work in progress," that data quality is a serious issue, and that the credibility of the PBF mechanism is threatened by serious deficiencies in data validity, appropriateness, and management (Macro International et al., 2008). In response to these concerns, the Global Fund, MEASURE Evaluation, and other partners developed a Data Quality Audit (DQA) tool, which was introduced in selected countries in 2008.

The DQA uses a cluster sampling technique to assess the accuracy of immunization coverage data and the same national level estimate (verification factor) of data quality. A system of cross checks that compare reported data with alternative data sources is one Global Fund DQA innovation. The DQA has been applied to HIV/AIDS in Mali, Belarus, and Comoros; to Malaria in China, Rwanda, and the Andes Regional Program; and to Tuberculosis in the Philippines and China. The Global Fund reports that most data are within the acceptable range for accuracy, but that many of the elements of a sound monitoring and evaluation system are not in place to ensure quality reporting. The method has been found to be effective in finding gaps in the reporting system and explicit cross-checks help uncover inconsistencies. Several weaknesses are as follows: the DQA is time consuming (discrepancies and inconsistencies need to be thoroughly investigated), the accuracy statistic is ambiguous (large over-reporting can be canceled out by large under-reporting), variability in data flow within countries precludes standard implementation of methods, and it is possible that the tools may have been applied inconsistently across teams. One important finding is that "spot checks" to assess reporting validity have been rarely practical: it is costly and difficult to find people who have reportedly received services (MEASURE Evaluation, 2009).

*Source: Evaluation of the Global Fund Partner Environment, at Global and Country Levels, in Relation to Grant Performance and Health Systems Effects, Including 16 Country Studies, Final Report (2008), MACRO et al.; Data Quality Audit/Assessment Tools, The Global Fund, MEASURE Evaluation (2009), Presentation by David Boone; Making Performance-Based Funding Work for Health (2007), Low-Beer D et al., PLoS Medicine, 4:8, 1308-1311*

## ***At-A-Glance: Global Program on Output-Based Aid***

---

The Global Partnership on Output-Based Aid (GPOBA) is a World Bank-administered multi-donor Global Program that was established in 2003 to fund, design, demonstrate, and document Output-Based Aid (OBA) approaches to improve delivery of basic infrastructure and social services to the poor in developing countries. OBA aims to address a funding gap between the affordability of service to the poor and the full cost of service delivery by paying output-based subsidies. Typical outputs are working water connections (including a number of months of satisfactory service delivery) or safe childbirth packages (including pre-natal care, safe delivery and post-natal care). To make sure that outputs have been delivered and meet agreed minimum quality standards, GPOBA always requires independent verification of outputs.

Output verification in GPOBA projects is typically delegated to specialized consultants or NGOs hired by GPOBA. In cases where the payments of project implementing agencies are not dependent on output delivery, the verification agent can also be selected by the implementing agency directly. In most cases the verification agent 1) undertakes a desk review of output reporting provided by service providers to make sure that all reported outputs are consistent with the criteria agreed upon, and 2) physically verifies a statistical sample of outputs reported.

The focus of output verification varies from sector to sector. Infrastructure projects typically involve one-off connection subsidies that have to meet minimum service delivery requirements. In these projects, verification agents visit a number of beneficiary households to make sure that connections were made and meet physical standards, and verify that the minimum service delivery criteria are met. Health and education projects typically involve on-going subsidies for service delivery. OBA projects in health frequently involve reimbursements for services provided and can have a very high number of billable interventions. In these projects, the focus is both on verifying that a sample of interventions billed was performed (similar to what is typically verified in other sectors), and on ensuring the on-going quality of service delivery (e.g., by verifying that clinics meet accreditation criteria on an on-going basis and follow applicable treatment protocols).

Early lessons indicate that it is important for service providers to provide adequate information on outputs delivered, that the verification agent needs to be seen as sufficiently independent to be credible, and that the agent needs to be selected in a timely manner so that the verification process does not lead to delays in disbursements.

*Source: Lars Johannes, World Bank*