Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation

Paulin Basinga, Paul J Gertler*, Agnes Binagwaho, Agnes L B Soucat, Jennifer Sturdy, Christel M J Vermeersch*

Summary

Background Evidence about the best methods with which to accelerate progress towards achieving the Millennium Development Goals is urgently needed. We assessed the effect of performance-based payment of health-care providers (payment for performance; P4P) on use and quality of child and maternal care services in health-care facilities in Rwanda.

Methods 166 facilities were randomly assigned at the district level either to begin P4P funding between June, 2006, and October, 2006 (intervention group; n=80), or to continue with the traditional input-based funding until 23 months after study baseline (control group; n=86). Randomisation was done by coin toss. We surveyed facilities and 2158 households at baseline and after 23 months. The main outcome measures were prenatal care visits and institutional deliveries, quality of prenatal care, and child preventive care visits and immunisation. We isolated the incentive effect from the resource effect by increasing comparison facilities' input-based budgets by the average P4P payments made to the treatment facilities. We estimated a multivariate regression specification of the difference-in-difference model in which an individual's outcome is regressed against a dummy variable, indicating whether the facility received P4P that year, a facility-fixed effect, a year indicator, and a series of individual and household characteristics.

Findings Our model estimated that facilities in the intervention group had a 23% increase in the number of institutional deliveries and increases in the number of preventive care visits by children aged 23 months or younger (56%) and aged between 24 months and 59 months (132%). No improvements were seen in the number of women completing four prenatal care visits or of children receiving full immunisation schedules. We also estimate an increase of 0·157 standard deviations (95% CI 0·026–0·289) in prenatal quality as measured by compliance with Rwandan prenatal care clinical practice guidelines.

Interpretation The P4P scheme in Rwanda had the greatest effect on those services that had the highest payment rates and needed the least effort from the service provider. P4P financial performance incentives can improve both the use and quality of maternal and child health services, and could be a useful intervention to accelerate progress towards Millennium Development Goals for maternal and child health.

Funding World Bank's Bank-Netherlands Partnership Program and Spanish Impact Evaluation Fund, the British Economic and Social Research Council, Government of Rwanda, and Global Development Network.

Lancet 2011; 377: 1421–28

See **Comment** page 1383

*These authors shared first authorship

National University of Rwanda School of Public Health, Kigali, Rwanda (P Basinga PhD); **University of California, Berkeley, CA, USA** (Prof P J Gertler PhD); **Ministry of Health, Kigali, Rwanda** (A Binagwaho MMed); **and The World Bank, Washington DC, USA** (A L B Soucat MMed, J Sturdy MA, C M J Vermeersch PhD)

Correspondence to:
Prof Paul J Gertler, Haas School of Business, MC1900, University of California, Berkeley, CA 94705, USA
gertler@haas.berkeley.edu

Introduction

Despite a substantial increase in development assistance for health during the past decade, most low-income countries are unlikely to reach the health-related 2015 Millennium Development Goals.[1] Only ten of 67 countries with high child mortality rates are on track to meet the fourth Millennium Development Goal—a two-thirds reduction of mortality in children younger than 5 years by 2015.[2] And, in most developing countries, the rate of decrease in maternal mortality is much lower than the rate needed to achieve the fifth Millennium Development Goal—a three-quarters reduction of maternal mortality rates by 2015. To accelerate progress towards meeting these goals, developing countries need to increase access to and quality of maternal and child health services.

An intervention that shows promise for improving access and quality of such health services is performance-based payment of health-care providers (payment for performance; P4P).[3] P4P schemes provide financial incentives to health-care providers for improvements in utilisation and quality of specific care indicators, and can affect the provision of health care in two ways: by giving incentives for providers to put more effort into specific activities, and by increasing the amount of resources available to finance the delivery of services. However, P4P schemes could have a detrimental effect on a health service. For example, when P4P payments depend on completion of reports, providers might spend more time on administrative duties and less time ensuring that patients receive the best quality care.[4]

In this study we assessed the potential of a P4P scheme to increase use and quality of key maternal and child health services. The impact evaluation was done prospectively in parallel with the rollout of a national P4P programme in Rwanda.

Methods

The Rwandan P4P scheme

In 2005, after encouraging reports from pilot P4P schemes run by non-governmental organisations, the Rwandan Government decided to implement a national P4P scheme to supplement primary health centres' input-based budgets. In this P4P scheme, payments are made directly to facilities and are used at each facility's discretion. The 14 key maternal and child health-care output indicators for which P4P payments are given are listed in table 1. Some of these output indicators are reasons for a visit, such as prenatal care or delivery, whereas others are services provided during a visit, such as tetanus vaccination during prenatal care. The Rwandan Ministry of Health defined the indicators and payments on the basis of national health priorities, available budget, and the experience of non-governmental organisations with previous P4P schemes.[5,6]

Facilities submit monthly activity reports and quarterly requests for payment to the district steering committee, which is responsible for verification of data and authorisation of payment. For referral indicators, a facility must provide verification from the hospital to which a patient was referred, stating that the referral was appropriate and that the patient was treated. The committee verifies all reports by sending auditors to facilities every 3 months on an unannounced, randomly chosen day to verify that the data reported are the same as those in the facility's records. Between October, 2008, and November, 2008, the Ministry of Health did a one-off tracking survey and interviewed roughly 1000 patients to verify the accuracy of the records; it noted that false reporting was less than 5%.[7]

A facility's overall quality is measured as an index of both structural and process-related measures of quality of care for various types of services (table 2).[8] Structural measures are the extent to which the facility has the equipment, drugs, medical supplies, and personnel necessary to deliver a specific medical service, whereas process measures rate the clinical content of care actually provided for services. The specific services and the structural and process indicators are based on the Rwandan clinical practice guidelines.[9–12]

Data used to calculate each facility's overall quality score are collected through the national monitoring system, in which district hospitals monitor and supervise the quality of health centres in their districts. Every 3 months, every facility is visited by a district hospital team on an unannounced, randomly chosen day, and assesses its quality through direct observation and a review of patients' records with a standardised assessment method (table 2). At the end of each visit, the team discusses their findings with the facility's staff and management, and provides recommendations to improve the quality of the facility. A total score—the quality index, which ranges from 0 to 1—is then generated for the overall quality of service provided by the facility each quarter. A facility's quality index is used to calculate the amount of money it will receive through the P4P scheme. If a facility meets all of the quality criteria, the index equals one and the facility receives the full P4P payment, but if the facility does not meet all of the quality criteria, payment is reduced accordingly—eg, if a facility has a quality index of 0·80, it will receive 80% of the payment for the 14 output indicators. In this way, the P4P scheme pays for both facility output and facility quality.

Experimental design

80 health facilities were randomly assigned to start the P4P scheme between June, 2006, and October, 2006

	Amount paid per unit (US$)
Visit and outreach indicators	
Number of curative care visits	$0·18
Number of first prenatal care visits	$0·09
Number of women who had four prenatal care visits	$0·37
Number of first-time family planning visits (new contraceptive users)	$1·83
Number of women who received 1-month resupply of contraceptives	$0·18
Number of deliveries in the facility	$4·59
Number of visits for child (aged <59 months) growth monitoring (preventive care)	$0·18
Content of care indicators	
Number of children who had a complete vaccination course*	$0·92
Number of women who received appropriate† tetanus vaccination during prenatal care	$0·46
Number of women who received a second dose of malaria prophylaxis during prenatal care	$0·46
Number of high-risk pregnancies referred to district hospital for delivery during prenatal care	$1·83
Number of emergency transfers to hospital for obstetric care during delivery	$4·59
Number of malnourished children referred to district hospital for treatment during preventive care visit	$1·83
Number of other emergency referrals to hospital during curative treatment	$1·83

*BCG at month 0; poliomyelitis at months 0, 1·5, 2·5, and 3·5; pentavalent (diptheria, tetanus, pertussis, hepatitis b, and *haemophilus influenzae* type b) at months 1·5, 2·5, and 3·5; and measles at month 9. †A second, third, fourth, or fifth tetanus shot. Data from reference 5.

Table 1: **Output indicators and payments**

(intervention group), and 86 health facilities were assigned to be control facilities and would continue to receive traditional input-based financing for an additional 23 months until the rollout of the scheme was complete (control group). Rollout of the scheme was staggered for logistical reasons. Because the P4P scheme was introduced in treatment districts during a 5-month period, data for this assessment were available for between 23 months and 18 months of exposure, according to when a specific facility was switched to P4P-based funding.

Because our aim was to assess the effect of the incentive-based bonus (P4P) scheme separately from the effect of an increase in financial resources, the amount of resources for the intervention and comparison facilities had to be held constant. Traditional input-based budgets allocated to the facilities in the control group were increased by the average amount of P4P payments that facilities in the intervention group received every 3 months during the 23-month assessment window.

For each of the 166 facilities, we did two surveys: one at study baseline and one 25 months after study baseline. Surveys consisted of facility questionnaires and household questionnaires. For facility questionnaires, women attending each facility on the day of the interview, and who agreed to participate in the study, were questioned. For household questionnaires, 13 households with children younger than 5 years were randomly selected from each facility's catchment area, giving a total of 2158 households. Maternal baseline characteristics data were taken from facility questionnaires and household questionnaires. Child baseline characteristics data were taken from household questionnaires.

For the household sample, we first sampled 13 zones (each containing roughly 15–20 households) from each facility's catchment area. We then listed all households in the sampled zones and selected one household with at least one child younger than 6 years from each zone. If the head of the household refused to participate in the study, the household was excluded from the sample and a second household was selected until the sample was fulfilled.

All surveys were done by trained enumerators hired by external firms specialised in data collection who were masked to whether they were interviewing in an intervention or control area.[13] In the end-of-study survey, households that could not be found or interviewed were replaced with randomly selected households from the same zones.

Outcome measures

This analysis focuses on the following outcome measures intended to measure a subset of the payment indicators in table 1 and table 2: prenatal care visits and institutional delivery, quality of prenatal care, and child preventive care visits and immunisation. Prenatal care quality is measured by compliance of data from the household questionnaires and facility questionnaires with Rwandan clinical practice guidelines. All outcome measures were assessed by use of data collected in baseline surveys and surveys at the end of the study, collected independently from the

	Weight of service in quality index	Share of weight allocated to structural measures*	Share of weight allocated to process measures†	Means of assessment
General administration	0·052	1·00	0·00	Direct observation
Cleanliness	0·028	1·00	0·00	Direct observation
Curative care	0·170	0·23	0·77	Medical record review
Delivery	0·130	0·40	0·60	Medical record review
Prenatal care	0·126	0·12	0·88	Direct observation
Family planning	0·114	0·22	0·78	Medical record review
Immunisation	0·070	0·40	0·60	Direct observation
Growth monitoring	0·052	0·15	0·85	Direct observation
HIV services	0·090	1·00	0·00	Direct observation
Tuberculosis services	0·028	0·28	0·72	Direct observation
Laboratory services	0·030	1·00	0·00	Direct observation
Pharmacy management	0·060	1·00	0·00	Direct observation
Financial management	0·050	1·00	0·00	Direct observation
Total	1·000	..	..	..

*The extent to which a facility has the equipment, drugs, medical supplies, and personnel necessary to deliver the listed service. †The clinical content of care (appropriate delivery of services during consultation). Data from reference 5.

***Table 2*: Services and weights used to construct the quality score for P4P formula**

	Intervention (n=620)	Control (n=670)	Difference (p value)
Maternal characteristics			
Age <20 years (%)	3%	2%	0·01 (0·322)
Age >35 years (%)	29%	31%	−0·02 (0·573)
Primary education or higher (%)	10%	11%	−0·02 (0·471)
Lives with partner (%)	94%	91%	0·04 (0·211)
Number of pregnancies (mean [SD])	4·32 (4·58)	4·33 (5·24)	−0·01 (0·969)
Household characteristics			
Health insurance (%)	55%	52%	0·03 (0·668)
Number of people in household (mean [SD])	5·15 (2·76)	5·40 (3·08)	−0·25 (0·145)
Distance between household and health facility (km; mean [SD])	3·21 (6·32)	3·39 (7·56)	−0·18 (0·663)
Ownership of land (%)	93%	87%	0·06 (0·126)
Maternal care use			
Any prenatal care (%)*	95%	96%	−0·01 (0·774)
Four or more prenatal care visits (%)*	18%	11%	0·07 (0·029)
Number of prenatal care visits (mean [SD])*	2·76 (1·58)	2·62 (1·80)	0·14 (0·180)
First prenatal care visit in first trimester (%)	11%	9%	0·02 (0·547)
Institutional delivery (%)*	35%	36%	−0·01 (0·801)
Delivery attended by qualified provider (%)	27%	28%	−0·01 (0·729)
Quality of prenatal care			
Tetanus vaccine during prenatal visit (%)	71%	67%	0·04 (0·331)
Total quality score (mean [SD])†	0·45 (0·45)	0·46 (0·49)	−0·01 (0·616)
Standardised total quality score (mean [SD])†	−0·13 (1·68)	−0·10 (1·83)	−0·04 (0·681)

Sample sizes are given in the column heading unless otherwise indicated. *Intervention group sample size=617; control group sample size=658. †Intervention group sample size=931; control group sample size=987.

***Table 3*: Maternal sample baseline characteristics**

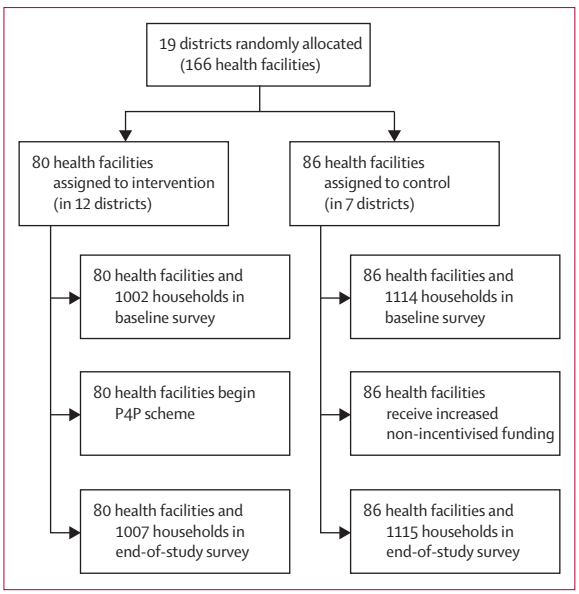

Figure: **Study design**

operation of the P4P programme. Detailed information about the measurement of outcomes is provided in the webappendix (pp 1–2).

See Online for webappendix

Written consent was obtained when possible; however verbal informed consent was accepted. For household questionnaires, both the head of the household and the women interviewed provided informed consent. The research protocol for this study was approved by the Rwanda National Ethics Committee.

	Intervention (n=1242)	Control (n=1321)	Difference (p value)
Child characteristics			
Age (months; mean [SD])	25·86 (11·01)	26·64 (8·38)	−0·78 (0·060)
Sex (female; %)	50%	50%	0 (0·872)
Parental characteristics			
Mother's height (cm; mean [SD])	157·8 (14·4)	158·1 (16·4)	−0·31 (0·631)
Mother's age (years; mean [SD])	31·07 (12·3)	31·28 (13·5)	−0·22 (0·678)
Mother completed primary education (%)	9%	12%	−0·03 (0·152)
Father lives in household (%)	91%	88%	0·03 (0·330)
Household characteristics			
Health insurance (%)	55%	52%	0·02 (0·745)
Number of household members (mean [SD])	5·28 (3·01)	5·40 (3·36)	−0·12 (0·355)
Number of household members younger than 5 years (mean [SD])	2·01 (1·05)	1·98 (1·15)	0·03 (0·457)
Landowners (%)	93%	88%	0·05 (0·131)
Child preventive care utilisation			
Younger than 23 months—preventive visit in previous 4 weeks (%)	0·21%	0·24%	−0·02 (0·556)
Aged 24–59 months—preventive visit in previous 4 weeks (%)	0·08%	0·14%	−0·06 (0·116)
Aged 12–23 months—fully immunised (%)	0·68%	0·63%	0·05 (0·517)

Table 4: **Child (0–59 months) sample baseline (2006) characteristics**

Randomisation

Administrative districts with pre-existing P4P schemes managed by non-governmental organisations were excluded from this assessment.[14] The remaining districts were then grouped into eight blocks based on rainfall, population density, and livelihood data from the 2002 Census.[15] Blocks covered between two and four districts, depending on district characteristics and size. The blocks were then divided into two sides, and one side of each block was randomly assigned to either the intervention or control group. Randomisation was done by coin toss.

Just before implementation of the baseline survey, the administrative district boundaries were redefined by the government in a decentralisation process.[16] As a result, some of the districts selected for our assessment were combined with districts that already had the exisitng P4P schemes. Because P4P schemes could not be removed from health facilities in which they had already been implemented, and because P4P was managed at the district level, the government enrolled all facilities in newly formed districts that had existing P4P schemes into the first phase of the rollout. As a result of this district reorganisation, we had to switch the assignment (intervention or control) for eight districts from four blocks, and add one block to the sample. In the end, the study's nine blocks included 19 districts in total, of which 12 were assigned to the treatment group and 7 were assigned to the comparison group. Of Rwanda's 401 primary care facilities, 80 were assigned to the intervention groups and 86 to the control group.

Statistical analysis

We estimate a multivariate regression specification of the difference-in-difference model in which an individual's outcome is regressed against a dummy variable, indicating whether the facility received P4P that year, a facility fixed effect, a year indicator, and a series of individual and household characteristics described in table 3 and table 4. We calculated robust standard errors, clustered at the district by year level to correct for correlation of the error terms across facilities within districts. Detailed information about covariates and statistical power of the study is found in webappendix pp 2–3. All statistical analyses were done with STATA (version 10).

Role of the funding source

The sponsors of the study had no role in data analysis, data interpretation, or writing of the report. All authors had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of data analysis. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

The figure shows the study design. 2·1% of households in the intervention group and 1·9% of households in the

control group refused to participate in the interview. 88% of the baseline households were re-interviewed at the end of the trial. Slight differences exist in sample sizes between baseline and follow-up because incomplete household surveys were dropped from the sample in each round. The rate of attrition in the number of households available for a second interview was not statistically different between the treatment group (11·8%) and control group (12·1%; p<0·0001).

Baseline characteristics were not statistically different between facilities in the intervention group and facilities in the control group (table 5). Log total expenditures in 2006 (study baseline) and 2008 (at the end of the study) were also much the same between facilities in the two groups (table 5). Data from the maternal baseline survey showed that a larger proportion of women who gave birth at facilities in the intervention group had had four or more prenatal care visits than had women who gave birth at facilities in the control group; all other baseline characteristics not statistically different (table 3).

We estimated no change with the P4P intervention in the probability that women received any prenatal care or in the probability that they had more than four prenatal care visits (table 6). However, we estimated a significant increase in the probability of institutional delivery, the probability of a woman being given a tetanus vaccination during a prenatal visit, and in standardised total quality scores (table 6).

Baseline characteristics were not statistically different between children in the catchment area of facilities in the treatment group and those in the catchment area of facilities in the control group (table 4). Among facilities in the treatment group, we estimated significant increases with the intervention in the probability that a child aged 23 months or younger visited a health facility for preventive care and in the probability that a child aged 24–59 months had a preventive visit (table 6). No difference was recorded in the probability that a child aged 12–23 months was fully immunised (table 6).

On average, facilities in the intervention group allocated 77% of the P4P funds to increase personnel compensation, amounting to a 38% increase in staff salaries; facilities in the control group allocated 73% of the additional input-based funds to increase personnel compensation (data not shown).

	Intervention group (mean [SD]; n=80)	Control group (mean [SD]; n=86)	Difference	p value*
Expenditures and budget shares				
Log total expenditures (RWFR) (2006)	15·81 (1·04)	15·61 (1·01)	0·20	0·418
Log total expenditures (RWFR) (2008)	16·91 (0·71)	16·99 (1·08)	0·08	0·568
Medical personnel budget share	0·46% (0·23%)	0·49% (0·26%)	0·03	0·555
Medical supplies budget share	0·22% (0·19%)	0·20% (0·19%)	0·01	0·705
Non-medical budget share	0·32% (0·25%)	0·30% (0·22%)	0·02	0·726
Staffing				
Medical doctors	0·05 (0·23)	0·05 (0·27)	0·00	0·940
Nurses (n)	6·31 (6·90)	5·48 (3·30)	0·83	0·409
Other clinical staff (n)	4·13 (3·09)	4·47 (4·05)	0·34	0·554
Non-clinical staff (n)	5·25 (3·56)	5·33 (5·09)	0·08	0·901

All expenditures are in Rwandan francs. *p values are for cluster-adjusted *t* test (continuous variables).

Table 5: **Baseline characteristics of health facilities**

	N	β (95% CI)	p value	% D*
Maternal care use†				
Any prenatal care	2309	0·002 (–0·021 to 0·025)	0·875	0·2%
Four or more prenatal care visits	2223	0·008 (–0·063 to 0·079)	0·825	4·4%
Institutional delivery	2108	0·081¶ (0·015 to 0·146)	0·017	23·2%
Quality of prenatal care‡				
Tetanus vaccine during prenatal visit	2856	0·051‖ (–0·002 to 0·103)	0·057	7·2%
Standardised total quality score	3826	0·157¶ (0·026 to 0·289)	0·020	N/A
Child preventive care use§				
Younger than 23 months preventive visit, previous 4 weeks	1971	0·119** (0·041 to 0·198)	0·004	55·9%
24–59 months preventive visit, previous 4 weeks	2902	0·111** (0·059 to 0·162)	0·000	131·6%
12–23 months fully immunised	872	–0·055 (–0·184 to 0·074)	0·390	–8·1%

N/A=not applicable. *The % D = (β / baseline mean) × 100, where the baseline mean of the dependent variable is the 2006 mean of the treatment group from tables 2 and 3. †The β is the estimated intervention effect controlling for a year dummy, facility-fixed effects, individual-level characteristics (age, education, partner lives in household, and number of pregnancies) and household characteristics (health insurance, number of household members, distance from the facility, land ownership, and assets value quartile). SEs were adjusted for clustering at the district-year level. ‡The β is the estimated treatment effect controlling for a year dummy, facility fixed effects, patient-level characteristics (age, education, partner lives in household, and insurance enrolment), and the source of the information (patient exit interview or household survey). SEs were adjusted for clustering at the district-year 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 five or fewer previous pregnancies. SEs were adjusted for clustering at the district-year level. §The β is the estimated treatment effect, controlling for a year dummy, facility-fixed effects, individual-level characteristics (age, sex), parental-level characteristics (height, age and education of mother, father lives in household) and household characteristics (health insurance, number of household members, number of household members younger than 6 years, land ownership, assets value quartile). SEs were adjusted for clustering at the district-year level. ¶p<0·05. ‖p<0·1. **p<0·01.

Table 6: **Estimated effect of P4P schemes on maternal and child health-care services**

Discussion

In this assessment of the P4P programme in Rwanda, our estimates suggest that P4P led to increased use and quality of several crucial maternal and child health care services, but had no effect on use of prenatal care or on the timely completion of child immunisation schedules. The estimates showed larger effects on services for which facilities receive larger financial incentives and those over which the provider has greater control (eg, prenatal care quality and tetanus vaccination during a prenatal care visit) and are less dependent on patients' health-seeking behaviour (eg, timely prenatal care visits).

In the Rwandan P4P scheme, the highest per-unit payment rate, at US$4·59, is given for institutional deliveries. Anecdotal evidence[17] suggests that because of this substantial per-unit payment, providers not only encouraged women to deliver in facilities during prenatal care encounters, but some also partnered with community health workers to promote institutional delivery. The

Panel: **Research in context**

Systematic review

We searched Medline and Google scholar for articles with the key words "pay for performance" and "performance based financing". We included any article that had quantitative estimates of programme effect, but excluded those that were purely descriptive or qualitative.

Interpretation

Our study is one of few rigorous assessments of a P4P scheme in a low-income setting, and is the first to isolate the effect of P4P incentives from effects associated with increase in resources. We report a statistically significant increase in utilisation of institutional delivery and child preventive care services, and improvements in prenatal care quality.

large estimated increase in visits for preventive care of children can also be attributed to the per-unit payment rates. Although the payment rate per child for a preventive visit is low at \$0·18, payment for referral of a malnourished child to hospital for treatment is high at \$1·83. Because, at the time of this study, almost 50% of children in Rwanda had stunted growth and could be appropriately referred to hospital for treatment,[18] roughly half of the child preventive growth monitoring visits yielded \$0·18 each, but the other half yielded \$2·01 each (payment for the visit plus payment for a referral). The strongest monetary payoff of the scheme is for prenatal care quality. Specifically, every tetanus vaccination and malaria prophylaxis course yields \$0·92, whereas improved compliance with prenatal care clinical practice guidelines raises a facility's overall quality score and thereby the amount of P4P payments actually received.

Despite the promise of P4P schemes, little evidence based on credible comparison groups is available on its effects, and no evidence has been obtained that isolates the effect of incentive-based payment from increased funding that is independent of incentives (panel).[19–22] In fact, a review of published studies[23] noted major methodological flaws that limit the ability to interpret the results as causal effects of P4P schemes. A study of primary care facilities in Haiti, which did not have a comparison group, showed increases in immunisation coverage and the number of attended deliveries after the introduction of performance-based bonus payments.[24] Similarly, studies of early Rwandan P4P pilot programmes (also without comparison groups) recorded large increases in the number of curative consultations and institutional deliveries, but no increase in measles vaccination or the number of people enrolling in family planning programmes.[19–21] A quasi-experimental study in Cambodia, which had a comparison group, showed that PSP schemes increased immunisation rates, but facilities in the intervention group received substantially more resources than did those in the control groups, suggesting that resources rather than the P4P mechanism could explain the positive outcome.[25] Our study is one of few rigorous assessments of P4P schemes in a low-income setting, and the first to isolate the effect of P4P incentives from effects associated with increase in resources. This distinction is important because if P4P achieves its results from increased financial resources rather than incentives, 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 (an estimated US\$0·3 per head[26]). In Rwanda, such administrative costs are 0·8% of total health expenditures per head and 1·2% of public and donor expenditures combined.[27]

The absence of effect of the P4P scheme on prenatal care use is also explained by poor financial incentives. The payment rate for an initial prenatal visit is only \$0·09. Because more than 95% of women at baseline made at least one visit, the \$0·09 payment provides little incentive for a facility to locate and give care to the few women who did not attend for prenatal care. Furthermore, the payment rate for completion of at least four prenatal visits is only \$0·37. Rwandan women tend to start prenatal care late in pregnancy (in most cases in the fifth or sixth month of pregnancy) and have between two and three visits on average; therefore, providers who want to achieve at least four visits for every woman would have to attract women to the facility earlier in pregnancy, and ensure that they attend more frequently than they do at present. Our findings suggest that providers chose not to expend the effort necessary to achieve such attendance for the \$0·37 per-unit payment.

We found no effect of P4P schemes on child vaccination rates; however, at baseline immunisation rates were close to 65%, and the government implemented an intensive national vaccination campaign in 2006 that raised immunisation rates to 78% by the end of our study.[28] An increase beyond the baseline would have needed substantial effort on the part of the providers to enter the community, identify unvaccinated children, and vaccinate them. Moreover, completion of a full immunisation programme for a child takes a lot of effort, with many clinic visits over the course of a year and complex administrative procedures.

One possible alternative explanation for our findings is that increased monitoring and supervision led to increased provider effort. However, monitoring and supervision of facilities is done nationally by district hospitals, independently of the P4P scheme. Moreover, we recorded improvements in only some indicators—if the recorded effects of P4P were attributable simply to monitoring we would expect to see increases in both low-priced and high-priced services. Another possible concern is that the P4P scheme was implemented in the context of a larger health sector reform. From 2001 to 2008, Rwanda implemented several other programmes to address major bottlenecks in the health service delivery system. In 2005, the government implemented a programme (Imihigo) to improve a set of

overall development indicators, including family planning, institutional delivery, and enrolment in health insurance plans.[29] Furthermore, in 2006 the government began a national campaign to rapidly increase coverage of child immunisation, vitamin A supplementation, and insecticide-treated bednets. However, our findings are unlikely to be biased by these programmes because they were implemented nationally and the districts assigned the intervention were mostly chosen randomly. Moreover, we controlled for the enrolment of households in health insurance plans in the analyses.

Our study had several possible limitations. First, the original randomised design was compromised by the decentralisation process, which could have inadvertently caused some confounding bias in the estimates. However, because baseline characteristics were much the same between the individuals and facilities in the two groups, we can be confident that results were not biased. Second, whether the incentive effects for prenatal care quality extend to other services is not clear. Third, as in any observational study, recall errors by individuals who were interviewed could affect the accuracy of our estimates. However, we believe such errors would affect both the intervention and control groups equally. Fourth, the Rwandan P4P programme paid facilities rather than individual practitioners. Because we have no rigorous evidence on the effect of individual incentive payments in a similar context, we cannot assess the relative efficiency of those two options. Finally, the analysis presented here does not include evidence of the effect of P4P schemes on health outcomes.

Our findings show that P4P can increase the utilisation and quality of maternal and child health services and thus accelerate progress towards the Millennium Development Goals for maternal and child health. The analysis also sheds light on the debate about some of the benefits and shortfalls of P4P. Several specific lessons emerge from this study. First, higher payments provide stronger incentives. Second, incentives have a larger effect on services in which providers have more control over delivery, such as prenatal care quality. Therefore, larger incentive payments are warranted not only for services that are more important for the improvement of health outcomes, but also for those in which more provider effort is needed. Third, programmes should pay more for verifiable clinical content indicators. Such indicators are closely related to outcomes, are measurable, and are within the control of the provider. For services that depend more on patients' care-seeking behaviour, the programme could provide financial incentives directly to the patient rather than to the provider. Indeed, substantial evidence exists that conditional cash transfers to families increase preventive care use and improve health outcomes.[30–37] Fourth, another feasible option is to give community health workers an incentive to identify patients and encourage them to visit clinics.

One of the more important findings of this analysis is that the P4P scheme led to increased quality of prenatal care. Although health workers might be able to do a medical procedure or give a consultation (ie, prenatal care), they might not always be willing or motivated to expend the effort to do the procedure fully.[38,39] Evidence from this assessment suggests that, by conditioning the Rwandan P4P scheme on a quality index score, the incentive payment gave providers the motivation to translate their knowledge about prenatal care into better practice. These results are important because better quality care improves health outcomes, and in the case of prenatal care this affects both women and their children.[40–43] Indeed, mothers and children must have the best possible access to care, but the quality of the care they receive is also important.

Contributors

PB contributed collection, management, analysis, and interpretation of data, and searched for published studies. PJG, CMJV, AB, and ALBS designed the study. PJG and CMJV contributed to collection and management of data and to fundraising. ALBS contributed to interpretation of data and fundraising. JS contributed to collection and management of data and fundraising, and searched for published studies. PB, PJG, CMJV, ALBS, and JS provided administrative and technical support. PB and PJG took the lead in drafting the report, while the other authors provided critical comments and inputs. PJG and CMJV revised the paper with critical comments from the other authors.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

We thank the World Bank's Bank-Netherlands Partnership Program, the British Economic and Social Research Council, the Government of Rwanda, and the World Bank's Spanish Impact Evaluation Fund for their financial support; Claude Sekabaraga, Louis Rusa, Rigobert Mpendwanzi, Kathy Kantengwa, Gyuri Fritsche, and Bruno Meesen for useful comments on the study design, discussions on programme design and implementation, and continued support of the study; the Rwandan Ministry of Health, Management Sciences for Health, the Belgian Technical Cooperation, Cordaid, GTZ, Healthnet, USAID, and the World Bank and their staff for their collaboration in the implementation of the P4P rollout plan and supporting the evaluation; Lia Fernald, Anastasia Gage, Paul Hutchinson, Benjamin Loevinsohn, Nancy Mock, and Nancy Padian for valuable comments on early drafts; the National University of Rwanda School of Public Health and CIEDEP for data collection; and Elena Samonte-Hinckley, Rosangela Bando, Rafael Ramirez, and Ricardo Cavazos for research assistance. PB thanks the Global Development Network and the MacNamara Foundation for their financial support.

References

1 Horton R. Countdown to 2015: a report card on maternal, newborn, and child survival. *Lancet* 2008; **371:** 1217–19.
2 United Nations. The millennium development goals report. New York: United Nations, 2010.
3 Levine R, Eichler R. Performance incentives for global health. Washington DC: Center for Global Development, 2009.
4 Luoma M. Increasing the motivation of health care workers. The capacity project technical brief 7, 2006.
5 Ministry of Health, Republic of Rwanda. Guide for performance based financing: training module for actors involved in the implementation of the PBF program. Kigali: Ministry of Health, 2006.
6 Ministry of Health, Republic of Rwanda. Reference module for family planning training: for use of trainers, supervisors and providers at the health facility level. Kigali: Ministry of Health, 2008.
7 Health, Development and Performance (HDP). Rapport d' Enquête de Contre-Vérification par la Communauté dans les Districts de Nyamasheke, Nyaruguru et Rulindo. Kigali, 2008. http://www.pbfrwanda.org.rw/index.php?option=com_docman&task=cat_view&gid=24&Itemid=29&limitstart=45 (accessed March 5, 2011).

8 Donabedian A. The quality of care: how can it be assessed? *JAMA* 1988; **260:** 1743–48.

9 Ministry of Health, Republic of Rwanda. Politiques et Normes de Services de Santé, Vol 1–6. Kigali: Ministry of Health, 2009.

10 Ministry of Health, Republic of Rwanda. Normes du District de Santé au Rwanda. Kigali: Ministry of Health, 2003.

11 Ministry of Health, Republic of Rwanda. Normes du District Sanitaire au Rwanda. Kigali: Ministry of Health, 1997.

12 Ministry of Health, Republic of Rwanda. Standards de prestation des services au Centre de Santé: soins préventifs en SMI/PF/ Nutrition (Volume 1, Première Edition). Kigali: Ministry of Health, 1993.

13 Grosh M, Muñoz J. A manual for planning and implementing the living standards measurement study survey: living standards measurement study working paper, No 126. Washington DC: The World Bank, 1996.

14 Vermeersch C, de Walque D, Sturdy J, Gertler P. Performance based contracting in Rwanda: evaluation design and methods. Washington DC: The World Bank, 2010.

15 Government of Rwanda. 2002—General census of population and housing (population). http://imisrwanda.gov.rw/node/18 (accessed March 4, 2011).

16 MINALOC (Ministry of Local Government). Making decentralized service delivery work in Rwanda: putting the people at the center of service provision. Kigali: Government of Rwanda, 2006.

17 Rusa L, Ngirabega Jde D, Janssen W, Van Bastelaere S, Porignon D, Vandenbulcke W. Performance-based financing for better quality of services in Rwandan health centres: 3-year experience. *Trop Med Int Health* 2009; **14:** 830–37.

18 Rwanda 2005: results from the demographic and health survey. *Stud Fam Plann* 2008; **39:** 147–52.

19 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; **85:** 108–15.

20 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; **11:** 1303–17.

21 Soeters R, Habineza C, Peerenboom PB. Performance-based financing and changing the district health system: experience from Rwanda. *Bull World Health Organ* 2006; **84:** 884–89.

22 Doran T, Fullwood C, Gravelle H, et al. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med* 2006; **355:** 375–84.

23 Eldridge C, Palmer N. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. *Health Policy Plan* 2009; **24:** 160–66.

24 Eichler R, Auxila P, Uder A, Desmangles D. Performance-based incentives for health: six years of results from supply-side programs in Haiti (2007). Washington DC: Center for Global Development; working paper 121.

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

26 Ministry Of Health of Rwanda. Annual Report 2007: Performance-based Financing in the Rwandan Health Sector, CAAC/MOH: 74. Kigali: Ministry of Health, 2008.

27 Ministry of Health, Republic of Rwanda. Annual report 2007: performance-based financing in the Rwandan health sector, CAAC/MOH: 74. Kigali: Ministry of Health, 2008.

28 Ministry of Health, Republic of Rwanda. Annual report 2007. Kigali: Ministry of Health, 2008.

29 Basinga P, Sekabaraga C, Soucat A. Scaling up innovation in services delivery of Primary Health Care: Case of Rwanda. Thailand: Prince Mahidol Award Conference book, 2008: 84–101.

30 Fiszbein A, Schady N. Conditional cash transfers: reducing present and future poverty. Washington DC: World Bank, 2009.

31 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; **371:** 828–37.

32 Barber SL, Gertler PJ. The impact of Mexico's conditional cash transfer programme, Oportunidades, on birthweight. *Trop Med Int Health* 2008; **13:** 1405–14.

33 Gertler P. Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment. *Am Econ Rev* 2004; **94:** 336–41.

34 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; **298:** 1900–10.

35 Morris SS, Flores R, Olinto P, Medina JM. Monetary incentives in primary health care and effects on use and coverage of preventive health care interventions in rural Honduras: cluster randomised trial. *Lancet* 2004; **364:** 2030–37.

36 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; **291:** 2563–70.

37 Barham T. Providing a healthier start to life: the impact of conditional cash transfers on infant mortality (May 2005). http://ssrn.com/abstract=1023786 (accessed March 5, 2011).

38 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; **271:** 1340–45.

39 Leonard KL, Masatu MC. The use of direct clinician observation and vignettes for health services quality evaluation in developing countries. *Soc Sci Med* 2005; **61:** 1944–51.

40 Villar J, Ba'aqeel H, Piaggio G, et al. WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. *Lancet* 2001; **357:** 1551–64.

41 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; **76:** 1–14.

42 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; **15** (suppl 1)**:** 1–42.

43 Carroli G, Villar J, Piaggio G, et al. WHO systematic review of randomised controlled trials of routine antenatal care. *Lancet* 2001; **357:** 1565–70.