

Performance-based financing for health in Haut-Katanga

Impact Evaluation Report

Executive Summary

Introduction

The Democratic Republic of Congo (DRC) is the second largest country in Africa by area, with the fourth largest population at 66 million (World Bank, 2012). It is also among the poorest countries in the world: the country is ranked second from the bottom of the Human Development Index (186 out of 187 in 2012) (UNDP, 2012), with an estimated per capita income of US\$ 220 (current) in 2012 (World Bank, 2012). Impoverished by decades of war, instability and bad governance, it is not surprising that DRC is not on track to reach the health-related MDGs.

Since the democratic elections in 2006, the country has started a slow reconstruction phase and a decentralization process, with the election of provincial governments, including provincial ministers of health. The government's 2006 health sector strategy emphasizes development of the health zone system, indicating that contractual approaches and public-private partnerships will increasingly be options for health service financing and delivery. Developing and putting in place effective service delivery models such as Performance-based Financing (PBF) would be a strategy for improving health outcomes among the population.

Results-Based Financing (RBF) is an instrument that links financing to pre-determined results, with payment made only upon verification that the agreed-upon results have actually been delivered. Different RBF mechanisms include Output-Based Aid (OBA), Performance-Based Financing (PBF), or other provider payment mechanisms that link payments to service volume and quality, and conditional cash transfers. The central theme in all of these is that a principle entity provides a reward, conditional on the recipient undertaking a set of actions to produce a desired outcome.

Even though performance-related payment models have been implemented in developed and developing countries in various settings and forms, the scientific evidence base on the impact of these mechanisms on specific outcomes remains thin. To date, only one experimental study of the impact of PBF on health service provision and utilization has been completed. In Rwanda, PBF proved an efficient way to increase health service quality and utilization, resulting in improved child health outcomes (Basinga et al. 2011, Gertler and Vermeesch, 2013). A different program providing financial incentive to village committees to invest in health and education also increased health service provision in Indonesia (Olken et al. 2012).

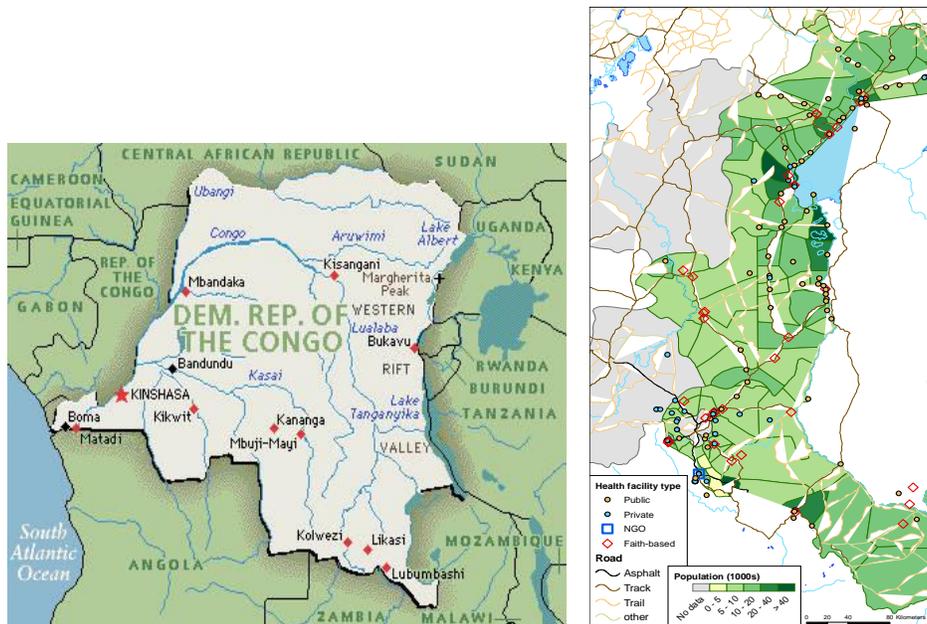
The Experimental Design of the Haut-Katanga Impact Evaluation

From 2009 to 2013, an impact evaluation was conducted in the Haut-Katanga district of DRC to study the effect of different types of provider payment mechanisms, including a PBF approach that provided financial incentives linked to service delivery volumes, on outcomes such as the quantity of services provided by contracted health facilities. The primary objective of the impact evaluation of the Haut-

Katanga pilot is to strengthen the evidence base on the relationship between performance-related payment mechanisms and health outcomes in DRC.

The district of Haut-Katanga (1.26 million people in the province of Katanga in the south-eastern corner of the DRC) is part of the \$335 million World Bank supported Health Sector Rehabilitation and Support project (*Projet d'Appui à la Réhabilitation du Secteur de la Santé - PARSS*) which covers eight districts across five different provinces of DRC. The objective of the project is to improve the access of the population to quality health care by supporting, inter alia, construction, the provision of equipment, the provision of medicine, training and salary top-ups. In addition, health facilities participating in the project subject to a user fee reduction policy where the prices of essential health services were either reduced or eliminated. .

Figure 1: District of Haut-Katanga



Source: Map on left is produced by www.greenwichmeantime.com; map on the right is produced using baseline data from the pilot.

Impact evaluations attempt to measure differences in outcomes of those participating in the intervention to those not participating. In the context of evaluating different provider payment mechanisms, outcomes of interest could include the quality of the service of providers, or the quantity and type of services used by communities. The main objective of the Haut-Katanga impact evaluation was to compare the effect of the two financing mechanisms (one conditional on the quantity of services delivered and the other a lump sum based on the number and cadre of facility staff) on health service delivery outcomes. More precisely, the impact evaluation aims at analyzing the effect of a financial incentive mechanism linked to service volume on (i) the availability of services, (ii) the price of health services and cost to patients, (iii) health workers' satisfaction, work-related stress and motivation, (iv) service utilization, (v) patient satisfaction and (v) the population health status.

Since it is impossible to compare same health facility or community with and without program at same point in time, impact evaluation attempts to measure what would happen to the beneficiary or participant in the absence of the intervention. This is called the counterfactual, and the estimated impact of the intervention is the difference between “treated” observations and counterfactual. Estimating a valid counterfactual is key to a successful impact evaluation. The impact evaluation needs to guarantee valid comparability of treatment and comparison groups, i.e. that on average they are the same prior to the intervention. When the intervention begins, the only remaining difference between the groups is that one receives the intervention and the other does not. In order to ensure that these comparison groups are on average the same prior to the intervention, the intervention must be randomly assigned across the potential beneficiaries. This means that the treatment group must be randomly chosen from a potential group of beneficiaries. The Haut-Katanga impact evaluation utilized this approach to randomly assign health areas (the unit of randomization) to intervention and comparison groups. The average causal effect of the performance-related subsidies on the outcomes of interest will, therefore, be captured by differences in outcomes across the intervention and comparison groups.

The Intervention Design

Different to the other seven districts included in the project (which did not have any PBF-type intervention), in Haut-Katanga district a PBF pilot was introduced that involved the introduction of performance-based payments to health centers and referral centers using a point system linked to the volume of targeted services delivered by contracted facilities. To measure the effect of this financial incentive payment mechanism (treatment group), an impact evaluation was conducted that introduced a predictable monthly transfer in the comparison group. In order to ensure neutrality in the level of financing between the two groups and allow for comparison of the two payment mechanisms independent of the level of money put into the system, the impact evaluation design was such that the total budget allocated to performance bonuses across health facilities in the intervention group was to be the same as the total budget for lump sum payments allocated across all health facilities in the comparison group (thus eliminating the risk of varying levels of financing between the two groups affecting outcomes of interest). Payments were to be made on a monthly basis after verification of declared results by facilities in the treatment group. The impact evaluation, thus, compares the results of a lump-sum monthly transfer to a transfer that is tied to the quantity of services provided through use of the point system.

The design of the performance-based strategy was kept simple so that it could be feasibly implemented in the difficult conditions of DRC. Consequently, the strategy lacked some of the more technically complex attributes that might be seen in other PBF interventions, such as (i) stringent quality measures tied to performance payments, (ii) business plans, (iii) quality investment units, (iv) individual health worker performance evaluations, (v) indice tool, (vi) intense coaching for the introduction of advanced strategies, (vii) fee-exemptions and waiver schemes financed through PBF targeting the poor, and (viii) rural hardship fee-adjustments, which while desirable, take a longer time and are more difficult to implement successfully. Other PBF schemes also and typically include the referral hospital, and the health administration in performance contracting. Furthermore, the budget used in this pay for performance scheme estimated at \$0.43 per capita per year (average monthly facility payments were \$550 and the average catchment area population was 12,900) was much lower than is typical in similar contexts where output budgets range between \$2 and \$3 per capita per year).

The incentivized services included seven services at the primary care level (outpatient first curative consultations, prenatal consultations, deliveries, obstetric referral, children completely vaccinated, tetanus toxoid vaccination, and family planning consultations) and three at the secondary care level (C-section, blood transfusion, and obstetric referrals to hospitals).

The “point system” scheme was such that each unit of service that was provided, of those services specified in the contract, was remunerated by a particular number of points. Summing the number of points a score was produced for each facility. Each facility’s score was then divided by the sum of all scores in the treatment group to produce the percentage of the total budget that should be allocated to each facility. Facilities’ payment was therefore determined by the quantity of services provided relative to the other health facilities. Important to note, payments in the treatment group were not linked to the quality of care provided by these facilities, the incentive scheme was linked to the relative volumes produced by each facility.

According to the design of the pilot, the volume of services was measured first by use of monthly Health Management Information System (HMIS) reports submitted by facilities in which the number of serviced delivered each month were reported. Subsequently verification of declared service volumes was conducted by verification agents through (i) comparing reported volumes with those found in health facility registers, and (ii) verifying that the information noted in the registers was true by conducting community verification by selecting a random sample of patients from registers and visiting them to ensure that they actually visited the facility and received the services noted in the facility registers. In principle, verification of facility registers was to be conducted monthly and community verification quarterly. In reality, community verification at occurred less frequently than expected and only with a sample of health facilities during each round of verification.

In contrast, the amount allocated to each facility in the comparison group was calculated based on the staff in the facility: a list of eligible workers was established at the beginning of the pilot jointly by the PARSS project and Ministry of Health; each worker received a given payment depending on his/her grade and experience. Health facilities were then randomly assigned to either the treatment or control group.

Finally, the comparison of the two payment mechanisms was conducted within the overall context of the PARSS intervention, so that both groups were benefited similarly from the overall PARSS intervention. Such interventions included: construction/rehabilitation, provision of equipment and medicine, staff training, and user fee reduction policy. However, the treatment and the comparison groups also varied in two aspects: payments in the treatment group were supposed to be contingent on verification of declared results while payments in the comparison groups were solely based on the number of health staff officially employed at the health facility and not dependent on the quantity of verified services provided. Second, in the treatment group health facility managers were granted financial autonomy to distribute the payment subsidies among facility staff as they pleased, whereas in the comparison group facility managers were expected to distribute bonus payments according to grade and classification.

Table 1: Description of the intervention

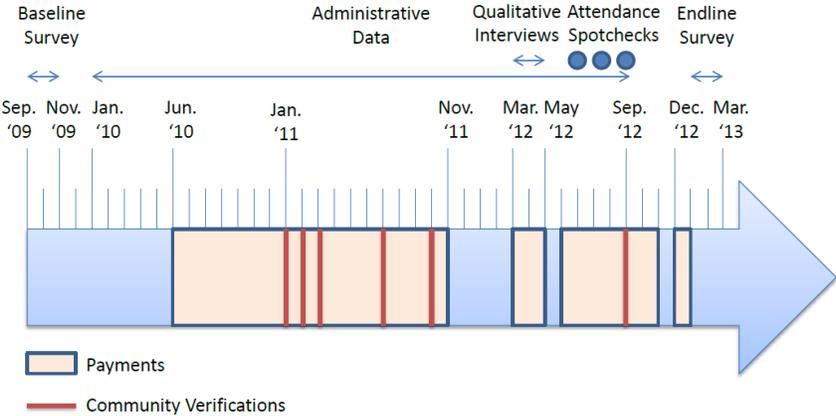
	Planned Design		Actual Design	
	Financial Incentive Group	Comparison Group	Financial Incentive Group	Comparison Group
Volume of Payment	Same		No difference in volume of payment overall – although higher payments in comparison group during first six months, budget neutrality respected for remainder of pilot	
Payment Mechanism	Payment based on the quantity of health services provided by the facility = Fee-for-service mechanism	Payment based on the size of the facility (the number of health workers on governmental payroll)	Payment based on the quantity of health services provided by the facility with a fixed total budget = Point system	As planned
Technical Verification	One visit every month by IRC/PARSS supervisors and the health zone administration - for verification of consistency between facility registers and declared data		As planned (no visit when there was no payment)	
Community Verification	One verification every month of 30 patients per health area in a subsample of health areas – for checks of accuracy of the registers to detect phantom patients. Sanctions on a case by case basis.	No verification planned	Although 28 verifications were expected, only 6 were conducted. No sanction except payment adjustment proportional to discrepancies.	2 community verifications for comparison purpose (no consequence on payments)
Autonomy	Facility manager could decide on payment allocation within the facility	Facility manager followed the governmental payroll which indicates who should receive what	As planned	
Equipment investment support	Some equipment provided by PARSS		As planned	
Drugs	Drugs for pharmacy stocking provided by PARSS		As planned	
Training	Health worker clinical training (by IRC)		As planned	
User fees	All health facilities in the pilot committed to reduce user fees		Overall reduction in user fee schedule despite substantial variation in level of reduction across facilities	

Baseline survey	Yes	Due to delay in project start-up and changes in health facility included in the pilot, The baseline sample did not match perfectly the pilot sample. Baseline survey used for balance checks but not for impact estimates
Endline survey	Yes	As planned

Implementation of the intervention was first managed by an international NGO, International Rescue Committee then replaced by the PARSS project implementation unit (UCP) within the Ministry of Health. Activities for IRC/PARSS included facility register and community verification of service volumes declared by facilities, led the calculation and processing of payments, and facilitated the process of community verification. The first payments were made in June 2010 and ended in September 2012 with several interruptions in the payment schedule over the two year period.

Five sources of data were used for the impact evaluation. First, a baseline survey was administered in Sept-Nov 2009. Second, administrative data from IRC/PARSS was collected from January 2010 to December 2012. Third, a qualitative survey was administered to a sub-sample of 30 health facilities in April-May 2012. Fourth, unannounced visits (spot-checks) were performed in July, August and September 2012 in order to collect objective data on worker attendance in the health facilities. Finally, a final survey was administered in December 2012-February 2013. The endline sample used for the main analysis included 96 health areas (123 health facilities) and 1,708 households. As the sample of health facilities taking part in the impact evaluation changed between baseline and the beginning of the pilot, only about 70% of health facilities were included in both the baseline and endline surveys. As a result only the endline survey was used for estimating the impacts of the program. The average treatment effect (ATE) was estimated for each outcome of interest. The ATE measures the difference in mean outcomes between health facilities in the treatment group and health facilities in the comparison group.

Figure 2: Flowchart of the Haut-Katanga PBF Pilot



Results

Despite (i) several interruptions in the payment schedule during the two years and (ii) the comparison group receiving substantially higher payment levels during the first six months of the intervention (June-November 2010), over the two year period the two groups of health facilities received the same level of bonus payments (on average \$550 per month per facility). All other PARSS-related investments were also equally delivered across the two groups. Therefore, the main differences between the two groups were (i) the financial incentives generated by the two different payment mechanisms, (ii) the level of community verification of results, and (iii) the level of financial autonomy facility managers had in using the payments to distribute financial bonuses to facility staff.

As payments were linked to the volume of services delivered, there was a natural risk of over-reporting by facilities in the treatment group. In order to curb this, payments were supposed to be contingent on rigorous and routine verification at the facility and community levels. In addition to a strong verification system, a system of financial sanctions was supposed to be integrated into the process in order to reduce providers' incentives to submit fraudulent reports and over-reporting of phantom patients. In reality, the system suffered from several important weaknesses. First, community verifications were substantially less numerous than planned - on average facilities only received 3 community verifications throughout the pilot, with large variations across facilities in the number of verifications conducted (some received zero, others up to eight). Second, there was actually no financial sanction associated with being caught for fraudulent over-reporting. Specifically, the reductions in payments were proportionally equal to the percentage of patients not being verified through community identification. For example, if 18% of patients were not found through community verification, the facility would only receive a reduction of 18% in their total payment and no additional sanctions were enforced.

The evaluation found that the propensity for facilities to over-report on service volumes decreased with the intensity of verification. In essence there were fewer discrepancies between declared and validated services as supervisors strengthened the attention paid to patient records. Relative to the comparison group, health facilities in the treatment group had a strong financial incentive to keep complete records of their patients, resulting in improved accountability as invalid patients were discounted from future payments. In contrast, the comparison group had no such direct incentive to maintain comprehensive patient records. Results from the evaluation show that a substantial proportion of health workers in the comparison group did not fill out consultation forms for each patient visiting the facility. As for the purpose of the impact evaluation completed consultation forms were used to validate patient volumes in both treatment and comparison groups, it is likely that service volumes were most likely underestimated in the comparison group. These opposite tendencies to over-report in the treatment group and under-report in the comparison group should be acknowledged when routine administrative data are used to compare the two groups.

In terms of financial consequences of the two payment mechanisms for the health facilities, it was found that the financial incentive payment mechanism in the treatment group led to higher payment volatility at the facility level than the fixed payment mechanism. As payments based on the point system were

adjusted relative to the performance of other facilities in the treatment group, the ambiguity of expected revenues may have led to feelings of insecurity and inability to financially plan among staff in the treatment group facilities. However, it was found that the flexibility provided to health facility managers in the treatment group with respect to bonus allocations among facility staff led to a more egalitarian distribution of payments, with a higher proportion of non-technical staff receiving bonuses in the treatment group than in the comparison group.

The introduction of the financial incentives in the treatment group led to concrete changes in health workers behaviors and the service delivery models they adopted. Facility staff were found to be present at facilities more often in the treatment group than in the control group and introduced different strategies to increase demand and utilization of health services by the population. One strategy adopted to increase the number of patients receiving services was to decrease the price of services, as it was found that fee schedules for targeted services were significantly lower in the treatment group than in the control group. A second strategy that facility staff in the treatment group used to increase service volumes was to offer more organized preventive health sessions at facilities in order to sensitize the population about the services offered by the facility and the health benefits associated with these services. Third, facilities in the treatment group were found to conduct more community-based outreach activities, such as the provision of preventative and curative services at the village level through monthly visits to communities. Overall, the financial incentives of the payment mechanism in the treatment group induced an intensification of activities used to increase demand and utilization of health services. Equally important, the evaluation found that introducing a financial incentive linked to service volumes did not have a negative effect on the quality of services, both for targeted and non-targeted health services (but quality did not increase either).

Surprisingly, the evaluation found that despite the introduction of different strategies to increase demand and access to health services by the local population, the efforts of the health workers in the treatment group did not lead to a significant increase in utilization or coverage of these services. It is crucial to think about the other barriers that impede the increase of the utilization of health services by the population. Given that the overall PARSS intervention included a package of components that included the reduction of user fees, the provision of drugs and equipment to facilities, infrastructure investment and rehabilitation, which were provided to both the treatment and comparison groups, it is possible that the additional incremental effect of either of the two payment methods had minimal to no effect on the population's health-seeking behavior. As all other components of the PARSS intervention were equal across the two groups and these interventions were introduced prior to the start of the evaluation and its baseline survey, the impact evaluation was not able to measure the effect of these other project components on health-seeking behavior or health status. Alternatively, the lack of response of the population might reflect other obstacles that were not addressed in the strategies implemented by the incentivized health facilities, such as geographic barriers, preference for alternative treatment options (traditional care, use of the informal sector, etc.), or general dissatisfaction with the quality of care provided by health facilities (quality of care was found to be the most important factor driving demand and utilization of facility-based health services). Also and compared to other PBF schemes, the general low level of performance based incentives, the lack of targeted coaching for quality and performance enhancement and the general absence of any PBF- health facility management tools might have conspired against any performance enhancements.

As facilities in the treatment group reduced the prices of their services more than comparison facilities without attracting more patients, the evaluation found that there was less total revenue in these facilities (42% less), even though on average the two groups received the same subsidy payment level from PARSS. According to evidence from the qualitative interviews, health facilities in the treatment group who reduced their fees as a strategy to increase demand found themselves in a situation where they were not able to re-adjust their price schedule and raise prices back to their original values as the population had become accustomed to the reduced prices (even though utilization did not increase) and they were fearful of reducing demand to even lower levels. Consequently, the financial incentive payment mechanism resulted in individual staff revenues being 34% less in the treatment than in the control group.

Finally, the financial incentive mechanism had a significant negative impact on job satisfaction of facility staff. The lower levels of job satisfaction found in the treatment group might be an effect of the limited response of the population to the newly introduced efforts to increase patient demand, which resulted in a reduction in facility and staff revenue (i.e. an increase in work led to a reduction in income). It is therefore not clear to what extent the negative effect of the financial incentive on job satisfaction is due to the exposure to the payment mechanism itself or to the difficulties induced by the overall PARSS pilot in the specific context of Haut-Katanga. It is also important to note that staff attendance was found lower in the PBF health facilities than in the comparison facilities several months after the pilot ended, which suggests that incentive-based payments deterred some of staff intrinsic motivation. Incentives should therefore be thought of as a permanent policy in order to steer clear of this motivational crowd-out.

Table 1: Summary of key results from the impact evaluation

Implementation of the payment mechanisms	
Understanding of the financial incentive payment mechanism	- Facility staff understood that the payment mechanism was linked to the quantity of services provided, but only limited understanding of how payments were calculated
Implementation of the payment mechanisms	- On average same level of payment in both groups - Payment irregularity in both groups (5 months without payment out of 28) - The treatment group had 26% higher payment volatility than the comparison group - Greater variations in payment levels in the treatment group - The autonomy associated with financial incentive system in the treatment group led to a more egalitarian distribution of the bonus payments among facility staff
Implementation of performance verification	- Same frequency of technical verifications - Same procedure for patient validation but different incentives to report accurate data that led to biased declarations - Very few community verifications overall in the PBF group (3 verifications per health facility) - No financial sanctions except payment adjustment for discrepancies
Strategies to increase the number of patients	The financial incentive mechanism induced more effort from the health workers to increase the number of patients for targeted services

Cost of health services	<ul style="list-style-type: none"> - User fees for targeted services and the prices of drugs were 20% to 60% lower in the treatment group than in the comparison group - The payment mechanism in the treatment group had no effect on the prices of non-targeted services
Availability of service delivery	<ul style="list-style-type: none"> - The financial incentive system led to no change in the number and type of health services offered - 43% more preventive sessions for targeted services in the treatment group than in the comparison group - No change in service delivery strategies for non-targeted services
Outreach activities	<ul style="list-style-type: none"> - Treatment group facilities offered 30% to 50% more outreach activities for targeted services than comparison group facilities, without any change for non-targeted services
Impact on service utilization	The strategies used to increase the number of patients in the treatment group proved unsuccessful
Number of patients	<ul style="list-style-type: none"> - The financial incentives had no impact on the number of patients visiting health facilities
Demand for health services	<ul style="list-style-type: none"> - The financial incentives had no impact on demand for health services
Impact on the quality of health services	The financial incentives in the treatment group did not affect the quality of health services
Technical quality of health services	<ul style="list-style-type: none"> - No differences found in the quality of care between the two groups
Patient satisfaction	<ul style="list-style-type: none"> - No differences in patient satisfaction found between the two groups
Impact on population health status	The payment mechanisms had no impact on the population health status
Impact on health facilities' resources	The payment mechanism in the treatment group led to lower resources for the facility and lower workers' payment
Total resources at the facility level	<ul style="list-style-type: none"> - The treatment group had 42% less total resources for the facilities than comparison group
Health worker income	<ul style="list-style-type: none"> - Health workers total income was 34% less in the treatment group than in the comparison group
Quality of facility infrastructure and equipment	<ul style="list-style-type: none"> - The quality and the quantity of equipment was poorer in the treatment group than in the control group (of about 0.6 standard deviation)
Impact on staff satisfaction and motivation	The financial incentives led to lower subjective workload but also lower job satisfaction compared to fixed payments. The importance attached to remuneration and the concern on its volatility proved higher.
Staff satisfaction, stress and anxiety	<ul style="list-style-type: none"> - Job satisfaction was 14% lower in treatment group than in control group - The treatment group had 28% fewer health workers feeling they have too much work than the comparison group - The treatment group had 72% more facility heads concerned by the volatility of health facility revenues
Staff motivation	<ul style="list-style-type: none"> - The financial incentives had a positive impact on staff attendance during the program (14% more during spot checks) but a negative one at the time of the survey (25% less 4 months after payments stopped) - The treatment group had 34% more health workers who attach importance to remuneration than the comparison group

Conclusion

This study examines the effect of financial incentives to health care providers linked to performance on health service delivery outcomes in the district of Haut-Katanga, DRC. The findings show that while the incentives did lead to increased availability of health services at the facility and community level, these changes did not lead to significant changes in the coverage or quality of services provided when compared with a lump sum payment mechanism that was not performance-related. The study also found that the purchasing of a specific sub-set of services in the intervention group did not have a negative effect on the availability or quality of non-targeted services. Equally important, the study found that the incentive-based payment mechanism had a negative impact on the availability of financial and non-financial resources for health facilities and their staff. As perceived quality of care was found to be an influential factor in household health-seeking behavior, the limited effect of the payment mechanism on quality outcomes may have contributed to insignificant differences in service coverage observed between the two groups. While this particular PBF approach didn't lead to significant changes, other comprehensive approaches designed and delivered by Cordaid and others in South-Kivu and Kasai have shown that PBF in DRC can lead to positive results. Furthermore, Experience from well-designed and well-implemented PBF schemes from Rwanda and Burundi, and increasingly from other countries and contexts (Nigeria; Benin; Chad; Zambia; Zimbabwe etc) point at the importance of good design and implementation.