PRELIMINARY This version: 04/27/2009

Impact of Paying Primary Health Care Centers for Performance in Rwanda

Paulin Basinga [*]
National University of Rwanda School of Public Health

Paul Gertler [* +]
University of California, Berkeley

Agnes Binagwaho
Ministry of Health, Rwanda

Agnes Soucat
The World Bank

Jennifer Sturdy
The World Bank

Christel Vermeersch
The World Bank

[*]Lead Authors, other authors listed alphabetically.

[+] Corresponding author:
Paul Gertler
Li Ka Shing Professor
Haas School of Business, MC#1900
University of California
Berkeley, CA 94705 USA
Gertler@haas.berkeley.edu

ABSTRACT

Background: Paying for performance (P4P) provides incentives payments to medical care providers for improvements in performance measured by specific utilization and quality of care indicators. P4P can affect health care in two ways; first by providing incentives for providers to put more effort into specific activities, and second by increasing the amount of resources available to finance the delivery of services. While the approach holds promise for improving health system performance, there is little rigorous evidence of its effectiveness, especially in low-income settings. In 2006, Rwanda began paying for performance at the health facility level in order to improve maternal and child health outcomes.

Methods: This study examines the impact of the incentives in the Rwandan P4P scheme on prenatal care utilization, the quality of prenatal care, institutional delivery, and child preventive care utilization using data produced from a prospective quasi-experimental design nested into the program rollout. The rollout was implemented in two phases in rural areas, with 80 facilities (treatments) enrolled in the P4P scheme in 2006 and another 86 facilities (comparisons) enrolled two years later. In order to isolate the incentive effect from the resource effect, the comparison facilities' traditional budgets were increased an amount equal to the average P4P payments to the treatment facilities. We collected baseline and end-line data from the 165 facilities and a random sample of 14 households in each facility's catchment area for a total of 2145 households.

Findings: We find that P4P had a large and significant positive impact on institutional deliveries and preventive care visits by young children, and improved the quality of prenatal care measured by process indicators of the clinical content of care and tetanus toxoid vaccination. However, we find no effect on the number of women completing 4 prenatal care visits or on the number of children fully immunized.

Interpretation: Our results provide evidence to support the hypothesis that financial performance incentives can improve both the use of and the quality of maternal and child health services. Because we isolate the incentive effect from the resource effect in P4P, we conclude that an equal amount of resources without the incentives would not have achieved the same gain in outcomes.

INTRODUCTION

The 2015 horizon of the Millennium Development Goals (MDGs) is now only seven years away and progress to date has been limited, especially in Sub-Saharan Africa.[1] Under-five mortality has decreased by only twenty percent in Sub-Saharan Africa between 1990 and 2005, whereas the MDGs call for a two-thirds decrease by 2015.[2] While official development assistance for health has increased dramatically over the last few years, health outcome indicators have not seen commensurate improvements.[3] Some explanations for this are that public spending often benefits richer groups disproportionately, [4] resources do not reach frontline providers, [3, 4] and health worker productivity and morale is low, often plagued by absenteeism. [5]

To reach the health MDGs, low-income countries need to improve access to and quality of care. One promising supply-side intervention that may accelerate progress is to pay health providers for performance. Paying for performance (P4P) provides financial incentives or bonuses to medical care providers to improve performance measured by specific utilization and quality of care indicators. P4P can affect health care in two ways; first by providing incentives for providers to put more effort into specific activities, and second by increasing the amount of resources available to finance the delivery of services. Proponents of the approach highlight that P4P strengthens the link between resources and productivity, and provides incentives to increase the quantity and quality of services supplied; a highly desirable outcome in developing countries where utilization of critical life saving services is still very low.[6] A number of low-income countries are piloting or scaling up P4P, including Afghanistan, Argentina, Democratic Republic of Congo, Benin, Eritrea, Ghana, Haiti, Kyrgyz Republic, and Zambia.[6]

In 2006, the Ministry of Health in Rwanda started implementing P4P on a large scale. This study examines the impact of P4P incentives on maternal and child health care in Rwanda using data produced from a prospective quasi-experimental design nested into the program rollout. Specifically, we examine the impact of the P4P incentives on prenatal care utilization, the quality of prenatal care, institutional delivery, and child preventive care. Not only is this among the first rigorous evaluations of P4P impact in a low-income setting, it is the first to isolate the impact of P4P incentives from the associated increase in resources. This is important because if P4P achieves its results from increased resources rather than incentives, then the same results could be achieved from an increase in traditional input-based budgets and there would be no reason to incur the administrative costs associated with P4P.

Despite the promise of P4P, there is little rigorous evidence on its impact and none that separates out the effect of incentives from increased resources. What little rigorous evaluation exists is in developed country contexts. In fact, a recent review of the evidence on P4P in hospitals in developed countries recognized that there has been little formal evaluation and many of the published studies have methodological flaws.[7] The 3 studies with control groups[8-10] demonstrate that hospitals participating in a P4P program had greater improvement in quality than observed in control hospitals. A number of other studies in developed countries show that P4P is associated with an increase in the use and quality of primary health care services.[11-14] However, one study found no impact of P4P on the quality of chronic disease management.[12]

To our knowledge, there has been no other rigorous evaluation of a P4P scheme in a low-income country, particularly no study which isolates the incentives effect (Eldridge, Palmer 2009). One study of a P4P program in Haiti, without a comparison group, shows that the time trend in immunization coverage and attended deliveries in NGO primary care facilities jumped after the introduction of bonus payments that were linked to performance targets.[15] Similarly, studies of the early Rwandan P4P pilot programs without comparison groups reported large increases in the time trend of the number of curative consultations, new family planning acceptors and institutional deliveries, but no increases in measles vaccination.[16-18]

METHODS

Rwanda is one of the poorest countries in the world, ranked 165 out of 177 countries on the Human Development Index with a GDP of US$ 819 PPP per capita in 2008.[19] The 1994 genocide decimated Rwanda's fragile economic base, severely impoverished the population, particularly women, and eroded the country's health system. Out of this devastation, Rwanda has made remarkable progress in improving maternal and child health. Between 2000 and 2005, infant mortality fell from 107 to 86 deaths per 1000 lives births.[20] Despite this progress, Rwanda remains far from achieving the 2015 MDG targets.

Performance Based Financing

The first P4P schemes, or Performance-Based Financing (PBF) as they are called in Rwanda, were launched in 2001 by nongovernmental organizations (NGOs) that believed their traditional fixed salary supplements to health workers were not achieving improvements in health service delivery.[16] In 2005, the Government of Rwanda (GoR) adopted the idea and through a collaborative process with development partners, designed and implemented a national model in the first half of 2006.[21-22]

The Rwandan national PBF scheme provides bonus payments to government and faith-based primary care facilities based on the provision of various types of services and the quality of those services. The formula used for payment to facility i in month t is:

$$Payment_{it} = \left(\sum_j P_j U_{j\,it} \right) \times Q_{it} \quad , \quad 0 \leq Q_{it} \leq 1$$

where P_j is the payment per output unit j (e.g. institutional delivery or child preventive care visit), U_{jit} is the number of patients receiving output j in facility i in period t, and Q_{jt} is the quality index of facility i in period t bounded between 0 and 1. If the overall quality index for the facility is 1, then health facilities are paid the maximum bonus for the services provided. By contrast, if the quality index is less than one, PBF payments are discounted for all services.

There are 14 maternal and child health PBF output indicators (U_{jit}), each with an associated per unit payment rate (

Table 1). The first 7 indicators capture the utilization of key services, including curative care, prenatal care, family planning, institutional delivery, and child preventive visits. The second set of 7 indicators capture aspects of the clinical content of the care provided during visits.[21,22] These content indicators are measures for the process quality of care, and include the number of children who were fully vaccinated during preventive visits, the number of pregnant women who received tetanus vaccines and malaria prophylaxis during prenatal care, the number of at-risk pregnancies referred to hospitals for delivery during prenatal care, the number of severely malnourished children who were referred to treatment facilities during preventive visits, and the number of general emergencies referred to the appropriate place for care.

The size of the monetary incentive (P_j) associated with each indicator was determined through a collaborative process between the MoH and development partners taking into consideration the pilot experiences, health priorities and available budget. The largest per unit monetary incentive is $4.59, paid for each institutional delivery and verified emergency referral of a woman to a hospital for obstetric services. While prenatal care is paid only $0.09 for the first visit, with a bonus of $0.39 for completion of 4 visits, the facility can earn a large payment if a provider provides specific content during a prenatal care consultation. Specifically, the provider can substantially raise the payment if she administers a tetanus vaccine and malaria prophylaxis (an additional $0.92), and the woman delivers in the facility ($4.59) or in the case of a risky pregnancy, delivers at a hospital ($1.83). Similarly, while the payment rate per preventive visit is low ($0.18), the rate for identifying a malnourished child and referring her for treatment is relatively high ($1.83). Since almost 50 percent of children in Rwanda are stunted[20] and could be referred, half of those visits could yield a payment of $2.01. In addition, if the provider fully vaccinates a child as part of a preventive visit, there is an additional \$0.93 payment for the facility.

Quality of care enters the payment formula through the multiplicative factor (Q_{jt}) that raises or lowers the payment for all outputs. The quality index component of the payment formula (Q_{jt}) is a function of both structural and process measures of quality of care[23] that are identified using the Rwandan clinical practice guidelines.[21,22] Structural measures are the extent to which the facility has the equipment, drugs, medical supplies and personnel necessary to be able to deliver a specific medical service, while process measures partly capture the clinical content of care provided for specific services.

The formula for the quality index is:

$$Q_{it} = \sum_k \omega_k S_{kit} \quad \text{with} \quad \sum_k \omega_k = 1 ,$$

where S_{kit} is the share of indicators for service k that are met by facility i in period t, and ω_k is the weight for service k. Note that the weights sum to 1. Therefore, if a facility has perfect structural and process quality, then all S_{kit} take on value 1 and the overall quality index is equal to 1. Each S_{kit} refers to the quality of a particular service such as prenatal care, curative care, delivery, etc.

Table 2 lists the services, their weights, the extent to which the underlying indicators are structural or process-based, and the method by which the indicators are assessed. Each S_{kit} is the facility's proportion satisfied of the structural and process indicators necessary to deliver that service as recommended in the Rwandan guidelines.

Facilities report their total monthly output for the 14 indicators (the *U*'s in

Table 1) to the district PBF steering committee, *Comité de Pilotage*, responsible for authorizing payment. In addition, for the referral indicators (hospital delivery, emergency obstetric, and malnourished child), the clinic must submit verification from the hospital that it treated the referred patient to trigger payment. In order to verify reporting, the committee audits each facility's report by sending auditors on a quarterly basis on an unannounced randomly chosen day. The auditors review the utilization registry and facility records in order to verify the data reported to the Comité de Pilotage is the same as that recorded at the facility level. False reporting is penalized by publicly identifying the facility and director who made the false report. An assessment of the extent of false reporting was conducted by validating facility records with face to face interviews of a random sample of patients and found very low levels of false reporting (Rapport d'Enquête de Contre vérification par la Communauté dans les districts de Nyamasheke, Nyanza, Nyaruguru et Rulindo » HDP, December 2008).

By contrast to the quantity indicators, the quality indices (Q_{it}) are assessed through the regular monitoring and supervision system, in which district hospitals have the responsibility of monitoring and supervising the quality of care at primary care facilities. In fact, District Hospital budgets are tied to the correct and timely execution of this supervision activity. Every quarter, qualified supervisors from the corresponding district hospital visit each facility on an unannounced randomly chosen day, and assess quality indicators through direct observation and review of patient medical records. At the end of the visit, they discuss their findings and the facility's quality score with its personnel, and provide practical recommendations on how to improve the quality of services where needed. In PBF districts, information from the visit is used to compute each facility's overall quality score Q_{it}.

PBF payments go directly to facilities, which decide how to use the funds, without restrictions. The overall amount of PBF payments is large in comparison to the facilities' budgets. In the sample of 80 facilities receiving PBF payments, we find the PBF payments increased overall expenditures by 22 percent on average. In addition, facilities chose to allocate 77 percent of the PBF funds to increase personnel compensation, amounting to a 38 percent increase in staff salaries. In addition, approximately XX% (?? SOURCE3 ??) of the total P4P budget and X% of the total expenditures was spent on administration, technical assistance, auditing and other related costs (PBF Annual Report 2007, CAAC/MOH)

Identification Strategy

Bearing in mind the lack of evidence on the impact of PBF, the GoR supported an independent evaluation that took advantage of the phased implementation of the national PBF model over a 24-month period. This arms length evaluation used a randomized stepped wedge design for the rollout of the program at the district level. In collaboration with the evaluation team, the GoR randomly allocated districts to treatment and comparison groups. Administrative districts with pre-existing pilot PBF schemes were excluded from the sampling frame. The remaining districts were then grouped into 8 pairs, where each pair had similar characteristics for rainfall, population density, and predominant livelihoods based on data from the 2002 Census. A coin was then flipped for each pair to allocate one side of the pair into the treatment group and the other into the comparison group.

At the time of the evaluation design, the GoR was in the process of redrawing administrative district boundaries in the context of a decentralization effort. The Ministry for Local Government, Good Governance, Community Development and Social Affairs was responsible for the restructuring of administrative districts (REFERENCE). In this process, some of the experimental areas were combined with areas that already had PBF from the pilot phase. Because PBF could not be "removed" from health facilities that were already implementing the system, and because PBF was to be managed at the district level, the GoR required that all facilities within those new districts be in the first phase (treatment) of the rollout. This led the evaluation team to switch the assignment of treatment and comparison for two of the eight pairs. In the end, 165 of Rwanda's 401 primary care facilities were included in the study, 80 in treatment districts and 86 in comparison districts.

The facilities in the treatment group started receiving PBF in 2006, while the facilities in the comparison group continued with traditional input-based financing for an additional 24 months before incorporating into the PBF system. Since a primary objective of the evaluation was to isolate the PBF incentive effect separately from the effect of an increase in resources, it was necessary to hold the level of resources constant across treatment and comparison facilities. To accomplish this, comparison facilities' traditional input-based budgets were increased by the average amount of PBF payments to treatment facilities on a quarterly basis during the entire 24-month treatment window. Therefore, if the design was correctly implemented, the total amount of resources should be the same for treatment and comparison facilities both at baseline and endline. Then, the only difference between the two groups is that increased payments to the treatment group were linked to incentives, while the comparison group received a lump sum increase in traditional budgets.

Data

We conducted a baseline survey pre-intervention and a follow-up survey 24 months later, both at the facility and at the household level. The intervention started in treatment districts between 1 and 5 months after the completion of the baseline survey, yielding an 18-23 month exposure period.

The facility survey included all 165 facilities, and collected information about budget, staffing and structural quality elements from the facility head. We also interviewed a randomly selected staff member responsible for delivering prenatal and childcare services on the day of the survey regarding his/her background and knowledge of the clinical care protocol. Finally, we conducted 8 to 12 exit interviews of pregnant women leaving the facility in order assess the process quality of prenatal care and patient socio-demographic characteristics.

In addition, we conducted a survey of a random sample of households living in each facility's catchment area. The household questionnaire collected information on the basic socio-demographic characteristics, maternal and child health status and health services utilization. The household sample was constructed by first sampling 14 zones (approximately 15 to 20 households) from each facility's official list of zones in their catchment area. We then physically listed all households in the sampled zones and randomly selected one household with at least one child under 6 years old from each zone. Therefore, the sample includes 14 randomly selected households with a child less than 6 years old per facility, for a total of 2145 households.

Less than 2 percent of sampled households refused to participate in the interview. Those households were replaced with randomly selected households with a child under six from the same zones. In the follow-up survey, 88 percent of the baseline households were re-interviewed.

The rate of attrition from the baseline sample was not different between the treatment and control groups (12 percent each), and treatment status was uncorrelated with the probability of attrition. Households that could not be found or interviewed were replaced with randomly selected households from the same zones.

Outcome Measures

Maternal health services utilization: We use maternal health services utilization for women who had a pregnancy in the 18 months preceding the interview. We restrict the sample to pregnancies in the last 18 months as these are the pregnancies that are most likely impacted by any changes in utilization and quality of care as a result of the PBF scheme. The indicators include a dummy variable indicating any prenatal care utilization, another dummy indicating completing four or more prenatal care visits, and a dummy indicating that the baby was delivered in a health care facility.

Quality of prenatal care: We assess the process quality of prenatal care delivered by computing the share of clinical content items that should compose a typical first prenatal consultation, as recommended in the Rwandan clinical practice guidelines for prenatal care, to the actual clinical content items delivered during a prenatal care consultation.[25] This measure has been used extensively in the literature to measure quality.[26-32] The items in the measure cover medical history questions, physical examinations, lab tests and follow-up procedures as established in the Rwandan clinical practice protocol. The specific items are:

- *Pregnancy history*: number of prior pregnancies, living children, and miscarriages, bleeding during previous labor, how the last child was delivered, and last child's birth weight.

- *Gynecological history*: STIs including HIV, pap smear, contraceptive use, last menstrual date, and related health problems.
- *Current obstetric symptoms*: contractions, vaginal bleeding, weight loss/gain, nausea, vomiting, and current medications.
- *Medical history*: high blood pressure, diabetes, contraceptive use, heart disease, disease, malaria, goiter, and tobacco and alcohol use.
- *Physical examination*: body height, body weight, check vital signs (blood pressure, temp, respiratory), palpitate abdomen, listen to fetal heartbeat, check for edema and measure uterus.
- *Laboratory tests* for hemoglobin (anemia), diabetes, urine protein, platelet count, HIV and STIs (syphilis /gonorrhea).
- *Prevention and case management*: advice about nutrition, tetanus vaccine, iron/folic acid supplementation, advice about danger signs for emergency help, HIV voluntary counseling/test, and complete prenatal card.

The prenatal quality score was computed at the individual patient level using two samples of patients. For the first, enumerators were posted at the exit of health facilities and interviewed pregnant women who visited the facility for a prenatal care visit. For the second, the same information was collected in the household survey from women who gave birth in the last 18 months and received prenatal care from the facility in whose catchment area they lived. In the analysis, we combined the exit interview and household survey data for first prenatal care visits to assess the impact of PBF on prenatal care quality. Cronbach alpha scale of reliability for the 38 item score in the combined sample is 0.78, indicating satisfactory internal consistency. In the impact analysis, we standardize the score by subtracting out the baseline mean and dividing by the baseline standard deviations. The unit of measurement then is interpreted as standard deviations in quality.

In addition to the prenatal care process quality index, we investigate separately whether women reported receiving the tetanus vaccine during prenatal care visits as the tetanus shot enters the payment formula as one of the *U*'s and as part of one of the *S*'s.

Child Preventive Care Utilization: We define utilization of child preventive care as whether the mother reported taking her child to a health facility for a preventive care visit in the four weeks prior to the household survey. Child preventive services cover immunization, vitamin A, distribution of mosquito nets and child growth monitoring with referral of malnourished children to higher levels of care for treatment. We analyze preventive visits separately for children 0-23 and 24-59 months old as the younger group is expected to have more visits than the older group.

We also examine the impact of PBF on full immunization, which, for 12-23 month olds, is measured by a dummy variable coded 1 if the child received all vaccines required by the protocol and 0 otherwise. According to the national protocol, by month 12 the child must receive the tuberculosis vaccine (BCG), three series of diphtheria–tetanus–pertussis- *Haemophilus influenza* type b and hepatitis B vaccine and polio. The measles vaccine is given at month 9. Vaccination status was assessed based on the vaccination card and not maternal reports. About 4 percent of the mothers could not produce the child's vaccination card and those observations were dropped form the analysis.

Statistical Methods

Despite the fact most of the districts were randomized into the treatment and comparison groups, we view the evaluation design as quasi-experimental. As a result, we use the difference-in-difference method to estimate the program impact. This method compares the change in outcomes in the treatment group to the change in outcomes in the comparison group. By comparing changes, we control for observed and unobserved time-invariant characteristics as well as for time-varying factors that are common to the comparison and treatment groups. The change in the comparison group is an estimate of the true counterfactual – i.e. what would have happened to the treatment group if there had been no intervention.

Formally, we estimate the following regression specification of the difference-in-difference model for individual outcomes:

$$Y_{ijt} = \alpha_j + \gamma_{2008} + \beta \cdot PBF_{jt} + \sum_k \lambda_k X_{kit} + \varepsilon_{ijt}$$

where Y_{ijt} is the outcome of interest for individual *i* living in facility *j*'s catchment area in year *t*; PBF_{jt} is a dummy variable that takes value 1 if facility *j* received PBF in year *t* and 0 otherwise; α_j is a facility fixed effect; γ_{2008} is a dummy variable that takes value 1 if the year of observation is 2008 and 0 otherwise; the X_{kit} are individual and household characteristics; and ε_{ijt} is a zero mean error term. We compute robust standard errors clustered at the district by year level to correct for possible heteroskedasticity and correlation of the error terms across facilities within districts. The models were estimated using Stata version 10 software.

Covariates

In all of the estimated models, we control for the total number of family members, the number of family members under six years old, family assets, the distance of the household from the facility, and whether the family is enrolled in mutuelle, an association that provides community-based health insurance. Assets are measured as the value of owned houses, durables in the house, farm animals, farm equipment and microenterprise equipment. Rather than entering the value of assets linearly, we use dummy variables to represent the quartiles of the asset distribution. We also include a dummy indicating if the household owned land.

In the prenatal care and delivery analyses, we also control for maternal years of schooling, marital status, whether the partner currently lives in the household, the number of prior pregnancies, and age defined in years. When analyzing the quality of prenatal care, we stratified women into a high and low risk groups, where we consider that women above the age of 35 and below the age of 20 have potentially high risk of pregnancy complications. In the prenatal care quality analysis, we also included a dummy variable indicating whether the observation was from the facility exit interview or household survey.

In the children's utilization and outcomes analyses, we include controls for whether the mother and father currently live in the household, their age and years of schooling, the child's age and sex. Age was entered as a series of dummy variables that represent 3-month age increments in the utilization models.

Ethical Review and Informed Consent

The Rwanda National Ethics Committee (RNEC) in Kigali approved the research design and methodology. In each round of the evaluation, interviewees were invited to participate in the survey after receiving a detailed explanation of the objectives, procedures, risks and benefits and, if they agreed, were asked to sign an informed consent declaration. In cases where the interviewees granted consent but refused to sign the informed consent, a witness was identified to verify informed consent.

Role of Funding Sources

The research was funded through grants from The British Economic and Social Research Council, The Global Development Network, The Government of Rwanda through a Japanese PHRD grant, The World Bank's Bank-Netherlands Policy Program, and The World Bank's Spanish Impact Evaluation Fund. The design of the evaluation in 2005 was developed in close consultation with the Rwandan PBF project team to ensure that the program monitoring and evaluation needs were met. However, all data collection, management, analysis, and interpretation, preparation, and review of this manuscript were independent of any influence by the PBF program, its staff or the funders.

RESULTS

We begin by confirming that the evaluation design achieved balance at baseline of observed characteristics between the treatment and comparison groups. Table 3 reports the means of facility level characteristics in 2006 before treatment facilities started receiving PBF. We find no significant differences between the treatment and comparison groups in terms of total expenditures, allocation of the budget across medical personnel, medical supplies and non-medical purposes, and the numbers of physicians, nurses and other types of personnel.

As a key goal of the evaluation design was to isolate the PBF incentives from the effect of increased resources, we must also confirm the program compensated the comparison facilities with an increase in their traditional budget equal to the increase in treatment facilities' resources. In the second row of Table 3, we report the mean 2008 log expenditures for treatment and comparison facilities and find no statistically significant difference in the means after the introduction of PBF in the treatment facilities. While expenditures did more than double from 2006 to 2008, the rate of increase was the same for the two groups. This validates our assertion that we can interpret the estimated impacts of PBF on outcomes as a result of a change in incentives, as opposed to an increase in resources.

Prenatal Care and Institutional Delivery

We now turn to the sample used to estimate the impact of PBF on the use of prenatal care services, the quality of prenatal care, and institutional delivery. Our sample consists of women 15-49 years old who gave birth in the 18 months prior to the survey. We observe 620 births in catchment areas of treatment facilities and 670 in comparison areas. Table 4 compares the baseline means of these outcomes and the socio-economic characteristics of the sample for the treatment and comparison groups. We find no statistical differences between the groups in terms of age, education, marital status, parity, health insurance, household size, or household assets at baseline. These are the control variables we use in subsequent multivariate regression models for prenatal care and institutional delivery.

Table 4 also shows that the treatment and comparison baseline samples are fairly well balanced in terms of the use of prenatal and delivery services, and the quality of prenatal care services. Although 95 percent of women obtained some prenatal care, only about 10 percent of women initiated prenatal care in their first trimester and the mean number of visits was less than 3. Moreover, only about 18 percent of women in treatment areas and 11 percent in comparison areas completed the 4 visits recommended by Rwandan, and WHO, guidelines, and this difference is statistically significantly different from zero at the 0.05 level. In terms of location of delivery, less than 35 percent of women delivered in a health center. Finally, in terms of prenatal care quality, while close to 70 percent of women received a tetanus shot, providers executed only 46 percent of the activities recommended in the Rwandan clinical practice guidelines during a prenatal care visit.

Table 6 reports the estimated PBF program impacts on maternal care. First, as indicated by the coefficient on the wave dummy, there were statistically significant increases from 2006 to 2008 in the entire sample for prenatal care, institutional deliveries and quality of prenatal care. There appears to be no impact of PBF on the probability of any prenatal care or on the probability of completing 4 or more visits. However, we estimate a statistically significant impact on the probability of institutional delivery of 0.074 percentage points, with a 95 percent confidence region of (0.004, 0.143), which represents a 21 percent increase from baseline. We also find significant impacts on the quality of prenatal care. We estimate that the impact on the probability of getting a tetanus vaccine is 0.054 (0.007, 0.265) percentage points, which is a 7.6 percent increase from the baseline. We also estimate an increase of 0.14 (0.015-0.265) standard deviations in the standardized prenatal quality score.

Child Preventive Care and Vaccination

In the second part of our analysis, we estimate the impact of PBF on use of preventive medical care by children. The sample includes children less than 59 months at the time of the survey. Again, the sample is very well balanced between the treatment and control groups in terms of baseline socio-economic characteristics including child age and sex, parental characteristics, and household characteristics (Table 5). We use those characteristics as control variables in the regression analysis. Age is entered as a series of dummy variables rather than linearly.

The samples are also well balanced in terms of baseline values of the outcome variables. On average, children 0-23 months old had 0.22 preventive visits in the 4 weeks prior to the survey. Children 24-59 months old living in treatment areas visited the health center 0.08 times in the last 4 weeks on average, while for those living in comparison areas the number was 0.14. While the difference in the treatment and comparison group rates are not significantly different in the 0 to 23 months old group, the difference is significantly different at the 0.11 level for the older group. Close to 65 percent of children 12-23 months were fully vaccinated in 2006.

As in the case of maternal care, again we find there were statistically significant increases from 2006 to 2008 in the entire sample for child preventive care utilization. In addition, we find large significant impacts of PBF on children's use of preventive health care (Table 6). We estimate that PBF increased the probability that a child 0-23 months visited a health facility for preventive care by 0.134 (0.045, 0.224), which is a 64 percent increase over baseline. Similarly, we estimate that PBF increased the probability that a child 24-59 months had a preventive visit by 0.106 (0.050, 0.111), which is a whopping 133 percent increase from the baseline. However, we did not find a significant effect on the probability that a child age 12-23 months was fully immunized.

DISCUSSION

We provide evidence that the incentives in the Rwandan PBF program are significantly associated with increased use and quality of a number of high impact maternal and child health care services, including quality of prenatal care, institutional delivery and child preventive care utilization.[33-38] These results suggest that financially rewarding providers for their performance yields improvements in the areas targeted and, hopefully, on maternal and child health outcomes.

Context

It is important to note that these effects were in the context of a larger health sector reform. From 2001-2008, the Rwandan MoH, in collaboration with international development partners, designed and implemented a number of programs to address major demand and supply side bottlenecks in health service delivery and utilization. These programs include decentralized performance-based budgeting at the district level, called "Imihigo". Imihigo was implemented nationally and provided incentives to the districts for improvements in health care utilization. Similarly, health insurance (mutuelle) enrollment was scaled up nationally so that by 2008 close to 75 percent of Rwandans had health insurance (World Bank-Ministry of Health of Rwanda 2009 Country Status Report on Health and Poverty in Rwanda). Finally, a national campaign took place in 2006 to rapidly increase the coverage with immunization, vitamin A supplementation and Insecticide Treated Nets. The results of our analyses confirm an important effect of the national efforts on health services utilization as the evolution of indicators between 2006 and 2008 is almost always positive and statistically significant. However, our results are not biased by these programs as Imihigo and other health campaigns were implemented nationally and we directly control for the enrollment of households in health insurance schemes in the analyses.

Quality of Care

One of the more important results of this analysis is the effect of PBF on the quality of care delivered. Although health workers may be competent to perform a medical procedure or consultation, (i.e. prenatal care) they may not always be willing or motivated to expend the effort to perform all the required components of that procedure.[39] By conditioning the PBF payment on a quality index score, the evidence suggests the Rwandan P4P scheme gave providers the incentive to translate their knowledge about prenatal care into better practice. In addition, the PBF explicitly separates out the tetanus vaccine during prenatal care for payment, and indeed we find a significant positive impact.

This is important primarily because better quality care yields better health outcomes, and in the case of prenatal care this impacts both pregnant women and unborn children.[40-44] During prenatal care in Rwanda, quality care involves diagnosing and managing pregnancy-related complications such as hypertension and high-risk deliveries. In addition, women receive preventive measures including screening for and treatment of underlying conditions and diseases such as anemia, malaria, sexually transmitted infections -including syphilis and HIV- , and underlying mental health problems. Prenatal care is also used to provide health education and promotion to women and their families. Specific topics include healthy lifestyles, healthy diet, support and care in the home, prophylactic treatments such as iron supplementation, vitamin A, folic acid and use of insecticide-treated bed nets. In addition, women should also receive a tetanus toxoid vaccination, which effectively protects newborn babies and young infants born to mothers with tetanus antibodies against tetanus by acquired maternal antibody. [45,46]

Utilization

While the PBF incentives have strong effects on the use of certain services, they have little effect on others. Specifically, we see that PBF incentives are significantly associated with increases in institutional deliveries and the use of preventive childcare, but not with prenatal utilization. This is probably related to the structure of the incentives. Deliveries have the highest unit payment rate at $4.59. During focus groups, providers reported that they found deliveries to be so lucrative that they not only encouraged women to deliver in the facility during prenatal care, but that they also commissioned community health workers to conduct outreach in the community and bring in pregnant women to deliver in the facility. This is one example of how the P4P incentive scheme resulted in creative solutions to increase utilization of PBF indicators.

Similarly, the large increase in preventive child visits can also be explained by the higher payment rate. While the payment rate per preventive visit is low at $0.18, the rate for finding a malnourished child and referring them for treatment is very high at $1.83. Since almost 50 percent of children are stunted [20]and could be referred, half of those visits could yield a payment of $2.01.

Why do we find no effect of the incentives to increase the number women who had any prenatal care and the number who completed 4 visits? First, the payment rate for the initial prenatal visit is very low at $0.09, and is only $0.39 for completing 4 prenatal care visits. Since over 95 percent of women at baseline made at least one visit, the $0.09 payment provides little incentive to find the few remaining women who do not use prenatal care. Second, the real money was in the content or quality of care. Specifically, facilities receive $0.92 for every tetanus and malaria vaccines, $1.83 for referring risky pregnancies to hospitals for delivery, and $4.59 for every woman who delivers in the facility. Since women start prenatal care late in Rwanda (5th or 6th month) and have between 2 to 3 visits on average, most of the activities to get the majority of the payment could be accomplished in the 2 to 3 visits, and there was little extra incentive to get the women back to complete the 4 visits, or find women to come in earlier during their pregnancy.

The differential effects of the program may also be explained by the relative difficulty and the amount of effort needed to increase these services. In the case of deliveries and well-baby care, providers were able to advise women already in prenatal care to deliver in the facility and bring their newly born children back for preventive care, whereas they would have to go out into the community to find recently pregnant women and convince them not only to start prenatal care early, but also return for four visits.

Vaccinations

One surprising non-result is that we did not find an impact of PBF on child vaccinations. For vaccination, this can be explained in part by the fact that child vaccination rates at baseline were already close to 65 percent. In addition, the government implemented an intensive vaccination program during the intervention period.[47] An increase beyond this would have required substantial effort on the part of the providers to enter the community, identify unvaccinated children and provide them with vaccinations. It may not have been worth this effort for the providers given the small per unit financial incentive.

Policy Implications

Our findings illuminate the debate on some of the benefits and shortfalls of P4P. The results of this analysis suggest paying health facilities for performance is a feasible and effective method for improving health system performance. More importantly, the role of incentives in P4P is key. Because the comparison facilities' regular budgets were increased by an amount equal to the P4P payment to the treatment group, we were able to isolate the P4P incentive effect from the resource effect. This implies that the same results could not have been achieved by simply increasing the amount of resourcing without the incentives.

A number of specific P4P lessons also emerge from this study. First, higher payments seem to provide high-powered incentives where providers have more control, such as prenatal care quality, convincing women to deliver in facilities once they attend a prenatal care visit, and promoting preventive child care once women are in the facility. Higher incentive payments are thus warranted for services that are more important in terms of leading to better health outcomes, and where the provider effort required to improve those services is high.

Second, the choice of indicators is important. programs might consider paying more for verifiable process indicators of clinical content rather than structural indicators of quality of care. Process indicators are more closely related to outcomes, they are measurable and feasible to pay for, they are solely within the control of the provider, and the evidence shows that providers respond to the incentives for process quality. In addition, the Rwandan program may consider high payments for starting prenatal care in the first trimester and the completion of 4 visits as opposed to the current policy of paying for the first visit and the completion of 4 visits. While 95 percent of women have some prenatal care, few start in the first trimester. Since most start in the 5^{th} or 6^{th} month, there is little time to complete the recommended 4 visits.

Third, no P4P scheme is an all-in-one solution for bottlenecks in health service delivery and utilization, and complementary interventions and strategies (on the demand and supply side) may be needed to increase health care utilization and quality. In fact, financial incentives to other players in the health system might complement incentives to the providers. In the case of early prenatal care utilization, one possibility is to provide incentives directly to pregnant women themselves through conditional transfers. Indeed, there is substantial evidence that conditional cash transfers to families increases their use of preventive medical services[48,49,50] and improves health outcomes [51](GERTLER 2004 AER). Another feasible intervention might be to give community health workers a financial incentive to identify pregnant women in the community and convince them to seek their first prenatal care visit during the first trimester and complete at least 4 prenatal care visits during the pregnancy. In Rwanda, this has potential, as there is already a well-developed network of community health workers working in the catchment area of health facilities.

Limitations

There are a number of limitations to the study. First, the original randomized designed was changed due to the political decentralization process and this could have inadvertently caused some confounding bias in the estimates. This is one of the challenges for effectiveness studies in the context of the scale-up of national programs. However, while small-scale efficacy trials are easier to control, they only inform us about what is possible given the best-case scenario. Efficiency studies, such as this rigorous evaluation of a national health program, are important for external validity and inform us about what is the likely impact in the best and worst conditions. Given that the sample is well balanced on observable characteristics and outcomes at baseline, we can be confident in the results of the difference-in-difference analysis.

The second limitation is we are not able to directly measure more of the outputs paid by the PBF program (unclear what this means ?), such as number of malnourished child referred. We are relying on the prenatal care quality and tetanus vaccinations as proxies for the full range of process quality outputs. It is possible that the results (which results) on prenatal care quality cannot be extrapolated to other areas.(??? what is the link between this sentence and the previous one)

ACKNOWLEDGEMENTS

We gratefully acknowledge funding from the World Bank's Bank-Netherlands Partnership Program, the British Economic and Social Research Council, the Government of Rwanda through a PHRD grant, and the World Bank's Spanish Impact Evaluation Fund. Dr. Basinga also gratefully acknowledges financial support from Global Development Network and the MacNamara Foundation. We are indebted to Claude Sekabaraga, Louis Rusa, Rigobert Mpendwanzi, Kathy Kantengwa, Gyuri Fritsche, and Bruno Meesen for useful comments on the study design, discussions on program design and implementation, and continued support of the study. We also thank Anastasia Gage, Paul Hutchinson, Nancy Mock, as well as mentors and participants from the Global Development Network, for valuable comments. We also thank the Belgian Technical Cooperation, Cordaid, GTZ, Healthnet, USAID, and the World Bank and their staff for cooperating in the implementation of the PBF rollout plan and supporting the evaluation. For data collection efforts, we thank the National University of Rwanda School of Public Health and CIEDEP. Finally, we gratefully acknowledge Elena Samonte-Hinckley, Rosangela Bando, Rafael Ramirez and Ricardo Cavazos for research assistance.

Table 1: Output Indicators (U's) and Unit Payments for PBF Formula

	OUTPUT INDICATORS	Amount paid per unit (US$)
Visit Indicators: Number of …		
1	curative care visits	0.18
2	first prenatal care visits	0.09
3	women who completed 4 prenatal care visits	0.37
4	first time family planning visits (new contraceptive users)	1.83
5	one-month contraceptive resupply	0.18
6	deliveries in the facility	4.59
7	child (0 - 59 months) growth monitoring (preventive care) visits	0.18
Content of care indicators: Number of …		
8	women who received appropriate tetanus vaccine during prenatal care+	0.46
9	women who received 2nd dose of malaria vaccine during prenatal care	0.46
10	at risk pregnancies referred to hospital for delivery during prenatal care	1.83
11	emergency transfers to hospital for obstetric care during delivery	4.59
12	children who completed vaccinations on time (child preventive care)	0.92
13	malnourished children referred for treatment during preventive care visit	1.83
14	other emergency referrals	1.83

+ Appropriate is defined to any women who obtains her third, fourth or fifth tetanus shot.

Table 2: Services (*S's*) and Weights (*ω's*) Used to Construct the Quality Score (Q) for PBF Formula

	Service	Weight	Share of weight allocated to structural components	Share of weight allocated to process components	Means of assessment
1	General administration	0.052	1.00	0.00	Direct observation
2	Cleanliness	0.028	1.00	0.00	Direct observation
3	Curative care	0.170	0.23	0.77	Medical record review
4	Delivery	0.130	0.40	0.60	Medical record review
5	Prenatal care	0.126	0.12	0.88	Direct observation
6	Family planning	0.114	0.22	0.78	Medical record review
7	Immunization	0.070	0.40	0.60	Direct observation
8	Growth monitoring	0.052	0.15	0.85	Direct observation
9	HIV services	0.090	1.00	0.00	Direct observation
10	Tuberculosis service	0.028	0.28	0.72	Direct observation
11	Laboratory	0.030	1.00	0.00	Direct observation
12	Pharmacy management	0.060	1.00	0.00	Direct observation
13	Financial management	0.050	1.00	0.00	Direct observation
	Total	1.000			

Table 3: Health Facility Baseline (2006) Characteristics

	Treatment		Control		Difference	P-Value*
	Mean	**Std dev**	**Mean**	**Std dev**		
Expenditures and Budget Shares						
Log Total Expenditures (2006)	15.81	(1.04)	15.61	(1.01)	0.200	0.418
Log Total Expenditures (2008)	16.91	(0.71)	16.99	(1.08)	-0.083	0.568
Personnel Budget Share	0.46	(0.23)	0.49	(0.26)	-0.031	0.555
Medical Supplies Budget Share	0.22	(0.19)	0.20	(0.19)	0.013	0.705
Non-medical Budget Share	0.32	(0.25)	0.30	(0.22)	0.018	0.726
Staffing						
Medical Doctors	0.05	(0.23)	0.05	(0.27)	0.003	0.940
Nurses	6.31	(6.90)	5.48	(3.30)	0.826	0.409
Other Clinical Staff	4.13	(3.09)	4.47	(4.05)	-0.335	0.554
Non-clinical Staff	5.25	(3.56)	5.33	(5.09)	-0.076	0.901
Observations	80		86			

Note: All variables, except Log Expenditures 2008, are measured at baseline (2006).

*P-values are for cluster-adjusted t-test (continuous variables).

Table 4: Maternal Sample Baseline (2006) Characteristics

	Treatment		Control		Difference	p-value*
	Mean	**Std Dev**	**Mean**	**Std Dev**		
Maternal characteristics						
Age	30.89	(7.05)	31.22	(6.85)	-0.33	0.594
Primary education or more (=1)	0.42	(0.03)	0.35	(0.04)	-0.01	0.889
Currently married/union (=1)	0.94	(0.23)	0.91	(0.29)	0.04	0.214
Partner present (=1)	0.98	(0.14)	0.97	(0.17)	0.01	0.325
Number of pregnancies (Parity)	4.32	(2.46)	4.33	(2.43)	-0.01	0.969
Number of living children	3.39	(1.93)	3.51	(2.62)	-0.12	0.490
Health insurance (=1)	0.55	(0.50)	0.52	(0.50)	0.04	0.591
Household characteristics						
Number of household members	5.15	(1.70)	5.40	(1.94)	-0.25	0.145
Own land (=1)	0.93	(0.25)	0.87	(0.33)	0.06	0.126
Value of Assets (household, animals, farm and enterprise equipment)						
Quartile 1 (=1)	0.22	(0.42)	0.22	(0.42)	0.00	0.994
Quartile 2 (=1)	0.25	(0.43)	0.25	(0.43)	0.00	0.885
Quartile 3 (=1)	0.27	(0.45)	0.28	(0.45)	-0.01	0.885
Quartile 4 (=1)	0.25	(0.43)	0.24	(0.43)	0.01	0.772
Maternal care utilization						
Any prenatal care (=1)	0.95	(0.21)	0.96	(0.20)	-0.01	0.773
Started in first trimester	0.11	(0.32)	0.09	(0.29)	0.02	0.544
4 or more prenatal care visits (=1)	0.18	(0.38)	0.11	(0.31)	0.07	0.036
Number of prenatal care visits	2.76	(0.84)	2.62	(0.80)	0.14	0.180
Delivery in facility (=1)	0.35	(0.48)	0.36	(0.48)	-0.01	0.801
Quality of prenatal care						
Total Quality Score	0.46	(0.02)	0.47	(0.02)	-0.01	0.6661
Standardized Total Quality Score	-0.13	(0.06)	-0.10	(0.06)	-0.03	0.7348
Tetanus vaccine (=1)	0.71	(0.45)	0.67	(0.47)	0.04	0.309
Observations	620		670			

*P-values are for cluster-adjusted t-test or chi-squared tests of difference.

Table 5: Children 0-5 years old Baseline (2006) Characteristics

	Treatment		Control		Difference	P-Value*
Number of Observations	**1384**		**1482**			
Variable	**Mean**	**Std Dev**	**Mean**	**Std Dev**		
Child characteristics						
Age (in months)	30.45	(0.58)	29.95	(0.58)	0.50	0.51
Female (=1)	0.50	(0.01)	0.49	(0.01)	0.01	0.59
Parental characteristics						
Mother present (=1)	0.97	(0.01)	0.96	(0.01)	0.01	0.13
Mother's age	31.62	(0.33)	32.08	(0.32)	-0.45	0.38
Mother's years of education	5.57	(0.19)	5.61	(0.20)	-0.03	0.99
Father present (=1)	0.87	(0.02)	0.90	(0.02)	-0.03	0.30
Father's age	36.10	(0.39)	36.58	(0.40)	-0.48	0.33
Father's years of education	6.06	(0.31)	6.27	(0.34)	-0.20	0.65
Household characteristics						
Number of household members	5.06	(0.09)	5.21	(0.09)	-0.15	0.26
Number of members 0-5 years old	2.18	(0.03)	2.21	(0.03)	-0.03	0.44
Own land (=1)	0.87	(0.02)	0.92	(0.02)	-0.05	0.15
Value of Assets (household, animals, farm and enterprise equipment)						
Quartile 1 (=1)	0.22	(0.03)	0.23	(0.02)	-0.01	0.77
Quartile 2 (=1)	0.28	(0.03)	0.25	(0.02)	0.03	0.27
Quartile 3 (=1)	0.26	(0.04)	0.28	(0.03)	-0.02	0.71
Quartile 4 (=1)	0.23	(0.02)	0.24	(0.02)	-0.01	0.71
Child preventive medical care	0.22	(0.03)	0.23	(0.02)	-0.01	0.77
Visit by child 0-23 months in last 4 weeks (=1)	0.21	(0.02)	0.24	(0.03)	-0.03	0.56
Visit by child 24-59 months in last 4 weeks (=1)	0.08	(0.02)	0.14	(0.02)	-0.05	0.11
12-23 month old is fully immunized (=1)	0.62	(0.06)	0.67	(0.06)	-0.05	0.52

*P-values are for cluster-adjusted t-test or chi-squared tests of difference.

Table 6: Estimated Impact of PBF on Maternal and Child Health Care Services

	N	β	(95% C.I.)	P-Value
Maternal Care Utilization*				
Any prenatal care (=1)	2223	0.002	(-0.022 - 0.025)	0.894
Made 4 or more prenatal care visits	2223	0.010	(-0.063 - 0.083)	0.790
Institutional delivery (=1)	2108	0.074	(0.006 - 0.142)	0.033
Quality of Prenatal Care**				
Standardized Quality Score	3683	0.14	(0.015 - 0.265)	0.030
Tetanus vaccine during prenatal visit (=1)	2810	0.054	(0.007 - 0.100)	0.024
Child Preventive Care Visit in last 4 weeks***				
Visit by child age 0-23 months (=1)	1973	0.134	(0.045 - 0.224)	0.004
Visit by child age 24-59 months (=1)	3645	0.106	(0.050 - 0.161)	0.000
Child 12-36 months is fully immunized (=1)	732	-0.065	(-0.178 - 0.047)	0.248

* Coefficients for the treatment effect were estimated while controlling for a year dummy, facility fixed effects, individual-level characteristics (age, education, number of children, civil status, presence of partner, health insurance) and household characteristics (number of household members, value of assets, land ownership and distance from the facility). Standard errors were adjusted for clustering at the health facility level.

** Coefficients for the treatment effect were estimated while controlling for provider-level characteristics (age, gender, competency score), patient-level characteristics (age, education, civil status, insurance enrollment). Standard errors were adjusted for clustering at the health facility level. The number of observations in the tetanus model is less than in the quality score model because tetanus is only given to women with 5 pregnancies or less.

*** Coefficients for the treatment effect were estimated while controlling for individual-level characteristics (age, gender, insurance enrollment), parental-level characteristics (mother/father present, age, education) and household characteristics (number of household members, value of assets, land ownership and distance from the facility). Standard errors were adjusted for clustering at the health facility level.

References

(1) Horton R. Countdown to 2015: a report card on maternal, newborn, and child survival. *Lancet* 2008 Apr 12;371(9620):1217-1219.

(2) World Bank. Global Monitoring Report. 2006.

(3) Gottret P and Schieber G, 2008, Health Financing Revisited, The World Bank

(4) World Bank 2004, Making Services Worl For Poor People, World Development Report

(3) Filmer D, Hammer J, Pritchett L. Health Policy in Poor Countries: Weak Links in the Chain. 1998;1874.

(4) Kalk, A., J. K. Mayindo, et al. (2005). "Paying for health in two Rwandan provinces: financial flows and flaws." Tropical Medicine and International Health **10**(9): 872-878 *Reinikka Ritva and Svensson Jakob (2001).* Explaining Leakage Unique survey data are used to explain variation in leakage of Public Funds across schools. POLICY RESEARCH WORKING PAPER 2709 . The World Bank Development Research Group Public Services for Human Development

5) Chaudhury N, Hammer J, Kremer M, Muralidharan K, Rogers FH. Missing in action: teacher and health worker absence in developing countries. *J.Econ.Perspect.* 2006 Winter;20(1):91-116.

World Development Report 2004 "Making Services Work for Poor People" The World Bank

(6) Levine R, Eichler R. Performance Incentives for Global Health. 2009.

Brenzel L, Caryn Bredenkamp, Joseph Naimoli, Amie Batson, Rachel Skolnik, and Anthony R. Measham, 2009 " TAKING STOCK: WORLD BANK EXPERIENCE WITH RESULTS-BASED FINANCING (RBF) FOR HEALTH" World Bank

(7) Mehrotra A, Damberg CL, Sorbero ME, Teleki SS. Pay for performance in the hospital setting: what is the state of the evidence? *Am.J.Med.Qual.* 2009 Jan-Feb;24(1):19-28.

(8) Grossbart SR. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Med.Care Res.Rev.* 2006 Feb;63(1 Suppl):29S-48S.

(9) Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, et al. Public reporting and pay for performance in hospital quality improvement. *N.Engl.J.Med.* 2007 Feb 1;356(5):486-496.

(10) Glickman SW, Ou FS, DeLong ER, Roe MT, Lytle BL, Mulgund J, et al. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA* 2007 Jun 6;297(21):2373-2380.

(11) Crawley D, Ng A, Mainous AG,3rd, Majeed A, Millett C. Impact of pay for performance on quality of chronic disease management by social class group in England. *J.R.Soc.Med.* 2009 Mar;102(3):103-107.

(12) Rubinstein A, Rubinstein F, Botargues M, Barani M, Kopitowski K. A Multimodal Strategy Based on Pay-Per-Performance to Improve Quality of Care of Family Practitioners in Argentina. *J.Ambul.Care.Manage.* 2009 April/June;32(2):103-114.

Eldridge, C, Palmer, N. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. Health Policy and Planning Advance Access originally published online on February 7, 2009 Health Policy and Planning 2009 24(3):160-166.

(13) Khanduja K, Scales DC, Adhikari NK. Pay for performance in the intensive care unit--opportunity or threat? *Crit.Care Med.* 2009 Mar;37(3):852-858.

(14) Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann.Intern.Med.* 2006 Aug 15;145(4):265-272.

(15) Eichler R, Auxila P, Uder A, Desmangles D. Performance-Based Incentives for Health: Six Years of Results from Supply-Side Programs in Haiti. 2007.

(16) Meessen B, Kashala JP, Musango L. Output-based payment to boost staff productivity in public health centres: contracting in Kabutare district, Rwanda. *Bull.World Health Organ.* 2007 Feb;85(2):108-115.

(17) Meessen B, Musango L, Kashala JP, Lemlin J. Reviewing institutions of rural health centres: the Performance Initiative in Butare, Rwanda. *Trop.Med.Int.Health* 2006 Aug;11(8):1303-1317.

(18) Soeters R, Habineza C, Peerenboom PB. Performance-based financing and changing the district health system: experience from Rwanda. *Bull.World Health Organ.* 2006 Nov;84(11):884-889.

19 Schwartz JB, Bhushan I. Improving equity in immunization through a public-private partnership in Cambodia.*Bull World Health Organ* 2004; **82:** 661–67.

(19) Human Development Indicators 2008 Rwanda.

(20) Rwanda 2005: results from the demographic and health survey. *Stud.Fam.Plann.* 2008 Jun;39(2):147-152.

(21) Ministére de la Santé, République du Rwanda. Guide de l'Approche Contractuelle. Modules de Formation des Acteurs Concernés pour le Démarrage PBF dans Rwanda. 2006.

(22) : Rusa, L. and G. Fritsche (2007). Rwanda: Performance-Based Financing In Health. Emerging Good Practice in Managing for Development Results: Sourcebook -2nd edition, the World Bank**:** 55-60

(22) Ministère de la Santé Rwanda. Module de référence de formation continue en planification familial. A l'usage des formateurs, superviseurs et prestataires au niveau des formations sanitaires. Mars 2008.

(23) Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988 Sep 23-30;260(12):1743-1748.

(24) Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16(3):297-334.

(25) Ministry of Health Contractual approach unit. Quarterly quality assessment grid for health centers. 2008.

(26) Barber S. Does the quality of prenatal care matter in promoting skilled institutional delivery? A study in rural Mexico. *Matern.Child Health J.* 2006 Sep;10(5):419-425.

(27) Barber SL, Bertozzi SM, Gertler PJ. Variations in prenatal care quality for the rural poor in Mexico. *Health.Aff.(Millwood)* 2007 May-Jun;26(3):w310-23.

(28) Barber SL, Gertler PJ. Empowering women to obtain high quality care: evidence from an evaluation of Mexico's conditional cash transfer programme. *Health Policy Plan.* 2009 Jan;24(1):18-25.

(29) Peabody JW, Gertler PJ. Are clinical criteria just proxies for socioeconomic status? A study of low birth weight in Jamaica. *J.Epidemiol.Community Health* 1997 Feb;51(1):90-95.

(30) Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann.Intern.Med.* 2004 Nov 16;141(10):771-780.

(31) Peabody JW, Rahman O, Fox K, Gertler P. Quality of care in public and private primary health care facilities: structural comparisons in Jamaica. *Bull.Pan Am.Health Organ.* 1994 Jun;28(2):122-141.

(32) Jishnu Das, Jeffrey Hammer. Which Doctor? Combining Vignettes and Item Response to Measure Doctor Quality. May 2004;Policy Research Working Paper 3301.

(33) Bell J, Hussein J, Jentsch B, Scotland G, Bullough C, Graham W. Improving skilled attendance at delivery: a preliminary report of the SAFE strategy development tool. *Birth* 2003 12//;30(4):227-234.

(34) Bulatao RA, Ross JA. Which health services reduce maternal mortality? Evidence from ratings of maternal health services. *Trop.Med.Int.Health* 2003 08//;8(8):710-721.

(35) de Bernis L, Sherratt DR, AbouZahr C, Van Lerberghe W. Skilled attendants for pregnancy, childbirth and postnatal care. *Br.Med.Bull.* 2003;67:39-57.

(36) Improve access to quality maternal health services. ; 97; Colombo: Technical Consultation on Safe Motherhood; 1997.

(37) Kayongo M, Butera J, Mboninyibuka D, Nyiransabimana B, Ntezimana A, Mukangamuje V. Improving availability of EmOC services in Rwanda--CARE's experiences and lessons learned at Kabgayi Referral Hospital. *Int.J.Gynaecol.Obstet.* 2006 Mar;92(3):291-298.

(38) Campbell OM, Graham WJ, Lancet Maternal Survival Series steering group. Strategies for reducing maternal mortality: getting on with what works. *Lancet* 2006 Oct 7;368(9543):1284-1299.

(39) Kak N,B. Burkhalter, M. Cooper. Measuring the competence of healthcare providers. 2001;Operations Research Issue Paper 2(1).

(40) Kogan MD, Alexander GR, Kotelchuck M, Nagey DA. Relation of the content of prenatal care to the risk of low birth weight. Maternal reports of health behavior advice and initial prenatal care procedures. *JAMA* 1994 May 4;271(17):1340-1345.

(41) Villar J, Ba'aqeel H, Piaggio G, Lumbiganon P, Miguel Belizan J, Farnot U, et al. WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. *Lancet* 2001 May 19;357(9268):1551-1564.

(42) Villar J, Bergsjo P. Scientific basis for the content of routine antenatal care. I. Philosophy, recent studies, and power to eliminate or alleviate adverse maternal outcomes. *Acta Obstet.Gynecol.Scand.* 1997 Jan;76(1):1-14.

(43) Carroli G, Rooney C, Villar J. How effective is antenatal care in preventing maternal mortality and serious morbidity? An overview of the evidence. *Paediatr.Perinat.Epidemiol.* 2001 Jan;15 Suppl 1:1-42.

(44) Carroli G, Villar J, Piaggio G, Khan-Neelofur D, Gulmezoglu M, Mugford M, et al. WHO systematic review of randomised controlled trials of routine antenatal care. *Lancet* 2001 May 19;357(9268):1565-1570.

(45) Roper MH, Vandelaer JH, Gasse FL. Maternal and neonatal tetanus. *Lancet* 2007 Dec 8;370(9603):1947-1959.

(46) Thwaites CL, Yen LM, Nga NT, Parry J, Binh NT, Loan HT, et al. Impact of improved vaccination programme and intensive care facilities on incidence and outcome of tetanus in southern Vietnam, 1993-2002. *Trans.R.Soc.Trop.Med.Hyg.* 2004 Nov;98(11):671-677.

(47) Rwanda Ministry of Health. Annual report 2007. 2008.

(48) Lagarde M, Haines A, Palmer N. Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries: a systematic review. *JAMA* 2007 Oct 24;298(16):1900-1910.

(49) Fernald LC, Gertler PJ, Neufeld LM. Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades. *Lancet* 2008 Mar 8;371(9615):828-837.

(50) Gertler P. Final report: the impact of Progesa on health. November 2000.

(51) Rivera JA, Sotres-Alvarez D, Habicht JP, Shamah T, Villalpando S. Impact of the Mexican program for education, health, and nutrition (Progresa) on rates of growth and anemia in infants and young children: a randomized effectiveness study. *JAMA* 2004 Jun 2;291(21):2563-2570.