Performance-based financing in the health sector of the Democratic Republic of Congo: Impact evaluation report

October 2022

Table of Contents

List of Tables

Acknowledgements

The impact evaluation of the DRC performance-based financing project was led by the principal investigators Günther Fink (The Swiss Tropical and Public Health Institute) and Gil Shapira (HHNGE), who conceptualized the study, oversaw its implementation, conducted the analysis, and wrote this report. The study was designed with Hadia Samaha (HSADR), Gyorgy Fritsche (HSAHN), and Michel Muvudi (HAEH2), who also provided support and inputs at all stages of its implementation. Other task team leaders and team members who provided support and valuable inputs during all stages of the project include Fatima El Kadiri El Yamani (HAEH2), Wezi Msisha (HAEH2), Avril Kaplan (HAWH3), Supriya Madhavan (HAEH2), Luc Laviolette (HMNDR), and Christophe Rockmore (HAWDR).

The study was conducted in close collaboration with the project implementation team of the Health System Strengthening for Better Maternal and Child Health Result Project (PDSS). The team particularly acknowledges the support provided by Dominique Baabo, Baudouin Makima Booto, Didier Ramanana Rahary, Delphin Antwisi, and Calude Sese. We thank the support by the Strategic Purchasing Technical Unit (Cellule Technique de l'Achat Strategique) of the DRC Ministry of Public Health, led by Charlie Tchomba Tulia. The study was also supported by the Department of Studies and Planning of the ministry and the Financing subgroup of the Inter-Donor Health Group.

Excellent research assistantship was provided by Saloua Benami, Salome Drouard (DECHD), Pablo Amor Fernandez (DECHD), and Cyrus Sinai. We would also like to acknowledge the work of Saidou Hamadou Theophile and Gervais Chamberlain Yama (EAWPV) in coordinating the baseline survey.

We thank Stephen Brenner (Center for Disease Control and University of Heidelberg) for his work on developing and piloting the tools for the direct clinical observations of family planning consultations and labor and delivery services. We also acknowledge the valuable contributions by Emma Clarke Deelder in the area of measurement and analysis of quality of care.

The baseline and endline data collections were implemented by the non-governmental organization Medecins d'Afrique. The firm Real-Time Analytics (RTA) provided technical assistance in the deployment of Smart Survey System using Android tablets and Open Data Kit

technology during the implementation of the baseline survey. The midline survey was implemented by a consortium between KIT Royal Tropical Institute and ACT for Performance.

We would like to thank our colleagues who served as peer reviewers for their comments and suggestions. The study concept note was reviewed by Demien de Walque (DECHD), Waly Wane (HAWE2), and Gaston Sorgho (HAWH2). This report was reviewed by Damien de Walque, Peter Hansen (HHNGF), and Isidore Sieleunou (HHNGF).

The financial contributions of the Health Results Innovation Trust Fund (HRITF) and the Global Financing Facility for Women, Children, and Adolescents (GFF) are gratefully acknowledged.

Executive Summary

The Health System Strengthening Project and Performance-Based Financing

The Health System Strengthening for Better Maternal and Child Health Result Project (*PDSS – Le Projet de Développement du Système de Santé*) was launched in 2015 with the objective of improving utilization and quality of maternal and child health services in eleven provinces of out of the 26 provinces in the Democratic Republic of Congo (DRC). The project was introduced in a context characterized by many challenges:

- Mortality and fertility rates in DRC are among the highest in the world.
- Low health expenditure per capita in comparison to other low-income or African countries. The largest sources of health expenditure are external aid and out-of-pocket spending, each accounting for about 39% of the total expenditure (WHO).
- A small share of health workers receives formal salaries, and health facilities rely on user fees to cover both operational costs and staff remuneration.
- Health services are of low quality. Health facilities in many areas have bad infrastructure and low availability of equipment and other medical supplies (structural quality). Studies have also documented low process quality, that is, there is low level of compliance with clinical protocols by providers.

The main component of the project is the implementation of a provider payment system reform through Performance-Based Financing (PBF), launched in the end of 2016. Contracted health facilities receive quarterly payments conditional on the volumes of targeted services provided and on quality of care. *Quantity bonuses* were provided on a fee-for-service basis, for the provision of a package of preventive and curative services with a focus on reproductive, maternal and child health. *Quality bonuses* were calculated based on facilities' performance on a detailed quality checklist, and in proportion to the quantity bonus. The financing received through the intervention did not replace any source of funding previously received by the facilities. Contracted facilities could use up to 50% of their revenue to remunerate their staff and had to use the rest of their revenues to cover other operational costs, procure medical supplies and invest in the facilities.

The program introduced a multi-layered system of supervision, verification, and counter-verification. Quantity indicators were reported by each contracted health facility and verified by

provincial purchasing agencies, through review of facility registries and tracing of patients in a community verification exercise. Health zone and provincial teams determine the quality scores through completion of the quality checklists in health centers and hospitals, respectively. Finally, a counter verification agency validates the verified performance in a sample of facilities each quarter, on both random and risk-based bases, with penalties enforced in cases of gaps between the verified and counter-verified data. Health zone management teams, provincial teams, provincial purchasing agencies, and the counter-verification agency are all under performance contracts with payments conditional on timely completion of their corresponding tasks.

Impact evaluation design

The objective of this impact evaluation is to assess whether the PBF approach impacts utilization and quality of primary health services, in comparison to equivalent amounts of unconditional financing. The rational for comparing outcomes produced by facilities implementing PBF to those produced by facilities who receive equivalent amounts of unconditional financing is to isolate the impact of the PBF incentive mechanisms from the direct impact of the additional resources received through the program. The study therefore doesn't evaluate the project as a whole nor evaluates the impact of PBF in comparison to not providing any financial assistance to health facilities.

The study relies on an experimental design to identify the causal effects of the PBF program. PBF was randomly assigned at the health zone level using a cluster-randomized design. The randomization was conducted in provincial public randomization ceremonies. Within each health zone, all facilities received the same financing model. In the control health zones, facilities received quarterly transfers equaling the average transfer made to facilities in the PBF arm in the same province, adjusted for the population in the catchment areas.

Large-scale household and facility-based surveys were conducted in 2015, prior to the launch of PBF, and in 2021/22, after 5 years of implementation. The household survey data is used primarily to measure health service coverage. The detailed facility-based surveys are used to measure quality of care and other indicators related to facility-functioning. The facility-based surveys included general facility assessments, provider interviews, direct clinical observations of consultations and deliveries, and patient exit interviews.

The study was originally planned to cover 100 health zones in all 11 provinces supported by the project. The final study covered 58 health zones in only 6 of the provinces: Kwango, Kwilu, Mai-Ndombe, Haut-Katanga, Haut-Lomami, and Lualaba. The other provinces were removed due to an Ebola outbreak in 2018, after which health financing was modified in some areas, and given the higher-than-expected survey costs. The change in design doesn't affect the internal validity of the study because randomization was blocked at the province level. However, this change implies lower statistical power.

General trends

The data collected at baseline and endline of this study suggests generally very positive trends for both the PBF and control health zones supported by the project. The coverage of many reproductive, maternal and child health services substantially increased. Data collected at the health facilities demonstrate better structural quality and improved adherence to protocols of antenatal care and child curative care. While other factors could have affected these trends, it is reasonable to assume that the project contributed to these positive changes at least to some extent. Studies from different settings have demonstrated that provision of financing to health facilities can improve outcomes regardless of whether the financing is conditional on performance. This might be particularly important in DRC, given that frontline health providers receive almost no financing and have to rely on user fees from populations with very little resources.

PBF impact on services utilization

The PBF model had positive but moderate impacts on health service coverage. There was no significant impact on child health service coverage relative to the control group receiving matching unconditional facility financing. With respect to maternal health services, coverage in the PBF group was higher for all the indicators, but these differences were only statistically significant for initiation of antenatal care in the first trimester. PBF results were most positive for family planning, with substantial increases in the use of modern contraceptives as well as the likelihood of women discussing family planning with health providers. The overall impacts, however, mask some important heterogeneities. For example, among poorer women (below the household wealth median) there is an impact on the likelihood of delivering in a formal health

facility, while there is no such impact on the wealthier group among which coverage of institutional delivery is 95%. There are also some heterogeneities by region.

PBF impact on quality of health services

The PBF program positively impacted the structural quality of health facilities. Although the positive trends in availability of medical supplies in both PBF and control zones, the PBF program further increased the availability of basic equipment, family products and protocols at the health centers. With respect to measures of infection prevention and control, the programs improved the availability of incinerators and secured boxes for needles and sharps, and the likelihood that health facilities post sterilization and disinfection procedures. The impacts on these structural quality measures are concentrated at the health centers and are weaker in the hospitals, that had higher baseline measures for these indicators.

The PBF program also improved process quality, or adherence to clinical protocols measured through direct clinical observations. In contrast to the findings on structural quality, stronger impacts on process quality measures are estimated at hospitals. Such impacts are found for child curative care, antenatal consultations, family planning consultations, and deliveries. At the health centers, strong positive impacts are estimated for the quality of family planning consultations and weaker impacts are found for the quality of child curative consultations.

PBF impact on user fees and health-related spending

Impacts of the PBF program are also found on user fees and out-of-pocket expenditures, with potential important implications on financial access to health services. PBF facilities are more likely charge flat fees, exempt fees for poor patients, visibly post fee schedules, and have lower official fees. At the user's level, results are a bit weaker – reductions in out-of-pocket expenditures measured during exit interviews were mostly detected for child curative care. However, households in PBF zones report to be less likely to borrow money or rely on support from others to cover health-related costs.

Limitations

This study has some important limitations. First, the study was not designed to assess broader impacts of PBF on the health system such as improvements in data systems and data use, accountability, and capacity of health zones and provincial teams. Second, there are several ways

in which outcomes in the control zones could have been influenced by the implementation of PBF in neighboring health zones. The same provincial health teams supervised both control and PBF zones and in some instances some PBF tools were used in the control zones although that didn't impact financing amounts. There were also movements of facility and health zone staff between PBF and treatment zones. We are unable to estimate to what extent such factors affected our results, but it is reasonable to assume that if such effects exist, it will imply that we under-estimate the impact of the PBF program in the empirical results presented here. Finally, the reduction in the number of provinces covered by the study reduced statistical power and enabled analyzing program impact in only 6 out of the eleven originally included provinces. Heterogeneity analyses show that there can be substantial variations in impacts across provinces.

Further Investigation

The comparison between PBF and unconditional facility financing suggests better coverage and quality outcomes in the PBF health zones. The results are at least as good or better in the PBF group for each type of outcomes. However, the differences between the groups are moderate in some areas. Progress has also been achieved in the control group, suggesting that some constraints to improvement in utilization and quality of care were lifted even without tying financing to performance. PBF is more expensive to implement, primarily because of the costs associated with the provincial purchasing agencies (contract managers and verifiers) and the counter-verification. It is therefore important to understand the main health, systems and societal benefits associated with these additional costs. A cost-effectiveness analysis, building on the findings of this IE and using spending data by the project has been initiated and will shed a light on this question.

The analysis reveals many heterogeneities in PBF impacts. We find stronger impacts on process quality in hospitals relative to health centers, but stronger impacts on the availability of medical supplies and measures of infection prevention and control at the health centers. We also find heterogeneous impacts on service utilization by wealth and region. Investigation of these heterogeneities, with further analysis of the data in combination with qualitative approaches, might reveal important lessons for the design of more impactful PBF programs in DRC.

Chapter 1: Background and Program Design

Country context

The Democratic Republic of Congo (DRC) is the largest country in sub-Saharan Africa, by area, with an estimated population of 89.5 million in 2020 (World Development Indicators, WDI). The country has a long history of conflict, political upheaval, and instability and has a very high poverty rate despite its abundance of natural resources. According to World Bank projection for 2020, 73% of the population lived under the international poverty line of $1.9 USD a day. The country was ranked 175th out of 189 countries in the United Nations Development Program's 2019 Human Development Report.

Mortality and fertility rates in DRC are among the highest in the world. Maternal mortality rate is estimated at 473 per 1,000 live births (World Health Organization 2019) and the estimated under-5 mortality is 81 per 1,000 births (WDI). Total fertility rate is estimated to be 5.8 births per women, trailing only fertility rates in Niger and Somalia (WDI). According to the Multiple Indicator Cluster Survey from 2018, a quarter of women aged 20 to 24 had a live birth prior to age 18. In terms of nutritional status of children, 42, 23, and 7 percent of children are estimated to be stunted, underweight and wasted, respectively.

Health system

The country's 26 provinces are divided into 516 health zones (health districts) that typically have a single general referral hospital, a first level referral hospital for the health centers in the zone. The health zones are further divided into health areas (catchment areas). Each health area is served by a health center providing a basic package of inpatient and outpatient services, including uncomplicated deliveries and minor surgeries. Geographically large health areas have additional health posts that provide a more limited package of services. Some health areas are served by referral health centers, which have the capacity to provide a wider set of services. In addition to the public health sector, there is a large network of faith-based service providers and many private providers, mostly concentrated in urban areas.

The current health expenditure per capita was $21 USD in 2019, in comparison to an average of $34 USD among low-income countries and $79 USD among sub-Saharan African countries (WHO Global Health Expenditure Database). The largest sources of health expenditure are external aid and out-of-pocket spending, each accounting for about 39% of the total expenditure. Government health spending amounts to 15% of the total spending, much of which remains at the central level or is devoted to the administration of the health system. The government spending covers a small share of the income of frontline service providers and public health facilities generally do not receive routine funding or medical supplies from the government. Health facilities, whether public, private, or faith-based, therefore rely on

user fees to remunerate their staff, procure supplies, and cover any other operational costs. Facilities traditionally had considerable autonomy in setting fees. Since 2018, the Ministry of Public Health adopted the flat fee (frais forfaitaire) approach with the goal of standardizing fees within each health zone (Ministere de la sante publique 2018).

With respect to human resources for health, the country's health system is facing several challenges. First, there are many unregulated and low-quality training institutions (Ministere de Sante Publique 2010). Second, there is an unequal distribution of health workers in the DRC. Many health facilities, especially lower-level facilities in rural areas, are understaffed. On the other hand, there are facilities, mainly higher-level and in urban or peri-urban areas, that are grossly overstaffed (Bertone and Lurton 2015). Third, health workers' incomes are fragmented, heterogenous and often volatile. Even in the public sector, most workers are not on the Ministry of Public Health payroll and only receive an allowance called 'prime de risque' (World Bank 2014). Many health providers do not even receive such allowance. For all cadres apart from doctors, facility revenues generated through user fees is the main source of remuneration (Bertone and Lurton 2016).

Coverage of health services in DRC is mixed. Considering the high poverty rate, cost of services, and physical accessibility challenges in many of the provinces, a relatively high share of women receives at least some antenatal care (82%) and delivers in health facilities (82%) (INS 2019). However, only 17% of pregnant women initiate antenatal care in the first trimester of their pregnancy and 43% complete at least four antenatal care visits. Coverage of child vaccination is considerably lower than the coverage of maternal health services. Only 37% of children between 12 and 23 months received all basic vaccinations in 2018 (INS 2019).

The average quality of health services is low. A series of studies analyzing direct clinical observations data collected for the purpose of this impact evaluation document substantial gaps in adherence to clinical protocols and guidelines. For example, out of a sample of 992 observed child curative consultations, less than ten percent of consultations complied with 80% of recommended diagnostic steps, and only 15% verified the presence of general danger signs that are markers of severe or even fatal illness (Perales et al. 2020). Another study showed that providers gave the IMCI-recommended treatment in only 42% of cases of severe febrile diseases, severe pneumonia and severe dehydration (Clarke-Deelder et al. 2019). Observations of labor and delivery services reveal alarmingly low adherence to protocols related to infection prevention and control and postpartum care (Brenner et al. 2022). Although 77% of interviewed women reported to have received any antenatal care during their last pregnancy in a household survey, only 8% reported to receive effective care that included blood pressure measurement, testing of blood and

urine samples, iron supplementation, and counseling on pregnancy complications (Fink, Kandpal and Shapira 2022).

There are also steep wealth gradients in health outcomes, service coverage, and quality of health services. According to the DRC Demographic and Health Survey (DHS) 2013/4, under-5 mortality in the bottom wealth quintile was 54% higher than in the top quintile (MPSMRM, MSP and ICP 2014). Children in the bottom wealth quintile were also four times more likely to be underweight and six times more likely to be severely underweight in comparison to children in the top wealth quintile. Some of the variation in health outcomes could be explained by differential health service utilization. For example, while 96% of women in the top wealth quintile deliver in health facilities, only 69% of women in the bottom quintile do so (INS 2019). Fifty-four percent of children in the top wealth quintile are completely vaccinated in comparison to only 23% of children in the poorest group. On average, poorer users also appear to receive care of lower quality in comparison to their wealthier counterparts, conditional on receiving any care. A study of antenatal care shows that household wealth is associated with greater protocol compliance (Fink, Kandpal and Shapira 2022). This relationship is largely driven by the better quality of health facilities in wealthier areas. However, there is also a within-village wealth-quality relationship that is primarily driven by wealthier women seeking care at higher-quality facilities even if they are more distant.

History of performance-based financing in the health sector

There have been several pilots or implementations of PBF models on a relatively small scale by different development partners in DRC. Evaluations of two of these pilots were published in peer-reviewed journals. The first evaluation is of a small-scale PBF program implemented by Cordaid in two health zones in Sud-Kivu province. Outcomes in the treated health zones were compared to those in two neighboring health zones receiving financing not tied to performance (Soeters et al. 2011). The study concluded that the PBF program resulted in lower out-of-pocket spending for patients who received higher quality services, even though the PBF health zones received assistance of approximately $2 USD per capita compared to $9-12 USD per capita in the control zones. The study, however, provides minimal details on the design of the program and what was implemented in the control zones.

Another evaluation is of the pilot implemented in Haut Katanga province in the context of the World Bank's financed Health Sector Rehabilitation Support Project (*PARSS – Projet d'Appui à la Réhabilitation du Secteur de la Santé).* Contracted facilities were paid for delivery of targeted services on a fee-for-service basis, with an estimated budget of $0.51 USD per capita per year. The scheme was randomized at the health (catchment) area level, with control facilities receiving fixed payments determined by the number of health workers and their status. The study found that PBF facilities charged lower user fees and exerted more effort on outreach activities. Nevertheless, utilization in the PBF areas

was slightly *lower* than in the control areas, leading to lower facility revenues and lower satisfaction by providers. Below, the differences between the PARSS PBF model and the model evaluated by this study are elaborated.

The PDSS PBF Program

The Health System Strengthening for Better Maternal and Child Health Results (*PDSS – Le Projet de Développement du Système de Santé*), was launched in 2015 with the objective of improving utilization and quality of maternal and child health services in eleven provinces out of the country's 26. The project supports 165 health zones, estimated to have just under a third of the country's population. The main component of the project is the implementation of a provider payment system reform through PBF. In addition, the project provides institutional, financial and technical support to various entities at the national and provincial level to improve governance, health system administration and surveillance capacities.

Program design

In the PDSS PBF program, contracted health facilities receive quarterly payments conditional on the volumes of targeted services provided and on quality of care. The financing received through the intervention, equaling approximately $1.6 USD per capita per year,[1] was meant to provide additional financing to facilities, and thus did not replace any source of funding previously received by the facilities. However, the program introduced rules for how contracted facilities can charge user fees and use their revenues. In total, 166 hospitals and 2,545 health centers were contracted. The performance contracts were signed between the health facilities and provincial purchasing agencies (EUP – *Etablissement d'Utilite Publique*). These independent and not-for-profit entities were created by the project for the management of the contracts and verification of performance. The facility performance payments are transferred directly to facility bank accounts. Health facilities without access to a bank (26% of facilities) received deliveries of cash.

There are two types of performance contracts for health facilities. Health centers are contracted to deliver the Minimum Package of Activities (MPA), a set of core preventive and curative primary health services. General referral hospitals and some referral health centers are contracted to deliver the Complementary Package of Activities (CPA), including services to be delivered at first level referral such as complicated deliveries, blood transfusions, and surgeries requiring anesthesia.

[1] This is the average amount received by facilities in the years 2019-2021 as calculated by the project implementation unit. This amount only includes the transfers to the facilities and doesn't include transfer to administrative units and other program costs.

The PBF program introduced rules for how contracted health facilities can spend their revenues, whether received as PBF payments or through any other source (e.g. user fees). The spending by each facility is guided by a quarterly business plan. In the structured business plan tool, facilities must document their performance on the different indicators in the previous quarter, set targets for the new quarter, and define strategies for achieving these targets. The business plan should be elaborated with technical assistance from the health zone management team and needs to be validated by the local health development committees (CODESA – *Comité de développment de la santé*), that include members of the community. After completing the business plan, the budgeting is determined with the use of the Indexes Tool (*Outil d'Indices*).[2] In the first part of the tool, the facilities enter all sources of funds to determine total revenue in the preceding quarter. In the second part, facilities determine budget allocation that should be aligned with the priorities identified in the business plan. Facilities are allowed to spend a maximum of 50% of each quarter's revenue on personal bonuses for staff. Since 2020, facilities are also required to spend a minimum of 20% of the PBF payments on medications and other consumables. The rest of budget needs to cover all operational costs, investments in facility infrastructure and equipment, and any savings.

The program also dictates how individual performance bonuses are distributed among the facility's staff. Thirty percent of the performance bonus (or up to 15% of the overall quarterly facility revenue) is fixed in the sense that it doesn't depend on individual staff performance. The relative share each staff member receives is determined according to their title (e.g. medical doctor, nurse of a specific grade), responsibility within the facility (e.g. head of maternity, facility in-charge), and seniority. The other 70% of the individual performance bonus is distributed based on an individual performance evaluation of each staff member. Staff performance is scored by the facility-in-charge with a use of a performance checklist. The categories included in the checklist are professionalism (punctuality, attendance, and attire), team spirit (interpersonal relationship, teamwork, devotion, and initiative), competency (organization, quality of work, and workload), growth (responsiveness to previous feedback), and number of days worked. Once all the staff members are evaluated, the relative individual scores determine the share everyone gets.

To account for the challenging geographical context with lack of a road network in most areas, and the unequal distribution of the population and health providers, the performance payments paid to facilities are adjusted using so-called "equity classifications". Base performance fees were adjusted by up to 80% according to classification of both health zones and health facilities, as shows in **Table 1**. Health zones were divided into five equity classes depending on their distance from the provincial administrative center, the size and density of their population, and the support received by financial and technical

[2] The indexes tool is described in more detail in Fritsche, Soeters and Meessen (2014).

partners.[3] Then, health facilities within each zone were also divided into five equity categories according to the facilities' distance from the zone's hospital and the office of the health zone team, its population size and density, and the number of qualified health workers. For both types of classifications, 1 represented the least remote category and 5 represented the most remote category. The classification was conducted by the province health teams, with support of the provincial purchasing agencies. Initially, all health centers were classified in category 3. For each health center classified in category 5, another health center within the zone had to be classified in category 1. For each health center classified in category 4, another health center within the zone had to be classified in category 2.

Table 1: Performance payment adjustment according to health zone and health facility equity classification

Health zone equity classification	Health center equity classification: 1	2	3	4	5
1	100%	110%	120%	130%	150%
2	110%	120%	130%	140%	150%
3	120%	130%	140%	150%	160%
4	130%	140%	150%	160%	170%
5	140%	150%	160%	170%	180%

Note: The table presents the percentage of the base performance fees health facilities of different equity classifications receive. Facilities of equity classification 1 in health zones with classification 1 receive the base performance fees while facilities of all other classifications receive augmented fees.

The PBF intervention also aimed to address equity concerns at the local level by providing facilities higher financial incentives for treating the poorest in each health area. The poorest users are identified by a community targeting exercise led by Indigent Committees that were formed in the catchment area of contracted facilities. In each health area, up to 5% of households can be identified as indigents, who receive free services. In addition, facilities received higher performance payments for delivering certain services to indigent patients (Table 2). By the end of 2021, 57% of the contracted health facilities formed indigent committees and completed identification the members of the communities eligible for the fee exemptions.

Quantity Indicators (paid services)

Provision of preventive and curative services was incentivized on a fee-for-service basis, with fees set prospectively. The specific services and their relative weights are presented in **Table 2** and **Table 3**.[4] Facilities contracted to provide the MPA (health centers) were initially paid for 18 services. Services targeting child health included vaccination and growth monitoring consultations for children between 6

[3] Health zones and facilities serving areas with lower population density received higher bonuses.
[4] The rewards are presented in terms of weights because the actual monetary reward depend on the equity classifications.

and 59 months. Indicators targeting maternal health include initiation of antenatal care during the first three months of pregnancy, at least two tetanus toxoid injections during pregnancy, deliveries attended by skilled birth attendants, and postnatal consultations between three and seven days after delivery. There are two indicators incentivizing use of hormonal family planning methods, so that longer-term contraceptive methods receive a higher weight than short-term methods. One indicator incentivizes facilities for new users of contraceptive pills or injectables. Another indicator, with a higher weight, rewarded for new users of intrauterine device or implants. In terms of HIV-related indicators, the program incentivizes voluntary consulting and testing, prevention of mother to child transmission (PMTCT), and care for newborn of HIV positive women. The program also paid for detection of infectious tuberculosis cases and completion of their treatment.

While the program focused on improving coverage of reproductive, maternal and child health services, in addition to HIV and tuberculosis care, facilities were also rewarded for provision of general curative care. Health centers were paid for outpatient consultations, with a higher reward for treatment of identified indigent patients. Minor surgeries that do not require anesthesia were also rewarded. Finally, health centers were paid for referral of severe cases, as long as the facility could provide a referral confirmation from the higher-level facility. Finally, health centers were also paid for overseeing home visits conducted by community health worker (*relais communautaires*). Using a standardized tool, the CHWs were instructed to identify health needs of the households, assess water safety and hygiene practices, and to enquire about use of health services. CHWs were expected to encourage the household to use health services and adopt best practices.

Table 2: PBF quantity indicators and corresponding relative weights - Minimum Package of Activities

	Before Midterm Review (2016-2019)		**After Midterm Review (since 2019)**	
Category	**Indicator**	**Relative Weight**	**Indicator**	**Relative Weight**
Outpatient consultation	New case	1	Same	1
	New case for indigent patient	3	Same	4
Surgery	Minor surgery (without need for anesthesia)	5	Same	3
Referral	Referral of a severe case to a higher-level facility	15	Same	9.2
Vaccination	Child completely vaccinated in first year of life	25	Same	8.3

Antenatal care	Two or more Tetanus Toxoid injections during pregnancy	3	Same	2
	First consultation in first three months of pregnancy	2	Same	2
			Fourth Antenatal care visit	1.7
			Third dose of intermittent preventive treatment (Sulfadoxine/Pyrimethamine)	2
Delivery	Delivery with skilled birth-attendant with a filled partograph	17	Same	9.2
Postnatal care	Consultation between 3 and 7 days after delivery	4	Same	2
Family planning	New user of pills or injectables	13	Same	5
	New user of IUD or implants	15	Same	5
Growth Monitoring	Child 6-23 months	0.8	Same	1
	Child 24-59 months	0.4	Same	0.7
Home visit		2.5	Same	0.8
HIV	Voluntary counselling and testing	3	Same	1
	PMTCT	20	Same	10
	Care for newborn of HIV+ woman	25	Same	13.3
			Biannual follow-up with individuals on antiretroviral regime	20
Tuberculosis	Detection of positive case	75	Same	36.7
	Case treated and recovered	150	Same	75

Note: The rewards for the quantity indicators are presented in terms of weights because facilities received different dollar amounts based on their equity classification. For benchmarking, the base performance fee for outpatient consultation was $0.6 USD after the midterm review, and the base fee for an outpatient consultation for an indigent patent was $2.4 USD. For facilities in the highest equity classification, the corresponding fees were $1.1 and $4.4.

Hospitals and referral health centers contracted to deliver the CPA were initially incentivized for 22 indicators. Unlike the health centers that were only rewarded for outpatient care and minor surgeries, the higher-level facilities were also paid for provision of inpatient care, major surgeries requiring anesthesia, and blood transfusions. Higher rewards were given for inpatient care and surgeries provided to indigent patients. The facilities were paid for the timely initiation of antenatal care and for three types of deliveries: uncomplicated (Eutocic) deliveries, complicated (Dystocic) deliveries excluding caesarians, and caesarians. In addition to the two family planning indicators included in the MPA, the CPA also

rewarded male and female sterilization. With respect to HIV-related services, higher-level facilities were rewarded for the number of new users of antiretrovirals and antiretroviral follow-ups, in addition to the three indicators in the more basic package. Just like the MPA the CPA also rewarded detection and treatment of tuberculosis cases.

Following the midterm review of the project, some indicators and corresponding weights were revised. In both the MPA and CPA, two additional indicators related to antenatal care were added in 2019. These include the number of women receiving at least four antenatal consultations and the number of women receiving a third dose of intermittent preventive treatment against malaria (Sulfadoxine/Pyrimethamine). For the MPA facilities, an indicator was added for the number of HIV positive individuals on antiretroviral care who complete follow-up visits at least once every six months. At the health center level, the relative weights assigned to outpatient care of indigents and growth monitoring increased, while the weights for referral, vaccination, deliveries, and HIV and Tuberculosis care were reduced. In the CPA contract, the weights for outpatient care, major surgery, timely antenatal care, and anti-retroviral follow-up visits increased.

Table 3: PBF quantity indicators and corresponding relative weights – Complementary Package of Activities

	Before Midterm Review (2016-2019)		**After Midterm Review (since 2019)**	
Category	**Indicator**	**Relative Weight**	**Indicator**	**Relative Weight**
Outpatient consultation	Outpatient care for a referred patient	5	Same	10
Inpatient care	Day of hospitalization	1	Same	1
	Day of hospitalization for indigent patient	3	Same	3
Surgery	Minor surgery	6	Same	5.8
	Minor surgery for indigent patient	10	Same	10
	Major surgery	25	Same	50
	Major surgery for indigent patient	75	Same	75
Blood transfusion	Blood transfusion	6	Same	5.8
Antenatal care	First visit in first three months of pregnancy	2	Same	3
			Fourth Antenatal care visit	2

			Third dose of intermittent preventive treatment (Sulfadoxine/Pyrimethamine)	2
Delivery	Uncomplicated delivery	15	Same	15
	Caesarian delivery	70	Same	70
	Complicated delivery (Excluding Caesarian)	30	Same	30
Family planning	New user of pills or injectables	8	Same	8,3
	New user of IUD or implants	15	Same	15
	Sterilization (Masculine or feminine)	50	Same	50
HIV	Voluntary counselling and testing	2	Same	1.7
	PMTCT	15	Same	15
	Care for newborn of HIV+ woman	20	Same	20
	Initiation of antriretroviral treatment	15	Same	15
	Antiretroviral client follow-up	10	Same	25
Tuberculosis	Detection of positive case	75	Same	75.8
	Case treated and recovered	150	Same	150.8

Note: The rewards for the quantity indicators are presented in terms of weights because facilities received different dollar amounts based on their equity classification. For benchmarking, the base performance fee for inpatient care at the hospitals was $0.6 USD per day after the midterm review, and the base fee for such care for an indigent patent was $1.8 USD. For hospitals in the highest equity classification (category 5), the corresponding fees were $0.8 and $2.4.

Quality Indicators

The quality checklists used to determine facilities' overall quality score contained a range of indicators divided into sub-categories. The checklists contained both structural quality indicators, such as availability of different equipment and drug items, and indicators more closely related to process quality, most of which were related to health providers adherence to clinical and administrative guidelines. The checklists were completed through review of documents and registries, verification of the availability of different supplies, and interviews with health providers.

Appendix Table 1 and Appendix Table 2 present the different categories, examples of indicators in each category, and the weight given to each category in the overall score for facilities contracted to provide the MPA and CPA. Both checklists are divided into 15 categories with the MPA checklist having 374 indicators while the CPA checklist having 411. Fourteen of the categories are identical although with different content and relative weights in determining the final quality score. The MPA checklist includes an additional category for vaccination while the CPA checklist includes a category for surgery. In the

MPA checklist, the category with the biggest weight is outpatient care, accounting for 28 percent of the overall score while all other categories account for less than ten percent. The CPA checklist allocates 25 percent of the overall score to inpatient care, 17 percent to outpatient care and 13 percent to surgery.

After the project's midterm review, some changes were introduced also to the quality checklists. First, many categories were broken down into an even larger number of indicators. Second, the relative weights of the different categories changed. At the MPA level, for example, higher weight has been assigned to the management plan, laboratory, inpatient care, maternity, vaccination and HIV/TB categories. The relative weight of the outpatient care category reduced the most.

The quality bonus was determined proportional to the *quantity* payment. Facilities that scored less than 50 percent on the quality checklist did not receive any quality bonus. Facilities contracted to deliver the MPA could receive a maximum bonus of 25 percent, while PCA facilities could get a maximum bonus of 40% of the total quantity-based transfers made. The overall quarterly payment can be represented as follows:

$$\text{Quarterly Payment for facility j} = \begin{cases} e_j \sum_i s_{ij} f_{ij} & \text{,if } Quality\_Score < 50\% \\ e_j \left(\sum_i s_{ij} f_{ij} \right) \left(1 + \gamma_j Quality\_Score_j\right) & \text{,if } Quality\_Score > 50\% \end{cases}$$

where s_{ij} represents the quantity of incentivized service i that facility j provided during the quarter and f_{ij} represents the fee-per-service associated with the service, depending on the package of incentivized services. e represents the equity adjustment and ranges between 1 and 1.8. That is, $e_j \sum_i s_{ij} f_{ij}$ is the *quantity payment*. $Quality\_Score_j$ is the quarterly quality score in a range between 0 and 100 percent. γ_j equals 0.25 for the MPA facilities and 0.4 for the CPA facilities.

Supervision and verification

The program introduced a multi-layered system of supervision, verification and counter-verification. Quantity indicators were reported by each contracted health facility and verified by the provincial purchasing agencies (EUPs). These agencies verify information each quarter by reviewing health management information system reports and facility registries. Then, a random sample of patients are drawn from the facility registries for community verification. Three percent of the patients in the registries for each service are drawn, but the overall number cannot exceed 50 per facility per quarter. The sample is passed to contracted community associations in each health zones. For each listed patient, the community associations verify whether the person exists, whether they received the service, what price they paid, and ask about satisfaction with the received service. Health zone management teams (ECZS – *Equipe cadre zone de sante*) and the Provincial Health Division (DPS – *Division Provincial de la Sante*)

determine the quality scores through completion of the quality checklists. The ECZSs manage this process for health centers, whereas the DPSs do this for hospitals. Health zone management teams, provincial teams, and the provincial purchasing agencies are all under performance contracts with payments conditional on timely completion of their corresponding tasks.

A third-party counter-verification agency (ACVE - *Agence de Contre-Vérification Externe*) validates the verified performance in a sample of facilities. Counter-verification is both random and risk-based in areas where there are large disparities between claimed data reported by the health facility data verified by the EUPs. Due to the high costs associated with travel within the country, counter-verification is conducted in four provinces each quarter. Within each of these four province, five health zones are selected. In each selected health zone, the counter-verification is conducted in the zone's hospital and in four selected health centers. In total, counter-verification is conducted in 100 facilities each quarter. In case of disparities, penalties are applied according to rules set in the operations manual. In a first occurrence of a discrepancy in volumes reported (register review in the health facility) of over 10%, or inability to trace more than 5% of the patients chosen for community verification, a 50% penalty is applied to the quarterly performance payment. In a second occurrence, a 100% penalty is applied while the purchase contract is frozen until the facility management has been replaced. In a third occurrence, the performance contract with the health facility is halted. For the quality score, a discrepancy of over 10% will lead to penalties for both the health facility as for the health zone or province team that carried out the evaluation.

Launch and scale-up

The scale-up of the PBF approach was gradual. PBF was first introduced in the provinces of Kwango, Kwilu, and Mai-Ndombe in the last quarter of 2016. In the first quarter of 2017, the program launched in Equateur, Mongala, Sud-Ubangi, Tshuapa, and Maniema. In Haut-Katanga, Haut-Lomami, and Lualaba, PBF was introduced in the third quarter of 2017. The impact evaluation focused exclusively on Kwango, Kwilu, Mai-Ndombe, Haut-Katanga, Haut-Lomami and Lualaba.

At the beginning of the project, all contracted facilities received financial and material support, to enable them to work towards the performance goals.[5] First, facilities received an investment bonus equaling $2500 USD per health center, $3000 USD per referral health center, and $5000 USD per hospital. Second, each facility received a transfer of medications financed by the project and a transfer of family planning products financed by United Nations Population Fund (UNFPA). Although the medications and

[5] Receipt of the financial assistance was provided after the elaboration of a first business plan and signature of the contract.

family planning products were planned to be delivered at the launch of the program, some facilities received these inputs several years later, due to logistical complications.

Difference with the PARSS PBF model

The PBF model introduced in the context of the PDSS project is different from the model implemented in the Haut Katanga pilot in several important ways. The overall budget for the performance payments has substantially increased from $0.5 USD to more than $1.6 USD per capita per year. The program also pays for quality of care and not just for quantity of services. The PDSS PBF model introduced performance contracts at all health administration levels to strengthen verification and governance. Management tools (business plan, indices tools, individual performance evaluation) were introduced for management of resources and remuneration of staff. To address inequity, differential incentive amounts depend on equity assessment. Finally, investment bonuses were given to facilities in the beginning of implementation.

Chapter 2: Methods and data

Study objectives

The primary research question of the impact evaluation is:

What are the effects of the PDSS PBF approach on utilization and quality of primary health services, in comparison to equivalent amount of unconditional financing?

The rational for comparing outcomes produced by facilities implementing PBF to those produced by facilities who receive equivalent amounts of unconditional financing is to isolate the impact of the PBF incentive mechanisms from the direct impact of the additional resources received through the program. The study therefore doesn't evaluate the PDSS project as a whole nor evaluates the impact of PBF in comparison to not providing any financial assistance to health facilities.

There are multiple channels through which the PBF intervention can theoretically improve utilization and quality of health services and improve health outcomes. Under PBF, health providers are expected to exert more effort in the delivery of priority and high quality maternal and child health (MCH) service outputs in a client-focused manner. In order to attract a larger number of users, providers are also expected to reduce fees (whether official or not). The quality checklists can orient health providers in quality improvement and provide structure for supervision. The supervision and verification activities are expected to increase accountability and can further encourage managers to work for results. The business plan and indexes tools are intended to improve management by guiding facility-in-charges in identifying priorities, allocating resources, and evaluating staff performance.

Empirical strategy

The study relies on an experimental design to identify the causal effects of the PBF package. PBF was randomly assigned at the health zone level using a cluster-randomized design. Within each health zone, all facilities received the same treatment. In PBF zones, health centers and hospitals received quarterly transfers depending on the quantity of incentivized services and the facilities' scores on the quality checklist, as described above. In the control health zones, facilities received quarterly transfers equaling the average transfer made to facilities in the PBF arm in the same province, adjusted for the population in the catchment areas. In both treatment arms, the transfer

amounts were adjusted according to the equity classification (according to the scales presented in Table 1).

Apart from the conditionality of the financing, there were several additional differences between the treatment groups, summarized in **Table 4**. Facilities in the control zones did not receive supervision with the use of the structured quality checklist and there was no verification of their reported service volumes, at the facility or community level. Health zone teams in the PBF zones were paid according to a performance contract, in which completion of the quality checklists in each contracted facility is the main component. Health zone teams in control zones received unconditional quarterly financing equaling to the average amount received by the PBF health zone teams.

Table 4: Differences between treatment arms

	PBF	Control
Quarterly facility financing transfers	Conditional on the quality and quantity of services and on remoteness.	Unconditional on facility performance. Amount of transfer equal to the average amount received by the PBF facilities in the same province, adjusted for catchment population and remoteness.
Supervision	Enhanced supervision using structured quality checklists, verification of registries and community verification.	Routine supervision by health zone and province teams.
Facility spending guidelines	Facilities must elaborate a quarterly business plan prior to deciding on budget allocation. A maximum of 50% of facility revenue can be distributed as personal performance bonuses. A minimum of 20% of the PDSS transfer must be spent on medication and other consumables.	Since 2019, facilities must spend a minimum of 20% of the PDSS transfers on medications and other consumables. The project did not change spending guidelines otherwise.
Personal performance bonus	Distribution of personal performance bonuses determined with the use of the indexes tool. Thirty percent is determined according to title, responsibility and seniority. Seventy percent determined based on individual performance evaluation.	No guidelines introduced by the program.
Initial investment bonus at the beginning of implementation	Yes. 2500 USD to health centers, 3000 USD to referral health	Yes. 2500 USD to health centers, 3000 USD to referral health centers and 5000 USD to hospitals.

	centers and 5000 USD to hospitals.	
One-time shipment of medications and family planning products	Yes	Yes
Financing of health zone teams	According to a performance contract. Transfers depend on the completion of the quarterly quality verification in all contracted health facilities	Unconditional. Amount of transfer equal to the average amount received by health zone teams in the same province.

Facilities in the control zones also didn't have to adhere to the elaboration of a quarterly business plan and the guidelines on budget allocation and provision of individual performance bonuses. On the other hand, all PDSS-supported facilities (PBF and control facilities) received the initial investment bonus and the one-off deliveries of medications and family planning products. Since 2019, control facilities were required, as PBF facilities, to spend a minimum of 20% of their quarterly transfer on medications and other consumables.

Transfer amounts to facilities in both treatment arms are recorded in the program's PBF portal. Analysis of the data suggest that while there are quarters in which one group receives higher transfers, the average transfer amounts in the 2017-2021 period are similar in both groups.

Randomization

The study was originally planned to cover 100 health zones in 11 provinces. These were all the health zones originally supported by the project where PBF was not already piloted in the context of the previous World Bank health system strengthening project. The randomization was blocked by province. In Kwilu, the 22 health zones were split into 2 remoteness classes, and randomization done within these two blocks.

For transparency, the randomization was conducted in provincial public randomization ceremonies in the presence of representatives from all health zones. A representative of each zone drew a folded note from a transparent bucket, determining PBF assignment (see

Picture 1**).** The results of the random assignment were then written in a form signed by representatives of the PDSS project team, province health team, civil society, and development partners. The map in **Figure 1** presents the results of the randomization.

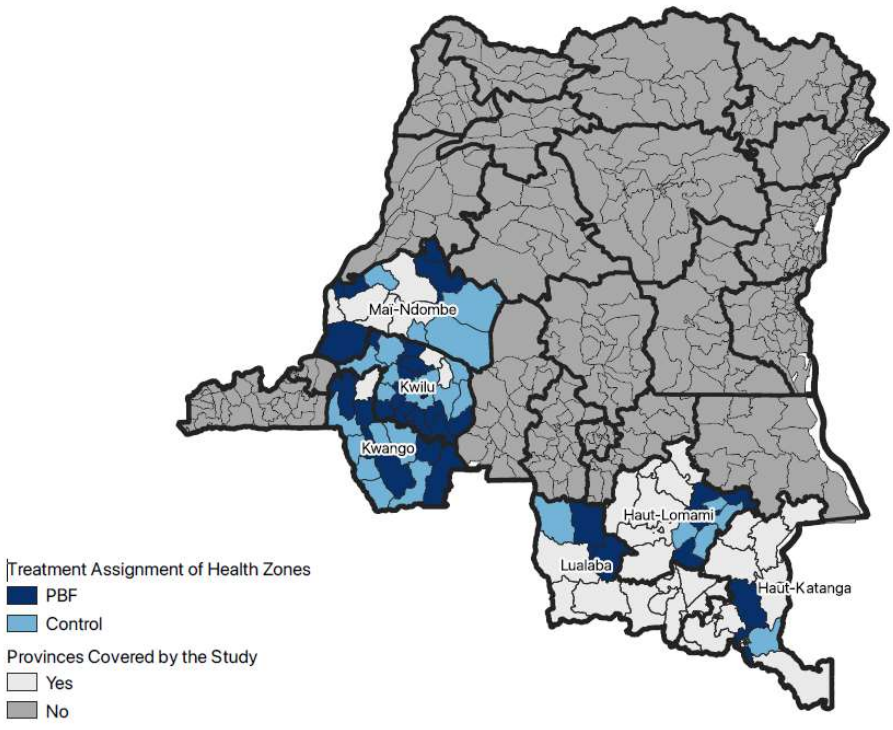
Picture 1: Public randomization ceremony in Kwango Province

Note: a representative of a health zone draws treatment assignment from a transparent bucket during a public randomization ceremony in Kwango province in December 2015.

While the random assignment of the financing approaches was strictly adhered to, there are several factors are important to keep in mind. First, individuals were not prevented from seeking care in neighboring health zones. While crossing of health zone boundaries is likely limited due to the large average geographical area of each health zone and average distances to neighboring zones, it is clearly possible that some households receive care from facilities in another zone. Second, there is movement of health providers and health zone administrators between health zones. Individuals who start in PBF facilities might acquire knowledge and practices that they

could use in control facilities. Third, and more generally, some transfer of knowledge and skills across zones seems likely. The project management team did, for example, observe the use of some of the PBF tools in control areas. These include the business plan tool and the indexes tool used to determine allocation of bonuses among staff. It is important to note that the use of these tools was not encouraged by the PDSS project or affected payments amount. These were initiatives of facilities, health zones, or provincial teams. Finally, there was an expectation that control facilities would be under the PBF approach after two years of project implementation. Given the delayed launch of the intervention and some of the circumstances described below, the PBF approach was not extended to the control health zones. Given that the control facilities were familiar with the program, it could be that they took actions to better prepare themselves to perform on the incentivized indicators that they wouldn't have taken if they didn't expect to be under the PBF regime eventually. All of the listed sources of control contamination listed above are expected to reduce the differences between the treatment and control groups. The evaluation results should therefore be interpreted as lower bound estimates of the true PBF impacts.

Figure 1:Assignment of PBF treatment at the health zone level



Pay-for-knowledge pilot

To further focus on quality of care, an additional randomized controlled trial was embedded into the PBF evaluation. Facilities selected for this intervention received quarterly knowledge (vignette) assessments, which then accounted for fifty percent of the PBF quality score and payment. The vignette study only took place in PBF zones in the provinces of Kwango, Kwilu, and Mai-Ndombe. All PBF hospitals in these provinces received the vignette intervention. In addition, 2-3 health centers in each health zone were randomly selected for the vignette intervention. More information on the intervention and on the design of the experiment are presented in Fink et al. (2022). This report focuses solely on the overall impacts of PBF, while the vignette intervention impact will be presented in a separate report later. It is worth highlighting here that the vignette intervention was only done in PBF areas, and thus may have contributed to observable differences in quality of care. This is more likely to affect outcomes at the hospital level, as 23 out of the 30 hospitals in the PBF group participate in the pilot. About a third of the PBF health centers are participating in the pilot. Excluding the health centers participating in the pay-for-knowledge pilot yielded very similar results overall, although the precision is reduced with the smaller sample.

Changes in study design

When the concept note for the impact evaluation was written and approved in the end of 2014, it was expected that the study would cover 100 health zone and that health facilities would start receiving payments in 2015. The original design included a baseline survey in 2015, a midline survey in 2017 to measure short-term impacts, and an endline survey in 2019. This original design had to be adapted several times due to the following reason.

First, the launch of the PBF program was delayed. As described above, the project was launched and scaled up between the fourth quarter of 2016 and the third quarter of 2017. Moreover, although the program launched, not all elements of implementation such as the counter-verification were immediately in place. A major challenge in the beginning of implementation was the long duration of the cycle of payments. That is, facilities received the payments many months after their performance scores were determined. Given these delays, the impact evaluation and task teams decided to postpone the midline survey to 2018 and to only cover the provinces of Kwango, Kwilu, and Mai-Ndombe. It was also decided to only measure short-term impacts at the facility level and to not conduct household survey during that survey.

In May 2018, there was an Ebola outbreak in Equateur province. As part of the project's response to the outbreak, some health zones in the Northwest of the country received additional financing to maintain health services and for recovery. Given this change in financing and the higher-than-expected survey costs incurred during the baseline, it was decided to remove from the study 39 health zones in the provinces of Equateur, Mongala, Sud-Ubangi, and Tshuapa. It was subsequently decided to also remove from the study's sample the three participating health zones from Maniema province. The reduction in the number of covered health zones from 100 to 58 reduces the statistical power of the study. As PBF was allocated at the health zone level, the standard errors in all of the regressions are clustered at the health zone level and the expansion of the sample sizes in the remaining zones would have limited impact on power. Nevertheless, as the randomization was blocked by province, the removal of provinces does not impact the internal validity of the study.

As the team was preparing for implementation of the endline survey in 58 health zones in 2020, the COVID-19 pandemic required suspension of operations. Preparations were renewed in 2021 after developing protocols for the protection of survey teams and respondents, renewal of ethical clearance, and approvals from local authorities. Fieldwork was further delayed by health workers strikes in 2021. Figure **2**: Study timeline portrays the timeline of the study.

Figure 2: Study timeline

		2015		2016				2017				2018				2019				2020				2021				2022	
Provinces	**Activity**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**
Kwango, Kwilu and Mai-Ndombe	Survey	Baseline																								Endline			
	PBF																												
Haut-Katanga, Haut-Lomami, and Lualaba	Survey	Baseline																											Endline
	PBF																												

Approvals, ethical clearance, and trial registration

The study was designed through a series of consultations with the project team, the Ministry of Public Health and development partners operating in the country. The concept note was peer reviewed in the World Bank and approved in September 2014.

Ethical clearance for each round of survey was provided by the National Health Ethics Committee. The endline survey was also reviewed and approved by the Swiss Ethics Committee (EKNZ).

The study was registered with the American Economic Association's Registry for Randomized Controlled Trials with ID AEARCTR-0002880.

Data and sampling framework

The evaluation relies on two main sources of data: household surveys and facility-based surveys. As described above, the main outcomes of interest for the evaluation were utilization and quality of health services. Service utilization was measured at the household level while quality of care was measured at health facilities. To facilitate linkage between the household and facility data, households were selected for the survey from within the catchment areas of the surveyed health centers.

In 2014, provincial health teams provided lists of the health areas in each health zone covered by the study. The lists also included the names of the facilities officially serving each area and a list of villages or urban neighborhoods in each area. These lists were used by the evaluation team for the sampling of both facilities and villages. In the first stage of sampling, 5 health areas were randomly selected within each health zone. The health center servicing the health area was added to the health facility sample. Within each selected area, a village or an urban neighborhood were randomly selected for the household survey.

To determine the required number of households to sample in each village or neighborhood, data from the 2007 Demographic and Health Survey were used, which was the most recent household data available at the time. Given that institutional deliveries were already highly common in the target areas, the study was powered to detect difference in percentage of children fully vaccinated. The rate of fully vaccinated children was 45% and the intra-class correlation was 0.1 at the cluster level, and 0.035 at the provincial level. It was assumed that sampling from five different villages in each health zone would allow bringing down the intra-class correlation to 0.05. A sample of 10 household per enumeration area would enable detecting an effect of 0.075 with power 0.8 and an effect of 0.081 with power 0.9 with an alpha of 0.05.

The baseline survey was fielded in the 3rd and 4th quarters of 2015 in 100 health zone originally included in the study. The sample target was 500 health centers and 5,000 households with at least one woman with a pregnancy in the two years preceding the survey (10 in each of the 500 catchment areas of the

covered health centers). In addition, hospital assessments were conducted in a random sample of 60 of the health zones. In the 58 health zones included in this final analysis, baseline surveys were conducted in 290 health centers, 29 hospitals, and with 2890 households. Because of the focus on maternal and child health, households were eligible for inclusion in the survey if they had a female member between the ages of 15 and 49, who was pregnant in the two years preceding the survey. The survey was conducted by Medecins d'Afrique, an NGO chosen through a competitive selection process.

A midline survey was conducted in the third quarter of 2018 only in Kwango, Kwilu, and Mai-Ndombe provinces that were the first provinces to start PBF implementation. The survey included only assessment of health centers. Households and hospitals were not covered. The same health centers visited during the baseline survey were re-visited to assess short-term impacts of the program. The target sample was 220 health centers (5 health centers in each of the 44 health zones in the three provinces). The survey was implemented by a consortium between KIT Royal Tropical Institute and ACT for Performance. Data from this survey is not used for most of the analysis presented in this report as we focus on the longer-term impacts of the program in 58 health zones.

The endline survey was conducted in two phases. It took place in 44 health zones in Kwango, Kwilu and Mai-Ndombe provinces in the third and fourth quarters of 2021. Fieldwork was conducted in 14 health zones in Haut Katanga, Haut Lomami and Lualaba in the second quarter of 2022. Both household surveys and health facility assessments were conducted. The same health facilities and villages visited during the baseline survey were re-visited. Given the change in the number of health zones (clusters), new power calculations were conducted using the baseline survey data. The calculations assumed power 0.8 and an alpha of 0.05. A sample of 12 household per village (and 60 per health zone) would enable detecting an effect of 0.1 on the share of children 13-24 months old receiving the third dose of the pentavalent vaccine, and an effect of 0.07 on the share of recently pregnant women who initiated antenatal care in the first trimester of their pregnancy.As during the baseline survey, the household survey targeted women with recent pregnancies. In order to reflect the greater focus on family planning and adolescent reproductive health by the World Bank's Health Nutrition and Population program in the country, additional samples were added. First, a sample of women 15-49 years old *without* a recent pregnancy was added to measure broader program impacts on use of family planning methods. Second, households with women of ages 15-20 were oversampled.

Survey tools

The survey instruments were based on those developed for the Impact Evaluation Toolkit of the Health Results Innovation Trust Fund.[6] The instruments were adapted to the country's context by the evaluation team, project team, and the baseline survey firm, in a process that included consultations with the different departments of the Ministry of Public Health and representatives of different development partners. The survey tools of baseline survey are available in the World Bank Microdata Library.[7] The survey tools and data from the endline survey will be uploaded to the Microdata Library by the 4th quarter of 2022.

Household Survey Tools

The household survey was composed of two main section. For the first, the household head or most knowledgeable household member available at the day of the survey were interviewed. For the second component, women 15 to 49 years old were interviewed about their health and that of their children under the age of 5. In addition, the survey teams weighed and measured the height of all children aged under 5 years of the interviewed women present in the household during the survey team's visit.

The main themes of the general household section were:

- Household composition
- Economic situation of the household
- Health expenditure and utilization of health services by the household

The main themes of the women interviews were:

- General health and case management of illness
- Fertility and family planning
- Care provided during recent pregnancies
- Interaction with health workers
- Immunization and vaccination of children
- Birth histories and child survival

[6] https://www.rbfhealth.org/resource/impact-evaluation-toolkit-provides-hands-guidance
[7] Baseline survey tools can be found at the following links:
https://microdatalib.worldbank.org/index.php/catalog/9333;
https://microdatalib.worldbank.org/index.php/catalog/9334

Health Facility Assessment

The main component of the facility-based survey is the facility assessment that seeks to collect data on key aspects of facility functioning and different aspects of quality of care. The main respondent for this module was the individual in charge of the health facility at the time when the survey team visits the health facility. The main themes covered by the facility assessment included:

- Facility staffing
- Facility infrastructure and equipment
- Availability of drugs, consumables and supplies at the health facility
- Supervision
- Record keeping and reporting to the Health Management Information System
- Facility management and autonomy
- Facility financing and user fees

In each facility, health workers with reproductive, maternal and child health service delivery responsibilities were randomly selected for a provider survey. The main themes covered by the provider survey tool included:

- Role and responsibilities of the interviewed health worker
- Compensation, including delays in salary payments
- Staff satisfaction,motivation and mental health
- Technical knowledge on Maternal and Newborn Health

While structural quality of care was measured through the general facility assessment, clinical quality of care was measured through the implementation of direct clinical observations. Enumerators with medical training used structured observation tools to record content of services without intervening. Specifically,

- Curative consultations for children under 5 were observed during the baseline, midline, and endline.
- Antenatal consultations were observed during the baseline, midline, and endline.
- For the endline survey, observations of family planning consultations and deliveries were added. The tools for these observations were developed by Dr. Stephan Brenner from

Heidelberg University and were piloted in four health zones during the implementation of the midline. Due to the associated costs, observations of labor and delivery services were only conducted in a subsample of facilities in the provinces of Kwango, Kwilu and Mai-Ndombe.[8]

Finally, patient exit interviews were conducted with the patients whose consultations were observed to collect background information on the patients and assess the patients' perception of quality of care. Respondent also reported on fees paid and content of care. If the patient is a child, the child's caregiver was interviewed. Exit interviews were not conducted for the observed deliveries.

Empirical specification

Exploiting the experimental design of the evaluation, the impacts of the PBF intervention on outcomes of interest are estimated using the following linear regression model:

$$y_{iz} = \beta_0 + \beta_1 * PBF_z + \gamma * B_z + \varepsilon_{iz},$$

where y_{iz} is an outcome for health facility/health worker/consultation/ individual i in health zone z, measured during the endline survey. PBF_z takes value 1 if the health zones is in the PBF group and 0 otherwise. B_z represents a vector of randomization block dummies, corresponding to provinces except for zones in Kwilu province that were divided into two blocks. ε_{iz} is a normally distributed error term, clustered at the health zone level. For some outcomes, explained below, we prefer an adjusted model with additional covariates. These include characteristics of facilities, providers, and patients.

We use the simple regression specification above in the report, rather than a difference-in-differences approach, for the following reasons. First, the use of the baseline data should improve the statistical power of the analysis, but in practice there are minimal changes in precision. Second, the difference-in-difference approach is not feasible for many outcomes given changes in the outcomes measured in the different survey rounds, and differences in sampling of hospitals and direct clinical observations. Therefore, given the good balance at baseline, we

[8] Within each health zone, facilities with higher volume of deliveries during the midline survey were prioritized. Many health centers have low volume of deliveries which would require the survey teams to spend more days on average in each facility to complete observations. The quality of delivery services is expected to be higher in higher-quality facilities.

prefer to report results from the more intuitive approach using only the endline data, rather than switching back and forth between specifications. In appendix 2, we show that the difference-in-difference results are very similar to those from our benchmark regression specification.

Internal validity: testing for baseline balance

The identification of the PBF impacts using the experimental design relies on the assumption that outcomes in the PBF and control groups would be identical in the absence of the PBF intervention. To assess whether this assumption is reasonable, we test whether the different groups were similar in baseline in terms of sample's characteristics and pre-program outcomes. In **Appendix Table 3**: Baseline Balance in household indicatorswe present means by treatment group and p-values of mean difference tests. The test is conducted with the same regression specification presented above with the baseline data. That is, we regress the different variables on a treatment coefficient, controlling for the randomization block dummies and clustering standard errors at the health zone level. Since only a subsample of hospitals were covered by the baseline survey, we use additional administrative data for the checks presented in Appendix Table 5. Statistically significant differences are detected for a very small number of indicators, suggesting that the randomization achieved a balance between the treatment groups on observable characteristics.

Chapter 3: Results

In this chapter, we present the experimental results on the impact of the PBF approach, in comparison to the unconditional financing received by the health facilities in the control health zones. For indicators collected in both the baseline and endline surveys, we also show the overall trends. While it is reasonable to assume that the investments and activities carried by the project are likely to have contributed to the trends, we cannot isolate the projects' impacts from other factors. The changes between baseline and endline should therefore be simply interpreted as trends in PDSS-supported health zones, capturing both generic improvements in indicators over time and benefits of additional financing.

Impacts on coverage of reproductive, maternal and child health services

Coverage of health services is measured at the household level. As can be seen in **Table 5**, positive trends were seen for almost all the services. The proportion of children 13-24 months who received the BCG vaccine and the third dose of the pentavalent vaccine increased by 6 and 17 percentage points between 2015 and 2021/22, respectively. The proportion of women who received any antenatal care and delivered in formal health facilities increased to above 90 percent. While antenatal care is almost universal, the share of women who initiate care in the first trimester increased to only 22 percent and we do not see a change in the share of women receiving four or more consultations. On the other, we observe an increase of 11 and 27 percentage points in the share of women who receive antenatal care that includes a tetanus shot and malaria chemoprophylaxis in pregnancy, respectively. The share of women who receive postnatal care after leaving the health facility where they delivered more than quadrupled but remained at only 41 percent. With respect to family planning services, there is a large increase in the percentage of women who report to have ever discussed family planning with a health provider but a much more modest increase in the percentage of women reporting to be current users of modern family planning methods. The service that stands out in terms of its trend is growth monitoring, that remains very low at 4 percent.[9]

We do not find that coverage of child health services in the PBF group is higher than that in the control group. There is also no difference in the share of mothers reporting their children under the age of five to have been measured to determine their nutritional status and on vaccination indicators.

[9] Women were asked whether the nutritional state of each child was measured during the 6 preceding months. In consultations that took place after the data collection, it was suggested that the phrasing of the question might have not been sufficiently clear. We cannot assess to what extent that might have affected the survey results.

With respect to maternal health services, coverage is higher among the PBF group for all services, but the difference is only statistically significant for one indicator. Women in the PBF zones are more likely by 8 percentage points (p-value = 0.01) to initiate antenatal care during the first trimester of their pregnancy. However, there are no significant impacts on attendance of a minimum of four consultations, or on receiving tetanus shot and anti-malarial drugs during antenatal care. In addition, there are no significant impacts on the coverage of institutional delivery and postnatal care.

Table 5: Impacts on reproductive, maternal and child health service coverage

	Baseline			Endline				PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.		PBF Impact	Standard Error	p-value
Child health services [a]										
Children under 5: growth monitoring in past 6 months	0.04	0.05	3175	0.03	0.04	5025		0.01	0.02	0.495
Children 13-24 months: All basic vaccinations	0.26	0.33	835	0.51	0.48	1402		0.00	0.05	0.977
Children 13-24 months: BCG vaccination	0.88	0.88	836	0.93	0.96	1429		0.03	0.03	0.351
Children 13-24 months: Penta 3	0.57	0.58	839	0.75	0.68	1429		-0.03	0.05	0.576
Maternal care during most recent pregnancy [b]										
Any ANC [c]	0.85	0.86	2251	0.92	0.93	4135		0.02	0.01	0.207
4+ antenatal consultations	0.42	0.44	2250	0.34	0.36	4134		0.04	0.05	0.435
ANC in first trimester	0.13	0.15	2244	0.18	0.28	4135		0.08***	0.02	0.001
ANC with tetanus shot	0.69	0.73	2251	0.78	0.80	4135		0.03	0.03	0.206
ANC with anti-malarial	0.44	0.48	2251	0.71	0.71	4135		0.01	0.04	0.784
Institutional delivery	0.82	0.83	2248	0.91	0.93	4089		0.03	0.02	0.197
Received postnatal care	0.08	0.10	2248	0.39	0.43	4135		0.03	0.04	0.491
Family planning for women with birth in previous 2 years										
Use modern contraceptive	0.04	0.04	2903	0.05	0.09	4480		0.04**	0.02	0.022
Ever discussed FP with health providers	0.43	0.45	2889	0.75	0.80	4480		0.08**	0.04	0.031

Family planning for women without birth in previous 2 years										
Use modern contraceptive			0.05	0.07	5106	0.05		0.03**	0.01	0.021
Ever discussed FP with health providers			0.58	0.63	5106	0.58		0.06	0.04	0.125

Note: Data from household surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
[a] Sample of children of women who were pregnant in the two years preceding the survey.
[b] Sample of women who had a pregnancy in the two years preceding the survey
[c] ANC = Antenatal care
Significance level: * 10%; ** 5%; *** 1%

The PBF program positively impacted outcomes related to family planning. We separate the analysis of these outcomes by whether women had a pregnancy in the preceding two years because that was the only inclusion criterion for the baseline sample. For both groups there is an impact of 4-5 percentage points on the use of modern family planning methods, statistically significant at the 1% level. There is also an impact on the share of women reporting to have ever discussed family planning with a health provider among the recently pregnant women, but not among women without a recent pregnancy.

Impacts on quality of care

We now turn to describing results related to quality of care, measured with the detailed data collected during the health facility assessments. Given the very different mix of services delivered, the quality analysis was conducted separately for health centers and hospitals. It is important to bear in mind here that the packages of incentivized services and the weights of the quality bonuses are different for hospitals; hospitals generally also started the program with better baseline outcomes and generally have better trained staff.

Table 6 shows substantial improvements in structural quality, for both the treatment and control facilities, in the availability of equipment, consumable medical supplies and clinical and administrative protocols. Each index represents the share of available items out of a list of considered items (the full lists are presented in the appendix). The improvements between the baseline and endline are bigger in the health centers relative to the hospitals that started from a higher baseline level. Out of the five categories, PBF had statistically significant impacts at the health center level on basic equipment (4 ppt, p-value=0.04), family planning products (19 ppt, p-value < 0.01), and protocols (11 ppt, p-value=0.01). There were no impacts on the availability of these items at the hospitals. It is important to note that facilities in both study arms were, as of 2020, required to spend a minimum 20% of their payments on

drugs. This might explain the lack of difference between the arms with respect to essential drugs. As for vaccines, facilities do not procure them but rather receive the inputs from the national vaccination program. Moreover, during the implementation of the project, many facilities received refrigerators connected to solar panels through a GAVI-funded program. This can explain why there was a large increase in vaccine availability on the one hand, but no difference between the two treatment groups on the other hand.

Table 6: Impact on availability of equipment, medical supplies, and protocols

	Baseline			Endline				PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.		PBF Impact	Standard Error	p-value
Health centers										
Basic equipment (out of a list of 27 items)	0.31	0.31	260	0.54	0.57	290		0.04**	0.02	0.041
Essential drugs (out of a list of 22 items)	0.44	0.46	190	0.68	0.70	290		0.01	0.03	0.586
Vaccines (out of a list of 6 items)	0.05	0.04	207	0.62	0.66	275		0.05	0.06	0.450
Family planning products (out of a list of 8 items)	0.10	0.15	207	0.35	0.53	290		0.19***	0.04	0.000
Clinical and administrative protocols (17 items)	0.30	0.30	286	0.43	0.54	290		0.11**	0.04	0.010
Hospitals										
Basic equipment (out of a list of 27 items)	0.58	0.60	26	0.75	0.76	56		0.01	0.02	0.600
Essential drugs (out of a list of 22 items)	0.69	0.71	29	0.85	0.83	56		-0.04	0.03	0.238
Vaccines (out of a list of 6 items)	0.01	0.00	29	0.13	0.07	53		-0.06	0.07	0.424
Family planning products (out of a list of 8 items)	0.19	0.31	29	0.53	0.54	56		0.00	0.06	0.959
Clinical and administrative protocols (17 items)	0.57	0.61	29	0.66	0.76	56		0.08	0.05	0.121

Notes: Data from facility-based surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level. The list of items included in each index is presented in the appendix
Significance level: * 10%; ** 5%; *** 1%

Like the increases in the availability of medical supplies, there are also positive trends with respect to most Infection Prevention and Control (IPC) measures (**Table 7**). Nevertheless, there were no improvements between 2015 and 2021/22 in the share of facilities with secured boxes for sharps and needles and in the share of health centers reporting to use an improved source of water as their main source. At the health center level, there is a significant positive impact of PBF on the availability of an incinerator (10 ppt, p-value = 0.099), the presence of containers for sharps and needles (12 ppt, p-value =0.016), and on having sterilization procedures posted (23 ppt, p-value <0.01). At the hospital level, there are positive impacts on posting sterilization procedures (24 ppt, p-value = 0.05) and disinfection procedures (28 ppt, p-value = 0.03).

Table 7: Impacts on infection prevention and control

	Baseline			Endline				PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.		PBF Impact	Standard Error	p-value
Health centers										
Improved source of water	0.32	0.36	289	0.30	0.31	290		0.00	0.07	0.995
Handwashing stations in consultation rooms	0.20	0.24	290	0.59	0.53	264		-0.07	0.07	0.351
Incinerator	0.10	0.11	288	0.19	0.30	290		0.12*	0.06	0.061
Sterilization procedures posted	0.02	0.03	289	0.09	0.35	290		0.27***	0.06	0.000
Disinfection procedures posted	0.17	0.15	256	0.41	0.53	264		0.12	0.09	0.165
Containers for sharps and needles	0.83	0.75	288	0.71	0.82	290		0.12**	0.05	0.016
Hospitals										
Improved source of water	0.53	0.43	29	0.46	0.57	56		0.10	0.13	0.457
Handwashing stations in	0.67	0.93	29	0.81	0.86	55		0.04	0.10	0.733

consultation rooms										
Incinerator	0.67	0.71	29	0.85	0.73	56		-0.14	0.11	0.224
Sterilization procedures posted	0.33	0.14	29	0.54	0.77	56		0.24**	0.12	0.051
Disinfection procedures posted	0.36	0.36	28	0.54	0.83	55		0.28**	0.12	0.028
Containers for sharps and needles	0.93	0.93	29	0.88	0.97	56		0.09	0.07	0.223

Notes: Data from facility-based surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
Significance level: * 10%; ** 5%; *** 1%

In Tables

Table 8 and Table 9, we present results on outcomes related to the process quality, measured through direct clinical observations of child curative care visits, first antenatal consultations, family planning consultations, and deliveries.

Table 8 shows results for health centers and **Table 9** shows results for hospitals. Each index represents a share of completed actions by health providers, out of a list of activities that should take place in each consultation or delivery. Unlike the results presented in the other tables, we prefer the adjusted regressions for these outcomes because PBF had some impacts on the utilization of services. To address potential imbalances, we control for facility characteristics, provider characteristics and user/patient characteristics. Overall, the additions of these controls have appeared to only have a negligible effect on the results.

To measure the content of curative care for children under 5, we focus on IMCI diagnostic protocols that should be adhered to regardless of a child's symptoms or disease. The first index is an overall assessment score, which covers 29 items, including questions the providers ask the caregivers and examinations they conduct, in addition to whether they greet the caregiver. The second index captures providers efforts to verify the presence of four critical danger signs (inability to drink, vomiting, lethargic behavior, and convulsions) that are markers for severe and potentially fatal disease. The final index is constructed of eight essential checks for completion of an IMCI assessment, including taking

temperature, measuring respiratory rate, undressing the child to inspect upper body, auscultating chest, skinfold test, checking arms and eyes, checking feet and ankles for swelling, and checking ear and neck.

Table 8: Impacts on content of care measured through direct clinical observations in Health centers

	Baseline			Endline				PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.		PBF Impact	Standard Error	p-value
Child curative care										
Overall assessment score	0.52	0.49	220	0.56	0.61	1663		0.03	0.02	0.114
Share of danger signs verified	0.52	0.45	220	0.48	0.55	1663		0.05*	0.03	0.095
Share of essential assessments conducted	0.51	0.45	220	0.56	0.63	1663		0.05	0.04	0.228
First Antenatal consultations										
Overall consultation score	0.49	0.49	133	0.52	0.54	1486		0.02	0.03	0.425
Share of history taking	0.69	0.61	139	0.60	0.65	1493		0.04	0.04	0.230
Share of physical exams	0.71	0.82	139	0.74	0.78	1493		0.04	0.04	0.359
Shae of diagnostic tests	0.20	0.29	137	0.27	0.27	1493		0.02	0.05	0.613
Share of preventive care	0.39	0.41	140	0.79	0.77	1493		-0.01	0.05	0.777
Share of counseling topics	0.65	0.46	140	0.66	0.66	1486		0.00	0.05	0.937
Family planning consultations										
Overall consultation score				0.72	0.85	570		0.13***	0.03	0.000
Respectful care score				0.84	0.93	570		0.09***	0.03	0.006
Method selection score				0.70	0.80	575		0.09*	0.05	0.064
Effective use score				0.68	0.85	575		0.16***	0.05	0.002
Continuity of use score				0.64	0.82	575		0.20***	0.05	0.000
Deliveries										
Overall delivery score				0.60	0.64	259		0.03	0.04	0.387
Postpartum care score				0.07	0.10	278		0.04	0.05	0.386

Newborn care score				0.74	0.80	279		0.05	0.06	0.411

Notes: Data from direct clinical observations conducted as part of the facility-based surveys. The PBF impact coefficients are estimated with an adjusted regression, with controls for facility, provider, and patient/client characteristics, and the randomization blocks and clustering of standard errors at the health zone level. The list of items included in each index is presented in the appendix
Significance level: * 10%; ** 5%; *** 1%

At the health center level, we estimate an impact of 5 percentage points (p-value = 0.095) in the verification of danger signs (

Table 8). While the PBF health centers perform better also on the other two indexes, the differences are not statistically significant. The estimated impact on the verification of danger signs in the hospitals is even bigger (8 ppt) but is not statistically significant given the smaller sample size of consultations observed at hospitals. However, we estimate a significant impact on the overall assessment score at the hospitals (5 ppt, p-value = 0.09).

Table 9: Impacts on content of care measured through direct clinical observations in Hospitals

	Baseline			Endline				PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.		PBF Impact	Standard Error	p-value
Child curative care										
Overall assessment score	0.63	0.60	71	0.65	0.69	366		0.05*	0.03	0.090
Share of danger signs verified	0.59	0.67	71	0.59	0.67	366		0.08	0.06	0.148
Share of essential assessments conducted	0.69	0.64	71	0.70	0.76	366		0.06	0.04	0.177
First Antenatal consultations										
Overall consultation score	0.65	0.52	38	0.57	0.63	284		0.06*	0.03	0.069
Share of history taking	0.83	0.59	39	0.62	0.73	284		0.12**	0.05	0.013
Share of physical exams	0.86	0.82	40	0.78	0.88	284		0.09**	0.04	0.025
Shae of diagnostic tests	0.53	0.53	38	0.47	0.51	284		0.03	0.06	0.635
Share of preventive care	0.67	0.75	40	0.78	0.80	284		0.02	0.07	0.736
Share of counseling topics	0.78	0.37	39	0.73	0.75	284		0.02	0.08	0.854
Family planning consultations										
Overall consultation score				0.79	0.89	159		0.07	0.05	0.173
Respectful care score				0.82	0.93	159		0.07	0.05	0.200
Method selection score				0.82	0.85	159		-0.01	0.05	0.799
Effective use score				0.78	0.90	159		0.10	0.07	0.169

Continuity of use score				0.72	0.88	159		0.14**	0.06	0.028
Deliveries										
Overall delivery score				0.67	0.74	135		0.07*	0.04	0.078
Postpartum care score				0.02	0.17	141		0.16**	0.07	0.031
Newborn care score				0.76	0.87	143		0.11*	0.06	0.070

Notes: Data from direct clinical observations conducted as part of the facility-based surveys. The PBF impact coefficients are estimated with an adjusted regression, with controls for facility, provider, and patient/client characteristics, and the randomization blocks and clustering of standard errors at the health zone level. The list of items included in each index is presented in the appendix
Significance level: * 10%; ** 5%; *** 1%

We restrict our analysis of the quality of antenatal care to the first consultations women receive during their pregnancies because the content of consultations does not depend on the content in previous ones. We use six indexes capturing medical history taking, physical exams, diagnostic testing, preventive care, and counselling. The overall consultation score includes the items in all of the other five indexes. During the endline survey, providers were observed to perform 53% of the overall protocol in the health centers and 60% of the protocol in hospitals. Between the survey rounds, providers' performance on the overall consultation score increased by 4 percentage points at the health centers and 2 percentage points at the hospitals. The biggest increases of 38 percentage points in the health centers and 8 percentage points in the hospitals was found for preventive care as providers are substantially more likely to give anti-tetanus shots, iron supplementation and malaria chemoprophylaxis. There is also a substantial improvement in inclusion of counselling about nutrition, birth preparation and danger signs. The overall index increased by 12 percentage points in health centers and 18 percentage points in hospitals between the survey rounds.

In terms of impacts, there is no statistically significant difference at the health center level for any of the ANC quality indexes. At the hospital level, impacts on three out of the six indexes are significant. The overall consultation score is higher by six percentage points in the PBF hospitals (p-value = 0.069), the share of history taking is higher by 12 percentage points (p-value = 0.01), and the share of physical exams if higher by 9 percentage points (p-value = 0.03).

The indexes we use to measure quality of family planning consultations are based of Jain et al. (2019). Respectful care measures whether providers address the clients with respect, respect their privacy, encourage question, and provide explanations. Quality of method selection depends on the questions providers ask the clients before recommending a method and on not showing a bias for a specific method. The consultation's quality with respect to effective use measures whether providers explain how the method works and how to use it, and discuss potential side effects and their management. In term of continuity of care, providers are expected to discuss when and where the method needs to be renewed, suggest a follow up consultation, and mention the possibility of

changing methods in the future. The overall consultation score includes the 17 items from all of the other indexes. Family consultations were not observed during the baseline survey, and we can therefore not address the evolution in quality. The average overall score during the endline survey was 80 percent in the health centers and 85 percent in the hospitals. At the health center level, there are statistically significant impacts of the PBF program on all the FP consultation quality indexes. In comparison to the measures at the control group, the PBF program increased the overall consultation score by 12 percentage points (p-value<0.01), the respectful care score by 8 percentage points (p-value = 0.02), the method selection score by 9 percentage points (p-value = 0.09), the effective use score by 16 percentage points (p-value<0.01), and the continuity of use score by 20 percentage points (p-value<0.01). The impact on quality of family planning consultations is more moderate at the hospitals. We only estimate a statistically significant impact on the continuity of use score of 14 percentage points (p-value = 0.03). We estimate coefficients in the range of 7-10 percentage points for three of the other indexes but given the sample size of consultations observed in hospitals, we can only detect larger impacts.

We use three different indexes to measure quality of labor and delivery services. The 'overall delivery score' is a comprehensive index based on Tripathi et al. (2015; 2019). It includes 21 items in the different ranging from preparation of supplies, assessment of women's medical histories and symptoms, use of partograph, to postpartum and newborn care. The postpartum care score is based on the WHO guide for immediate postpartum monitoring (WHO 2015). The newborn care score includes the four elements of WHO essential newborn care (immediate drying of baby, skin-to skin, tying or clamping cord, and initiation of breastfeeding; WHO 2017) plus additional items included in the WHO guide for essential practice for pregnancy, childbirth, postpartum and newborn care (WHO 2015). The average overall delivery score is 62 percent at the health centers and 71 percent at the hospitals. The postpartum care score is only around 10 percent in both type of facilities. The performance on this score low because the guidelines require providers to check on women four times during the first hour following the birth and an additional time in the following hour. While provider conduct some of the checks, they do not meet the required number and frequency of checks. With respect to the newborn care score, the average score is 78 percent at the health centers and 82 percent at the hospitals.

At the health centers, we do not detect significant impacts of PBF on quality of delivery services as there are only modest improvements in the PBF group relative to the control. At the hospitals, we find significant impacts on the three indexes. PBF hospitals had 7 percentage points higher overall delivery score (p-value = 0.08), 16 percentage points higher postpartum care score (p-value = 0.03) and by 11 percentage points (p-value = 0.07) higher newborn care score. It is important to note that observations of labor and delivery services were only conducted in the provinces of Kwango, Kwilu and Mai-Ndombe and in only a subset of health centers. Further, the number of observed deliveries was limited by low volume in many facilities, which limits statistical power with respect to these outcomes.

In **Table 10**, we present results on outcomes related to the delivery of maternal and reproductive health services. In both type of facilities, there has been an increase in the number of days during which antenatal care is provided but there is no statistically significant difference between the treatment arms. During the baseline survey, only 41 percent of health center and 72 percent of hospitals provided family planning services. These rates have increased to 80 and 95 percent. Health centers in the PBF zones are more likely by 23 percentage points to offer such services (p-value<0.01). PBF centers are also more likely to include family planning counselling as part or antenatal and postpartum care, and to have models to use as visual aids. PBF hospitals were significantly less likely to charge for family planning consultations.

Table 10: Impacts on delivery of reproductive and maternal health services, reports by facility-in-charges

	Baseline			Endline				PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.		PBF Impact	Standard Error	p-value
Health centers										
Days offering ANC services per week	1.31	1.23	283	1.77	2.31	288		0.42	0.28	0.143
ANC: cover FP			0	0.82	0.91	288		0.09**	0.04	0.021
Discuss FP during postpartum care			0	0.83	0.94	283		0.11***	0.04	0.005
Offer family planning service	0.39	0.42	288	0.68	0.91	290		0.24***	0.06	0.000
Charge for FP consultations			0	0.26	0.22	231		-0.06	0.05	0.271
Charge for FP products			0	0.23	0.22	231		-0.06	0.05	0.243
Have models for FP consultation			0	0.39	0.66	231		0.29***	0.09	0.001
Hospitals										
Days offering ANC services per week	1.80	1.62	28	2.58	3.07	56		0.38	0.64	0.550
ANC: cover FP			0	1.00	1.00	56				
Discuss FP during postpartum care			0	0.92	1.00	56		0.07	0.05	0.159
Offer family planning service	0.53	0.93	29	0.96	0.93	56		-0.02	0.06	0.740
Charge for FP consultations			0	0.44	0.18	53		-0.28**	0.12	0.020
Charge for FP products			0	0.28	0.25	53		-0.06	0.12	0.628
Have models for FP consultation			0	0.68	0.82	53		0.16	0.13	0.200

Notes: Data from facility-based surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
Significance level: * 10%; ** 5%; *** 1%

Impacts on health providers' work effort, motivation and satisfaction

Table 11 reports outcomes reported by health providers. During the endline survey, health providers reported to have been absent during 1.2 days on average in the preceding 30 days, to work an average of 5.6 days a week, and an average of 9.3 hours per day. When comparing baseline and endline measures, health providers report higher average number of absent days at the health centers but not in the hospitals. There is also an average decrease of about one hour in the number of hours worked per day. We do not find that the PBF program affected absenteeism and the amount of time worked, at least as measured through these self-reports. We also do not find evidence that the program impacted the satisfaction or motivation of staff. Staff motivation and satisfaction is similar across the facility types. It appears that there has been an improvement in staff motivation between 2015 and 2021/22 in both treatment arms.

Table 11: Impacts on health providers attendance, motivation, and satisfaction

	Baseline			Endline			PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.	PBF Impact	Standard Error	p-value
Health centers									
Days absent during last 30 days	0.86	0.69	562	1.20	1.15	1182	-0.05	0.29	0.874
Days worked in past 7 days	5.95	5.99	564	5.59	5.61	1182	-0.01	0.17	0.972
Hours worked in past 7 days	10.12	10.41	564	9.41	9.14	1182	0.08	0.46	0.859
General motivation (0-10 scale)	6.60	6.86	558	7.12	7.40	1182	0.17	0.30	0.570
Satisfaction index				0.55	0.57	1186	0.01	0.02	0.650
Hospitals									
Days absent during last 30 days	0.92	0.38	115	0.38	0.93	279	0.65	0.39	0.105
Days worked in past 7 days	6.12	6.04	115	6.00	5.85	279	-0.07	0.17	0.678
Hours worked in past 7 days	9.68	10.11	115	8.55	9.08	279	0.54	0.53	0.313
General motivation (0-10 scale)	5.95	6.39	114	6.94	7.26	279	0.24	0.41	0.561
Satisfaction index				0.53	0.58	279	0.04	0.03	0.218

Notes: Data from provider interviews conducted as part of the facility-based surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level. The list of items included in the satisfaction index is presented in the appendix
Significance level: * 10%; ** 5%; *** 1%

Impacts on user fees and health expenditure

We next turn to consider the impact of the PBF program on user fees and health expenditure by households. We first show outcomes reported by the health facilities in **Table 12**. Health facilities in the PBF zones are more likely to charge flat user fees, and accordingly are less likely to charge separate fees for laboratory tests, inpatient care, consumables and drugs. These impacts are stronger at the health center level relative to the hospitals. Both types of facilities are significantly more likely to post service fees in a public space in the facility. Health centers are also significantly more likely to report giving fee exemptions to poor patients. Finally, health centers in the PBF health zones post significantly lower fees for delivery and antenatal care services. However, they also post higher fees for curative care.

Table 12: Impacts on user fees

	Baseline			Endline			PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.	PBF Impact	Standard Error	p-value
Health centers									
Charge flat fees	0.51	0.49	289	0.51	0.75	290	0.21***	0.07	0.003
Charge for lab tests	0.84	0.88	185	0.62	0.24	290	-0.37***	0.05	0.000
Charge for inpatient care	0.41	0.41	289	0.53	0.23	290	-0.28***	0.06	0.000
Charge for consumables	0.68	0.65	289	0.68	0.33	290	-0.33***	0.07	0.000
Charge for drugs	0.81	0.80	289	0.86	0.45	290	-0.39***	0.07	0.000
Fees posted	0.60	0.57	289	0.63	0.81	290	0.15**	0.06	0.011
Fee exemptions for the poor	0.99	1.00	290	0.67	0.87	287	0.18***	0.05	0.000
FP consultation fee			0	573.92	452.21	228	-200	164	0.229
Delivery fee	8142	8207	279	8734	6718	288	-1914**	882	0.034
ANC fee	2586	2166	283	2493	1721	289	-735***	273	0.009
Curative care fee	2948	2580	273	3120	3837	244	596*	336	0.081
Hospitals									
Charge flat fees	0.20	0.14	29	0.58	0.73	56	0.14	0.13	0.307
Charge for lab tests	1.00	1.00	29	0.88	0.57	56	-0.32***	0.10	0.003
Charge for inpatient care	0.93	1.00	29	0.85	0.70	56	-0.15	0.11	0.203
Charge for consumables	0.73	0.79	29	0.65	0.33	56	-0.35**	0.13	0.010
Charge for drugs	1.00	1.00	29	0.92	0.73	56	-0.19**	0.09	0.040
Fees posted	0.53	0.79	29	0.69	0.97	56	0.29***	0.10	0.005

Fee exemptions for the poor	1.00	1.00	29	0.77	0.83	56		0.05	0.11	0.608
FP consultation fee			0	1000	714	52		-545	462	0.244
Delivery fee	12200	10150	29	11788	9917	56		-1959	1899	0.307
ANC fee	2913	2115	28	2540	2605	54		425	911	0.642
Curative care fee	2922	1083	15	4438	3563	16		-567	1660	0.737

Notes: Data from facility-based surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
Significance level: * 10%; ** 5%; *** 1%

In **Table 13,** we present results on health spending from the users' perspective. In the first two panel, we present reports on fees paid during the exit interviews collected at the health centers and hospitals. These measures suggest more moderate differences in user fees between the two groups. There is only a significant difference with respect to out-of-pocket expenditure on child curative care consultations at the health center level (-993 CDF, p-value=0.03), representing a reduction of 20% relative to the fees paid in the control group. The bottom panel of the table presents results from the household survey. Thirty-seven percent of the household reported to have had health costs in the preceding 12 months that exceeded the household financial means. Household in the PBF group were less likely to report such health spending shocks by 4 percentage points, but the difference is not significant. However, household s in PBF zones were less likely to have had to borrow money or to depend on gifts for covering health-related costs.

Table 13: Impacts on health spending

	Baseline			**Endline**				**PBF Impact**		
	Control Mean	**PBF Mean**	**Obs.**	**Control Mean**	**PBF Mean**	**Obs.**		**PBF Impact**	**Standard Error**	**p-value**
Paid fees reported during exit interviews at health centers										
Child curative care	5765	5479	220	4997	4032	1627		-993**	445	0.030
Antenatal care	3766	2476	129	1976	1909	1489		152	404	0.709
Family planning consultations			704	611	577	704		-28	287	0.922
Paid fees reported during exit interviews at hospitals										
Child curative care	8898	10180	71	13917	12486	364		-1112	1574	0.483
Antenatal care			3221	4199	283	3221		1099	1018	0.285
Family planning consultations			1980	1396	159	1980		-636	850	0.458
Reports on household health spending in preceding 12 months, from household survey										
Health spending exceeded household's financial means			0.37	0.34	8510	0.37		-0.04	0.05	0.412

Household had to sell goods/assets to cover health spending			0.34	0.33	8510	0.34		-0.02	0.04	0.705
Household had to borrow money to cover health spending			0.18	0.15	8510	0.18		-0.04*	0.02	0.089
Household relied on a gift to cover health spending			0.05	0.03	8510	0.05		-0.02**	0.01	0.020

Notes: Data from exit interviews conducted as part of the facility-based surveys and household surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
Significance level: * 10%; ** 5%; *** 1%

Impacts on users' satisfaction

We finally test whether the PBF program affected users' satisfaction with the services received at the health facilities. Satisfaction was collected during the exit interviews conduction after observed consultations for child curative care, antenatal care, and family planning. We consider three satisfaction indicators: with the quality of the received service, with the affordability of the service and an overall satisfaction index composed of 9 categories such as cleanliness, attitude of providers and waiting time (full list in the appendix). As can be seen in **Table 14**, almost all users express satisfaction with the quality of the service they had just received before being interviewed in both survey rounds. Between the baseline and the endline surveys, there has been an increase in the overall satisfaction score and with respect to the affordability of services. When comparing the two treatment groups, there are no statistically significant differences in the satisfaction measures except for the affordability of antenatal care. Women interviewed in the PBF health centers are more likely, by 9 percentage points, to report being satisfied with the affordability of the consultations.

Perceptions about facilities were also collected in the household survey. Heads of households (or most informed household members who responded to the main household modules) were asked about different aspects related to the nearest health facility. The respondents were defined to be satisfies with the different categories if they reported them to be either good or very good. As can be seen in the lowest panel of **Table 14**, measuring satisfaction at the households rather than at the facilities resulted in substantially lower satisfaction. Satisfaction in the PBF health zones is slightly higher for all categories

but none of the differences are estimated to be statistically significant. However, there has been an overall increase in satisfaction between the survey rounds, particularly with respect to costs.

Table 14: Impacts on users' satisfaction

	Baseline			Endline			PBF Impact		
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.	PBF Impact	Standard Error	p-value
Exit interviews after child curative care at health centers									
Satisfied with quality	0.95	0.92	218	0.92	0.92	1627	0.00	0.02	0.893
Satisfied with affordability	0.62	0.68	217	0.87	0.90	1556	0.02	0.03	0.419
Overall satisfaction index	0.85	0.80	199	0.89	0.88	1538	0.00	0.02	0.980
Exit interviews after antenatal care at health centers									
Satisfied with quality	0.95	0.87	129	0.95	0.90	1480	-0.04	0.03	0.144
Satisfied with affordability	0.78	0.71	125	0.73	0.82	1266	0.09**	0.04	0.048
Overall satisfaction index	0.88	0.80	121	0.89	0.90	1256	0.01	0.02	0.434
Exit interviews after family planning consultations at health centers									
Satisfied with quality			0.93	0.94	573	0.93	0.00	0.04	0.962
Satisfied with affordability			0.67	0.78	245	0.67	0.15	0.09	0.113
Overall satisfaction index			0.91	0.92	556	0.91	0.01	0.02	0.676
Reports by household heads' on nearest health facility									
Overall satisfaction index	0.52	0.53	2733	0.57	0.60	8343	0.03	0.02	0.268
Satisfied with equipment	0.40	0.40	2860	0.44	0.49	8453	0.05	0.04	0.265
Satisfied with drug availability	0.37	0.39	2877	0.39	0.44	8480	0.05	0.04	0.262
Satisfied with staff respectfulness	0.67	0.68	2809	0.68	0.69	8499	0.03	0.02	0.273
Satisfied with staff competency	0.67	0.70	2876	0.70	0.73	8448	0.04	0.03	0.188
Satisfied with cost of services	0.14	0.15	2882	0.33	0.37	8491	0.03	0.05	0.558
Satisfied with cost of drugs	0.13	0.13	2880	0.25	0.29	8474	0.03	0.04	0.440

Notes: Data from exit interviews conducted as part of the facility-based surveys and household surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
Significance level: * 10%; ** 5%; *** 1%

Heterogeneous impacts

The overall estimations might mask some important heterogeneities in program impacts. We test for two types of heterogeneities. First, we consider heterogeneity by region. DRC is a very diverse and decentralized country. The regions in DRC differ by ecosystems, economic activity, remoteness, ethnic group composition. In terms of local administration of the health system, each province has its own ministry of health. Ideally, we could test for impacts at each province, but we don't have sufficient

number of clusters in each province for a meaningful analysis. We therefore group Kwango, Kwilu and Mai-Ndombe in the Southwest of the country, that are all part of the former province of Bandundu until the decentralization reform in 2015. The other group corresponds to the former Katanga province in the Southeast of the country, which now comprises Haut-Katanga, Haut-Lomami, and Lualaba. The PBF in each of these groups of provinces was implemented by a single provincial purchasing agency, and therefore it is reasonable to assume that implementation of the program within each group was relatively similar. The second heterogeneity we consider is by household wealth. Using standard asset indices, we test whether impacts on utilization differed for women below and above the household wealth median. For brevity, we only present results on a subset of the outcomes.

Table 15:Heterogeneous impacts on coverage of reproductive and maternal health services by region

	Endline Samples' means					Impact in old Katanga Province				Impact in old Bandundu province		
	Old Katanga Provine	Obs.	Old Bandundu Province	Obs.		PBF Impact	Standard Error	p-value		PBF Impact	Standard Error	p-value
Women with pregnancies in past 2 years [a]												
Any ANC [b]	0.89	930	0.94	3205		-0.02	0.05	0.734		0.03**	0.01	0.030
4+ antenatal consultations	0.43	930	0.37	3204		0.03	0.07	0.683		0.04	0.06	0.495
ANC in first trimester	0.22	930	0.23	3205		0.15***	0.05	0.003		0.07**	0.03	0.015
ANC with tetanus shot	0.73	930	0.82	3205		0.06	0.06	0.330		0.03	0.03	0.358
ANC with anti-malarial	0.61	930	0.75	3205		0.03	0.06	0.648		0.01	0.05	0.904
Institutional delivery	0.88	914	0.93	3175		0.01	0.04	0.756		0.03	0.03	0.203
Received postnatal care	0.49	930	0.40	3205		0.06	0.09	0.520		0.02	0.05	0.659
All Women												
Use modern contraceptive	0.08	2095	0.07	7491		-0.01	0.02	0.515		0.05***	0.01	0.002
Ever discussed FP with health providers	0.83	2095	0.68	7491		-0.01	0.03	0.674		0.09**	0.04	0.044

Notes: Data from household surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
[a] Sample of women who had a pregnancy in the two years preceding the survey
[b] ANC = Antenatal care
Significance level: * 10%; ** 5%; *** 1%

Table 16: Heterogeneous impacts on health center outcomes by region

	Endline Samples' means		Impact in old Katanga Province		Impact in old Bandundu province

	Old Katanga Provine	Obs.	Old Bandundu Province	Obs.		PBF Impact	Standard Error	p-value		PBF Impact	Standard Error	p-value
Outcomes from health center assessments												
Basic equipment	0.56	82	0.57	300		0.05**	0.02	0.023		0.07***	0.03	0.009
Family planning products	0.63	82	0.37	300		0.06	0.05	0.250		0.21***	0.05	0.000
Improved source of water	0.22	82	0.38	300		0.13***	0.04	0.004		0.01	0.08	0.923
Incinerator	0.37	82	0.31	300		-0.09	0.09	0.307		0.16**	0.07	0.036
Containers for sharps and needles	0.78	82	0.76	300		0.07	0.11	0.566		0.18***	0.05	0.002
Flat fees	0.74	82	0.58	300		0.08	0.06	0.191		0.26***	0.08	0.002
Fee exemptions for the poor	0.85	81	0.69	298		0.06	0.05	0.300		0.24***	0.08	0.004
Child curative care assessment score	0.51	431	0.61	1225		0.07**	0.04	0.049		0.02	0.03	0.439
First antenatal consultation score	0.47	424	0.56	1346		-0.10***	0.03	0.004		0.07**	0.03	0.028

Notes: Data from facility-based surveys. The PBF impact coefficients are estimated with an adjusted regression, with controls for facility, provider, and patient/client characteristics, and the randomization blocks and clustering of standard errors at the health zone level. Regressions on child curative care assessment score and antenatal consultation score include controls for facility, provider and patient characteristics. The list of items included in each index is presented in the appendix
Significance level: * 10%; ** 5%; *** 1%

As can be seen in **Table 15**, there are differences in program impacts on utilization by region. Positive impact on timely antenatal care were found in both regions, with a larger magnitude in the eastern provinces (Katanga). On the other hand, the impacts on family planning are concentrated in the western provinces (Bandundu). In addition, when restricting the analysis to Kwango, Kwilu and Mai-Ndombe, there is also a significant impact on receiving any antenatal care during the pregnancy. **Table 16** shows differences in outcomes measured during the health center assessments. In both regions the availability of basic equipment increased, but the impact on availability of family planning products is only significant in the old Bandundu region where overall availability of such products is substantially lower. The PBF impact on fees is also stronger in Kwango, Kwilu and Mai-Ndombe, but the facilities in the other provinces are more likely to charge flat fees and except fees for the poor regardless of the PBF

program. Finally, we find diverging results on process quality. In the eastern provinces that were part of the old Katanga province, we estimate a positive impact on the overall assessment score of child curative care but a *negative* impact on the quality of first antenatal consultations. In the other group of provinces, we estimate a positive but non-significant impact on the quality of child curative consultations and a *positive* impact on the quality of antenatal care.

Table 17 shows results on utilization of reproductive and maternal health services by household wealth. The impact on timely initiation of antenatal care is detected for women below and above the wealth median. For women below the household wealth median, we also estimate an impact on institutional delivery. We also find a stronger PBF impact on the poorer group for ever speaking with a health provider about family planning. However, the improvement in the use of modern contraceptives is concentrated among the wealthier group.

Table 17: Heterogeneous impacts on coverage of reproductive and maternal health services by household wealth

	Endline Samples' means				Impact on women below wealth median			Impact on women above wealth median		
	Below wealth median	Obs.	Above wealth median	Obs.	PBF Impact	Standard Error	p-value	PBF Impact	Standard Error	p-value
Women with pregnancy in past 2 years										
Any ANC	0.91	2039	0.95	2096	0.03	0.02	0.151	0.01	0.01	0.685
4+ antenatal consultations	0.35	2038	0.40	2096	0.01	0.05	0.795	0.06	0.05	0.263
ANC in first trimester	0.22	2039	0.24	2096	0.08***	0.03	0.009	0.08***	0.03	0.004
ANC with tetanus shot	0.76	2039	0.84	2096	0.05	0.04	0.162	0.02	0.03	0.554
ANC with anti-malarial	0.68	2039	0.75	2096	-0.02	0.05	0.674	0.04	0.04	0.335
Institutional delivery	0.89	2024	0.95	2065	0.06*	0.03	0.069	0.00	0.01	0.819
Received postnatal care	0.37	2039	0.47	2096	0.02	0.05	0.714	0.04	0.04	0.396
All Women										
Use modern contraceptive	0.05	4574	0.09	5012	0.02	0.01	0.115	0.04***	0.02	0.008
Ever discussed FP with health providers	0.69	4574	0.73	5012	0.09**	0.04	0.030	0.04	0.03	0.211

Notes: Data from household surveys. The PBF impact coefficients are estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.
Significance level: * 10%; ** 5%; *** 1%

Shorter vs. longer-term impacts

A policy relevant question is how long it takes for the PBF impacts to materialize. We can partially address this question by comparing results from the midline survey to those from the endline survey. The midline survey was collected only in health centers in Kwago, Kwilu and Mai-Ndombe in the third quarter of 2018, about two years after PBF was launched in these provinces. As can be seen in **Table 18**, there were already some impacts on structural quality and service fees after two year, but there were further improvements between the midline and endline surveys. For example, availability of family products in the PBF health centers during the midline survey was 23 percentage points higher relative to the control facility during the midline. Availability in both treatment arms improved between 2018 and 2021, but the gap in availability between the two groups was maintained. The reductions in delivery and antenatal care fees also materialized by 2018. For some indicators, impacts are only detected in the longer-term. PBF impacts on the availability of incinerators and safety boxes for sharps are only statistically significant in the endline survey. The impacts on the likelihood of publicly posting service fees and on offering fee exemptions for the poor have also only materialized after 2018.

Table 18: Comparing midline and endline results at the health centers

	Midline Survey			Endline Survey			Midline impact		Endline impact	
	Control Mean	PBF Mean	Obs.	Control Mean	PBF Mean	Obs.	PBF Impact	Standard error	PBF Impact	Standard error
Availability of equipment, medical supplies and protocols										
Basic equipment	0.52	0.53	150	0.55	0.58	220	0.02	0.02	0.03	0.02
Essential drugs	0.64	0.64	145	0.68	0.70	220	0.01	0.04	0.01	0.03
Vaccines	0.55	0.54	180	0.64	0.67	205	-0.02	0.08	0.03	0.07
Family planning products	0.14	0.37	162	0.27	0.50	220	0.23***	0.04	0.23***	0.05
Protocols	0.40	0.51	205	0.42	0.51	220	0.12**	0.05	0.09*	0.05
Infection prevention and control										
Improved source of water	0.41	0.33	208	0.35	0.33	220	-0.08	0.08	-0.03	0.09
Incinerator	0.16	0.21	209	0.14	0.32	220	0.06	0.06	0.18**	0.07
Sterilization procedures posted	0.10	0.30	209	0.08	0.30	220	0.20***	0.07	0.23***	0.08
Disinfection procedures posted	0.35	0.47	192	0.41	0.50	202	0.11	0.09	0.09	0.10

Box for sharps and needles	0.79	0.81	207	0.71	0.84	220		0.02	0.08		0.13***	0.05
Delivery of Reproductive and maternal health services												
Days offering ANC per week	1.79	1.93	208	1.98	2.52	218		0.14	0.34		0.44	0.36
Offer family planning	0.52	0.73	209	0.62	0.88	220		0.23***	0.07		0.28***	0.07
Fees												
Charge flat fees	0.66	0.87	208	0.46	0.73	0.66		0.21***	0.08		0.25***	0.09
Charge for lab tests	0.64	0.30	165	0.68	0.24	0.64		-0.36***	0.11		-0.43***	0.06
Charge for inpatient care	0.36	0.14	208	0.49	0.20	0.36		-0.25***	0.06		-0.30***	0.07
Charge for consumables	0.51	0.14	208	0.64	0.33	0.51		-0.39***	0.08		-0.31***	0.08
Charge for drugs	0.68	0.36	208	0.85	0.46	0.68		-0.33***	0.10		-0.38***	0.08
Fees posted	0.61	0.73	207	0.63	0.78	0.61		0.13	0.08		0.15**	0.07
Exemptions for the poor	0.65	0.68	208	0.63	0.86	0.65		0.02	0.08		0.23***	0.06
Delivery fee	7149	5148	207	8687	6772	218		-1940***	584		-1945**	730
ANC fee	2211	1482	207	2738	1985	219		-655**	275		-674**	332
Curative care fee	2913	3193	203	3888	4339	178		347	532		423	442

Notes: Data from facility-based surveys. The PBF impact coefficients are estimated with an adjusted regression, with controls for facility, provider, and patient/client characteristics, and the randomization blocks and clustering of standard errors at the health zone level. Regressions on child curative care assessment score and antenatal consultation score include controls for facility, provider and patient characteristics. The list of items included in each index is presented in the appendix

Significance level: * 10%; ** 5%; *** 1%

Chapter 4: Discussion

General trends

The data collected at baseline and endline of this study suggests generally very positive trends for the health zones supported by the PDSS project overall. The coverage of most of the considered reproductive, maternal and child health services increased substantially between 2015 and 2021/22. Direct clinical observations at the health facilities also suggest substantially greater adherence to protocols of antenatal care and diagnostic of child illnesses over time. Improvements can also be seen with respect to other dimensions of facilities' quality such as the availability of medical supplies and measures of infection prevention and control. While other factors could have affected these trends, it is reasonable to assume that the project contributed to these positive trends at least to some extent. Studies from different settings have demonstrated that provision of financing to health facilities can improve outcomes regardless of whether the financing is conditional on performance (de Walque et al 2022).

Although the overall trends are positive, the data also show that many gaps remain, and that there are a few areas where relatively little progress has been achieved. For example, coverage rates of growth monitoring and use of modern family planning methods remain in the single digits. Furthermore, while more than 90 percent of the women in our endline sample report to have received any antenatal care, and although the improvements in the content of care, less than a quarter initiate care in the first trimester of the pregnancy and many still do not receive basic antenatal care components such as iron supplementation. Despite some improvements in infection and prevention control measures at the facility level, many health facilities use unimproved sources of water and do not have handwashing stations in consultation areas. We hope that the detailed data collected for the purpose of this evaluation will be used to inform future interventions to address these gaps.

Summary of PBF impacts

Overall, PBF had moderate positive impacts on health service coverage and clearly improved quality in some areas. We did not see any improvements in child health service coverage relative to the control group receiving matching unconditional facility financing. With respect to maternal health services, coverage in the PBF group was higher for all the indicators, but these differences were only statistically significant for initiation of antenatal care in the first trimester.

PBF results were most positive for family planning, with significant increases in the use of modern contraceptives as well as the likelihood of women discussing family planning with health providers.

The PBF program improved several measures of quality of care. We found positive impacts for the availability of medical supplies, for measures of infection prevention and control, and for provider adherence to clinical protocols. However, we find different effects at the health center and hospital levels. At the health centers. We found significant improvements in the availability of equipment, family planning products and protocols at health centers. These were not seen at hospitals, where general availability of such supplies was much higher. We also find larger impacts on infection prevention and control at the health centers, such as on the likelihood of having functioning incinerators and secure boxes for needles and sharps. On the other hand, we detect more pronounced impacts on the process quality, or adherence to protocols more specifically, at the hospital level, where quality of all types of observed services are higher in the PBF group than in the control group. The only service for which we detect stronger process quality impacts at the health center level is family planning consultations. While these findings suggest that the PBF program impacted differently health centers and hospitals, two caveats must be mentioned. First, the sample consists of only 56 hospitals which limits the statistical power. In some cases, we estimate positive but imprecise coefficients for impacts at the hospital level. In addition, a larger share of hospitals is participating in the pay-for-knowledge pilot which enhanced the PBF quality score focus on process quality. This might explain, to some extent, the stronger impacts on adherence to clinical protocols at the hospitals.

There is some evidence that PBF changed user fees and out-of-pocket expenditures. PBF facilities are more likely to charge flat fees, exempt fees for poor patients, visibly post fee schedule, and have lower official fees. Such changes could have important implications on financial access to health services in a context with high reliance on user fees. At the user's level, results are a bit weaker – reductions in out-of-pocket expenditures measured during exit interviews were mostly detected for child curative care. However, households in PBF zones are less likely to borrow money or rely on support from other to cover health-related costs.

We find meaningful heterogeneity in impacts by region. The results in Kwango, Kwilu and Mai-Ndombe provinces (that formed the Bandundu province until 2015) are better for most outcomes

relative to the results in the provinces of Haut-Katanga, Haut-Lomami and Lualaba (that were part of the Katanga province until 2015). These differences highlight that the context in which PBF is introduced and the way in which it is implemented are important. This study cannot identify the exact factors that caused these differences. Process evaluation and qualitative approaches that explore these factors might uncover important lessons for implementation of a more impactful PBF approach in DRC.

A previous evaluation of a PBF pilot in DRC, also funded by the World Bank, concluded that, relative to unconditional payments to facility, the PBF approach *reduced* utilization of services and providers' satisfaction (Huillery and Seban 2021). We do not find such results here. Instead, we find a moderate *increase* in service utilization and no impact on providers' satisfaction and motivation. In addition, we estimate some positive impacts on quality of health services. The differences in the results highlight the importance of the details of the PBF design in addition to the context in which it is implemented and the fidelity of implementation. The two PBF models are different in several important ways such as the volume of performance payments and budget per person, whether quality was incentivized, the contracting of health zone and province teams, and the introduction of management tools.

Comparison to results from other countries

The results from this impact evaluation can also be compared to a larger set of evaluations. However, it is important to first highlight some differences in the context and interventions. The differences in results between the two studies carried in DRC show the risk in comparing and grouping interventions that might differ along important dimensions. Out of the countries where impact evaluations of PBF have been implemented, DRC stands out in terms of the strong reliance of health facilities on user fees to cover operational costs and remunerate their staff. This reliance on operational income means that facilities in DRC have always enjoyed broad autonomy relative to facilities in other countries and that their revenues were already conditional on the quantity of services they provided. While in other countries PBF greatly increased facility autonomy, the program in DRC may have actually restricted autonomy to some extent by introducing stricter guidelines on how facilities can charge users, spend their revenue, and remunerate health providers.

A recent Policy Research Report re-analyzes the experiences of five countries - Rwanda, Zambia, Zimbabwe, Nigeria, and Cameroon – all of which compared PBF to unconditional facility financing (de Walque et al. 2022). The report classifies the unconditional transfers as Decentralized Facility Financing (DFF). All DFF interventions provided financing not tied to performance but they differed by whether they included enhanced supervision, by the disbursement amounts to DFF facilities relative to PBF facilities, and by how the additional financing could be spent by the facilities. In DRC, control facilities received the average quarterly transfer by the PBF facilities, adjusted for catchment area population and remoteness. Health zone and province teams were not given any guidelines from the project with respect to how to conduct supervision in these zones.

Like the findings from the other studies, the PBF program in DRC significantly increased coverage of only few of the incentivized service relative to the matched income transfers. The types of impacted services in DRC, however, are different. In three countries, Rwanda, Zimbabwe, and Nigeria, PBF positively impacted institutional deliveries (Basinga et al 2011; Friedman et al. 2016A; Khanna et al. 2021). The lack of impact on institutional deliveries might be explained by the fact that the baseline coverage in DRC (81%) was substantially higher than the coverage in the other countries when PBF was introduced. By the time of the endline survey, more than 90 percent of the women in the control group have delivered in formal health facilities. When considering women below the wealth median, where baseline coverage of institutional delivery was lower, positive PBF impacts were found. As in DRC, the program in Zambia was found to impact earlier initiation of antenatal care (Friedman et al. 2016B). The other studies did not find impacts on antenatal care utilization indicators.

The results from DRC stand out also because of the estimated impacts on family planning. Some evaluations comparing PBF to not providing any additional financing (business as usual) found impacts on the use of modern contraceptives (e.g. de Walque et al. 2021; Khanna et al. 2021). However, studies comparing PBF and DFF interventions generally have not found such differences. The changes in delivery of family planning services in the PBF zones in DRC are evident across the different sources of data. During the baseline survey, only 41 percent of health centers and 72 percent of hospitals offered any family planning services and availability of products was very low. Although there are also positive trends in the control zones, the PBF

program significantly impacted the availability of family planning services and products, integration of family planning in antenatal and postpartum care, and the quality of family planning consultations. Relative to the coverage of modern family planning in the control group, the PBF impact represents 66 percent increase. However, in absolute term, the impact of 4 percentage points is moderate, and coverage of modern family planning remains below 10 percent even in the PBF group. It is possible that the initial impact was achieved because there was demand for such services in areas where supply was not available. It is hoped that the improvements in the availability and quality of these services, and the increase in coverage already achieved, will gradually increase demand even further.

Comparisons of impacts on quality of care are less straightforward as the other five studies did not measure impacts on hospitals.[10] In terms of improvements in structural quality at the health centers, the results from DRC are on the higher side. Impacts on such measures were found in Nigeria and Zimbabwe, but not in Cameroon or Zambia.[11] The study from Rwanda is the only one to find PBF to perform better than unconditional financing in terms of process quality, with improved content of antenatal care. In this study, we find positive impacts on process quality, but with stronger impacts at the hospital level. At the health centers, we see strong evidence of improvements in the quality of family planning consultations and weaker evidence of improvements in the quality of child curative care.

Study limitations

This study has some important limitations. First, the study was not designed to assess broader impacts of PBF on the health system (Meessen et al. 2011), such as improvements in data systems and data use, accountability, and capacity of health zones and provincial teams. Second, there are several ways in which outcomes in the control zones could have been influenced by the implementation of PBF in neighboring health zones. The same provincial health teams supervised both control and PBF zones and there are instances in which some of the PBF tools such as the business plan, indexes tool and the individual performance evaluation were used in the control zones although that didn't impact financing amounts. There were also movements of facility and health zone staff between PBF and treatment zones. We are unable to estimate to

[10] In some countries where PBF was randomized at the health center level, district hospital teams played a role in determining quality at the health center level. Therefore, all hospitals were under performance contracts.
[11] Basinga et al. (2011) do not report on such indicators in Rwanda.

what extent such factors affected our results, but it is reasonable to assume that if such effects exist, it will imply that we under-estimate the impact of the PBF program in the empirical results presented here. Finally, we had to change the design of the study and reduce its geographical coverage from 100 to 58 health zones due to an Ebola outbreak and higher than expected survey costs. While the internal validity of the study was not affected by this, statistical power to identify impacts of smaller magnitudes was reduced. The reduction in spatial scope of course also means that we can analyze program impacts in only six out of the eleven originally included provinces. As shown in the comparison between the different regions, there can be substantial variations in impacts across provinces.

The evaluation's findings might have also been affected by the COVID-19 pandemic. The endline data was collected in 2021-2022, with households reporting on care received in the two preceding years. The pandemic had both direct and indirect impacts on the health system and the population. Disruptions to essential maternal and child health services have been documented in many settings, including DRC (Ahmed et al. 2021; Shapira et al. 2021). The disruptions in DRC are estimated to be mostly concentrated in the beginning of the pandemic and in major urban areas (that are outside of this study's geographical coverage). We do not have reasons to believe that the pandemic impacted the population, supply chains, and ability to provide services differently in PBF and control zones. The PBF program was not suspended and payments to both PBF and control facilities were not disrupted. However, if the pandemic weakened implementation of any component, or restricted facilities' scope to increase utilization or improve their infrastructure, this might reduce the estimated impacts.

Cost effectiveness

The comparison between PBF and unconditional facility financing suggest moderately higher coverage and better quality outcomes in the PBF health zones. The results are at least as good or better in the PBF group for each type of outcomes. However, the differences between the groups are moderate in some areas. Progress has also been achieved in the control group, suggesting that some constraints to improvement in utilization and quality of care were lifted even without tying financing to performance. This is consistent with findings from impact evaluations in other countries that compared outcomes under unconditional facility financing to those under 'business-as-usual', i.e. no additional financing (e.g. Friedman et al. 2016; de Walque et al. 2021; Khanna et al. 2021). PBF is more expensive to implement, primarily because of the costs

associated with the provincial purchasing agencies (contract managers and verifiers) and the counter-verification. It is therefore important to understand the main health, systems and societal benefits associated with these additional costs. A cost-effectiveness analysis, building on the findings of this IE and using spending data by the project has been initiated and will shed a light on this question.

The financing models compared in this study, as designed and implemented, are of course not the only two options available for implementation in DRC. To inform the policy discussion, there is a need to identify potential ways in which the cost of implementing PBF could reduce with minimal risk to effectiveness and identify potential cost-effective ways to increase the impact of PBF. Different models of verification are under consideration in the country. The Strategic Purchasing Technical Cell of the Ministry of Public Health is exploring the use of a risk-based verification and a project financed by the European Union is using a bi-annual rather than a quarterly verification. Both options could substantially reduce the verification costs, but it needs to be assessed whether reduced verification is associated with an increased risk of fraudulent data reporting. In terms of enhancing impact, several approaches are already in the process of being tested. One example is the use of providers' performance on clinical vignettes to determine facilities' quality scores. Such modification is associated with a very small additional cost. An evaluation of a pilot of this approach will be available in the coming months to assess whether greater improvements in quality of health services are achieved. The World Bank-funded Multisectoral Nutrition and Health Project (*PMNS – Projet Multisectoriel de la nutrition et de la Sante)* is combining PBF at the health facilities with community-based approaches carried by community health workers, community associations and contracted NGOs, with a focus on nutrition and family planning. It should be assessed whether the combination of these approaches will lead to an increase in the very low coverage of modern contraceptives and nutrition-related services.

Conclusion

The results presented in this report suggest that PBF can result in improved health service coverage and quality in some areas relative to unconditional financing to health facilities. These improvements are achieved although substantial improvements have also been achieved in the facilities receiving unconditional financing. Further research is needed to understand relative

cost-effectiveness of the program in the DRC, equity implication and factors driving the heterogeneities in impacts.

References

Ahmed, T., Roberton, T., Alfred, J.P., Baye, M.L., Diabate, M., Kiarie, H., Mbaka, P., Mohamud, N.A., Mwansambo, C., Ndiaye, Y. and Nzelu, C., 2021. Indirect effects on maternal and child mortality from the COVID-19 pandemic: evidence from disruptions in healthcare utilization in 18 low-and middle-income countries.

Basinga, P., Gertler, P.J., Binagwaho, A., Soucat, A.L., Sturdy, J. and Vermeersch, C.M., 2011. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *The Lancet*, *377*(9775), pp.1421-1428.

Bertone, M.P. and Lurton, G., 2015. Availability and remuneration of Human Resources for Health in the DR Congo—final report. *Kinshasa: Ministry of Public Health and World Bank*.

Bertone, M.P., Lurton, G. and Mutombo, P.B., 2016. Investigating the remuneration of health workers in the DR Congo: implications for the health workforce and the health system in a fragile setting. *Health policy and planning*, *31*(9), pp.1143-1151.

Brenner, S., Madhavan, S., Nseya, C.K., Sese, C., Fink, G. and Shapira, G., 2022. Competent and deficient provision of childbirth services: a descriptive observational study assessing the quality of intrapartum care in two provinces of the Democratic Republic of the Congo. *BMC Health Services Research*, *22*(1), pp.1-11.

Clarke-Deelder, E., Shapira, G., Samaha, H., Fritsche, G.B. and Fink, G., 2019. Quality of care for children with severe disease in the Democratic Republic of the Congo. *BMC public health*, *19*(1), pp.1-12.

de Walque, D., Kandpal, E., Wagstaff, A., Friedman, J., Neelsen, S., Piatti-Fünfkirchen, M., Sautmann, A., Shapira, G. and Van de Poel, E., 2022. Improving Effective Coverage in Health: DO Financial Incentives Work?

de Walque, D., Robyn, P.J., Saidou, H., Sorgho, G. and Steenland, M., 2021. Looking into the performance-based financing black box: evidence from an impact evaluation in the health sector in Cameroon. *Health Policy and Planning*, *36*(6), pp.835-847.

DRC Ministry of Public Health, 2010. Plan national de developpement des Ressources Humaines de la sante (PNDRHS 2011-2015).

Fink, G., Fritsche, G., Samaha, H., Sese, C. and Shapira, G., 2022. Financially incentivized knowledge assessments to improve provider compliance with treatment guidelines: a cluster-randomized controlled trial. *Trials*, *23*(1), pp.1-14.

Fink, G., Kandpal, E. and Shapira, G., 2022. Inequality in the Quality of Health Services: Wealth, Content of Care, and the Price of Antenatal Consultations in the Democratic Republic of Congo. *Economic Development and Cultural Change*, *70*(3), pp.1295-1336.

Friedman, J., Das, A. and Mutasa, R., 2016. Rewarding Provider Performance to Improve Quality and Coverage of Maternal and Child Health Outcomes: Zimbabwe Results-Based Financing Pilot Program. *Washington, DC: The World Bank*.

Friedman, J., Qamruddin, J., Chansa, C. and Das, A.K., 2016. Impact evaluation of Zambia's health results-based financing pilot project. *Washington, DC: World Bank Group*.

Fritsche, G. and Peabody, J., 2018. Methods to improve quality performance at scale in lower-and middle-income countries. *Journal of Global Health*, *8*(2).

Fritsche, G.B., Soeters, R. and Meessen, B., 2014. *Performance-based financing toolkit*. world bank publications.

Huillery, E. and Seban, J., 2021. Financial incentives, efforts, and performances in the health sector: Experimental evidence from the democratic republic of Congo. *Economic Development and Cultural Change*, *69*(3), pp.1115-1164.

INS, 2019. Enquête par Grappes à Indicateurs Multiples, 2017–2018, Rapport de Résultats de L'Enquête.

Jain, A., Aruldas, K., Mozumdar, A., Tobey, E. and Acharya, R., 2019. Validation of two quality of care measures: results from a longitudinal study of reversible contraceptive users in India. *Studies in family planning*, *50*(2), pp.179-193.

Khanna, M., Loevinsohn, B., Pradhan, E., Fadeyibi, O., McGee, K., Odutolu, O., Fritsche, G.B., Meribole, E., Vermeersch, C.M. and Kandpal, E., 2021. Decentralized facility financing versus performance-based payments in primary health care: a large-scale randomized controlled trial in Nigeria. *BMC medicine*, *19*(1), pp.1-12.

Meessen, B., Soucat, A. and Sekabaraga, C., 2011. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform?. *Bulletin of the World Health Organization*, *89*, pp.153-156.

Ministère du Plan et Suivi de la Mise en œuvre de la Révolution de la Modernité (MPSMRM), Ministère de la Santé Publique (MSP) and ICF International. 2014. *Enquête Démographique et de Santé en République Démocratique du Congo*, *2014*.

Ministère de la Santé Publique. 2018, Guide de la mise en œuvre de la tarification forfaitaire en RDC.

Perales, N.A., Wei, D., Khadka, A., Leslie, H.H., Hamadou, S., Yama, G.C., Robyn, P.J., Shapira, G., Kruk, M.E. and Fink, G., 2020. Quality of clinical assessment and child mortality: a three-country cross-sectional study. *Health Policy and Planning*, *35*(7), pp.878-887.

Shapira, G., Ahmed, T., Drouard, S.H.P., Amor Fernandez, P., Kandpal, E., Nzelu, C., Wesseh, C.S., Mohamud, N.A., Smart, F., Mwansambo, C. and Baye, M.L., 2021. Disruptions in maternal and child health service utilization during COVID-19: analysis from eight sub-Saharan African countries. *Health policy and planning*, *36*(7), pp.1140-1151.

Soeters, R., Peerenboom, P.B., Mushagalusa, P. and Kimanuka, C., 2011. Performance-based financing experiment improved health care in the Democratic Republic of Congo. *Health Affairs*, *30*(8), pp.1518-1527.

Tripathi, V., Stanton, C., Strobino, D. and Bartlett, L., 2015. Development and validation of an index to measure the quality of facility-based labor and delivery care processes in sub-Saharan Africa. *PloS one*, *10*(6), p.e0129491.

Tripathi, V., Stanton, C., Strobino, D. and Bartlett, L., 2019. Measuring the quality of maternal and care processes at the time of delivery in sub-Saharan Africa: development and validation of a short index. *BMC pregnancy and childbirth*, *19*(1), pp.1-12.

World Bank, 2014. *Revue des dépenses publiques en RDC – santé*. World Bank – unpublished report.

World Health Organization. 2015. "Pregnancy, childbirth, postpartum and newborn care: a guide for essential practice (3rd edition)." https://www.who.int/publications/i/item/pregnancy-childbirth-postpartum-and-newborn-care.

World Health Organization, 2017. *WHO recommendations on newborn health: guidelines approved by the WHO Guidelines Review Committee* (No. WHO/MCA/17.07). World Health Organization.

World Health Organization, 2019. Trends in maternal mortality 2000 to 2017: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division.

Appendix 1: Supplementary Tables

Appendix Table 1: Quality checklist components and weights for the minimum package of activities

Quality Category	Share of overall score before midterm review	Share of overall score after midterm review
1. General organization e.g. presence, completion and archival of different forms and reports, catchment area map, means of communication	8%	8%
2. Management plan Quarterly management plan completed with different elements such as outreach strategies and negotiated fees	2%	6%
3. Financial management e.g. financial and accounting documents filled and present, compensation structure clear to providers	4%	3%
4. Indigent committee e.g. reports of monthly indigent committee meetings, documentation of care provided to indigent patients	5%	4%
5. Hygiene and sterilization e.g. waste disposal, sterilization procedure and supplies, containers for sharps and needles, cleanliness	8%	6%
6. Outpatient consultations e.g. management of registries and forms, physical state of waiting and consultation rooms, review of registers for child curative care, availability of supplies	28%	21%
7. Family planning e.g. staff member trained, confidentiality at consultation room, availability of family planning products	9%	8%
8. Laboratory e.g. presence of lab technician, management of registries, availability of equipment and supplies	5%	6%
9. Inpatient care e.g. management of registries, condition of observation areas (cleanliness, lighting, beds)	2%	4%
10. Medications and consumables e.g. management of stock inventories, correct stocking of products, correct disposal of drugs	5%	5%
11. Tracer drugs Availability of pre-specified list of essential drugs and consumables	6%	3%
12. Maternity e.g. correct use of partograms, availability of equipment and supplies, cleanliness and confidentiality in delivery room	7%	12%
13. Vaccination e.g. availability of vaccines, storage of vaccines, management of registries and forms	5%	7%

14. Antenatal care e.g. availability of equipment and supplies, review of registries	3%	3%
15. HIV/TB e.g. confidential consultation room, trained staff members, review of registries	2%	5%

Appendix Table 2: Quality checklist components and weights for the complementary package of activities

Quality Category	**Share of overall score before midterm review**	**Share of overall score after midterm review**
1. General organization	4%	6%
2. Management plan	1%	3%
3. Financial management	7%	4%
4. Indigent committee	3%	3%
5. Hygiene and sterilization	7%	9%
6. Outpatient consultations	17%	17%
7. Family planning	5%	4%
8. Laboratory	2%	4%
9. Inpatient care e.g. management of registries, condition of observation areas (cleanliness, lighting, beds)	25%	16%
10. Medications and consumables	3%	3%
11. Tracer drugs	5%	3%
12. Maternity	6%	7%
13. Antenatal care	1%	5%
14. HIV/TB	1%	4%
15. Surgery	13%	12%

Appendix Table 3: Baseline Balance in household indicators

	Variable	Control Group	PBF Group	P-value of mean difference test	Observations
Household Level indicators	Household members	5.41	5.49	0.823	2888
	Above wealth median	0.51	0.49	0.611	2889
	Health expenditure in past 12 months exceeded household resources	0.23	0.19	0.099	2883
Characteristics of recently pregnant women[b]	Age	27.21	27.26	0.757	2924
	Education: none	0.30	0.29	0.982	2924
	Education: primary	0.29	0.29	0.717	2924
	Education: at least some secondary schooling	0.41	0.42	0.834	2921
	Married	0.88	0.91	0.090	2924
	Number of Children	2.92	3.02	0.168	2884
Reproductive and maternal health indicators for recently pregnant women	Last pregnancy: at least one antenatal consultation	0.85	0.86	0.839	2251
	Last pregnancy: at least 4 antenatal consultations	0.42	0.44	0.569	2250
	Last pregnancy: initiation of antenatal care in first trimester	0.13	0.15	0.937	2244
	Institutional delivery for last birth	0.82	0.83	0.948	2248
	Received postnatal care after last birth	0.08	0.10	0.773	2248
	Use modern family planning method	0.04	0.04	0.914	2903
	Ever discussed FP with health providers	0.43	0.45	0.195	2889
Characteristics of children under 5	Child's age in months	21.46	22.20	0.489	3997
	Female	0.50	0.51	0.580	3997
	Child's mother lives in the household	0.99	1.00	0.191	3997
Child health services	Growth monitoring in past 6 months	0.04	0.06	0.279	3970
	Children 13-24 months: All basic vaccinations	0.26	0.33	0.145	835
	Children 13-24 months: BCG vaccination	0.88	0.88	0.555	836
	Children 13-24 months: Penta 3	0.57	0.58	0.573	839

Note: Data from household surveys. The balance test is estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.

Appendix Table 4: Baseline Balance in health center indicators

	Variable	Control Group	PBF Group	P-value of mean difference test[a]	Observations
Sector	public	0.88	0.87	0.600	290
Availability of medical supplies and protocols	Equipment index	0.31	0.31	0.661	260
	Vaccines index	0.05	0.04	0.636	207
	Essential drugs index	0.44	0.46	0.648	190
	Family planning products index	0.10	0.15	0.191	207
	Protocols index	0.30	0.30	0.776	286
Infection prevention and control	Use improved source of water	0.32	0.36	0.636	289
	Hand wash stations	0.20	0.24	0.330	290
	Incinerator	0.10	0.11	0.585	288
	Sterilization procedures posted	0.02	0.03	0.514	289
	Disinfection procedures posted	0.17	0.15	0.666	256
	Container for sharps	0.83	0.75	0.061	288
Service fees	Chare flat fees	0.51	0.49	0.899	289
	Fee expemption for poor/indigent users	0.99	1.00	0.311	290
	Fee for uncomplicated delivery (CDF)	8142	8207	0.917	279
	Fee for antenatal consultation (CDF)	2586	2166	0.197	283
	Fee for curative consultation (CDF)	2948.03	2580	0.421	273
Provider characteristics	Age	40.41	42.69	0.020	565
	Manager	0.16	0.19	0.328	565
	Unit chief	0.45	0.41	0.376	564
	Female	0.46	0.44	0.509	565
	Nurse	0.76	0.82	0.328	565
	Midwife	0.04	0.03	0.378	565
	Years of experience	12.00	13.59	0.105	564
	Years in facility	5.15	6.43	0.039	560
Attendance	Days absent during last 30 days	0.86	0.69	0.453	562
	Days worked in past 7 days	5.95	5.99	0.770	564
Motivation	General motivation	6.60	6.86	0.343	558
Provider's knowledge measured through clinical vignettes	Vaccination calendar	0.76	0.76	0.778	556
	Severe dehydration scenario	0.67	0.66	0.488	565
	Ear infection scenario	0.71	0.73	0.186	564

	Severe lower respiratory infection scenario	0.69	0.68	0.369	564
	Non-severe pneumonia scenario	0.74	0.72	0.105	564
	Non-severe malaria scenario	0.71	0.68	0.048	564
	Pregnancy danger signs	0.47	0.47	0.932	564
Direct observations of antenatal care	ANC1 consultation score	0.49	0.49	0.928	133
	Fees paid (CDF)	3766	2476	0.209	129
Direct observations of child curative care	Assessment score	0.52	0.49	0.606	220
	Danger signs verified	0.52	0.45	0.389	220
	Fees paid (CDF)	5765	5479	0.719	220

Note: Data from health center assessments. The balance test is estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level.

Appendix Table 5: Baseline Balance in hospital indicators

	Variable	Control Group	PBF Group	P-value of mean difference test[a]	Observations
Sector	public	0.73	0.36	0.046	29
Availability of medical supplies and protocols	Equipment index	0.58	0.60	0.547	26
	Vaccines index	0.01	0.00	0.227	29
	Essential drugs index	0.69	0.71	0.365	29
	Family planning products index	0.19	0.31	0.161	29
	Protocols index	0.57	0.61	0.412	29
Infection prevention and control	Use improved source of water	0.53	0.43	0.457	29
	Hand wash stations	0.67	0.93	0.183	29
	Incinerator	0.67	0.71	0.530	29
	Sterilization procedures posted	0.33	0.14	0.041	29
	Disinfection procedures posted	0.36	0.36	0.900	28
	Container for sharps	0.93	0.93	0.295	29
Service fees	Chare flat fees	0.20	0.14	0.726	29
	Fee expemption for poor/indigent users	1.00	1.00		29
	Fee for uncomplicated delivery (CDF)	12200.00	10150.00	0.547	29
	Fee for antenatal consultation (CDF)	2913.33	2115.38	0.183	28
Provider characteristics	Age	42.73	41.62	0.412	115
	Manager	0.03	0.09	0.050	115
	Unit chief	0.68	0.60	0.083	115
	Female	0.50	0.44	0.469	115

	Nurse	0.73	0.78	0.826	115
	Midwife	0.00	0.00		115
	Years of experience	10.82	12.93	0.698	115
	Years in facility	10.92	9.13	0.351	115
Attendance	Days absent during last 30 days	0.92	0.38	0.243	115
	Days worked in past 7 days	6.12	6.04	0.813	115
Motivation	General motivation	5.95	6.39	0.389	114
Provider's knowledge measured through clinical vignettes	Vaccination calendar	0.71	0.79	0.070	112
	Severe dehydration scenario	0.65	0.65	0.678	115
	Ear infection scenario	0.72	0.70	0.175	115
	Severe lower respiratory infection scenario	0.71	0.68	0.092	115
	Non-severe pneumonia scenario	0.74	0.73	0.503	115
	Non-severe malaria scenario	0.70	0.73	0.166	115
	Pregnancy danger signs	0.50	0.49	0.464	115
Direct observations of antenatal care	ANC1 consultation score	0.65	0.52	0.388	38
	Fees paid (CDF)				
Direct observations of child curative care	Assessment score	0.63	0.60	0.631	71
	Danger signs verified	0.59	0.67	0.103	71
	Fees paid (CDF)	8898.48	10180.26	0.401	71

Note: Data from hospital assessments. The balance test is estimated with an unadjusted regression, with control for the randomization blocks and clustering of standard errors at the health zone level

Appendix 2: Difference-in-differences specification

In this appendix, we present results using an alternative regression specification based on a difference-in-differences approach, and compare the results to those presented in the tables above. The following regression model is used for the difference-in-differences approach:

$$y_{izr} = \beta_0 + \beta_1 PBF_z + \beta_2 \text{Endline}_r + \beta_3 PBF_z * \text{Endline}_r + \gamma X_{izr} + +\varepsilon_{izr},$$

where y_{izr} is an outcome for health facility/ individual i in health zone z, measured during survey round r. $Endline_r$ takes value 1 if the outcome is measured during the endline survey and 0 otherwise. X_{izr} is a vector of individual/household/facility characteristics, including health zone dummies. ε_{iz} is a normally distributed error term, clustered at the health zone level.

As can be seen in **Appendix Table 6** and **Appendix Table 7,** there is little change in results when using the difference-in-differences approach relative to using the estimation based only on the endline data.

Appendix Table 6: Comparison between estimation of the benchmark regression specification and a difference-in-differences specification, service coverage

	Estimation using Endline data only [c]				Difference-in-Differences specification [d]		
	PBF Impact	Standard Error	p-value		PBF Impact	Standard Error	p-value
Child health services [a]							
Children under 5: growth monitoring in past 6 months	0.01	0.02	0.495		0.00	0.02	0.858
Children 13-24 months: All basic vaccinations	0.00	0.05	0.977		-0.11	0.09	0.245
Children 13-24 months: BCG vaccination	0.03	0.03	0.351		0.00	0.04	0.986
Children 13-24 months: Penta 3	-0.03	0.05	0.576		-0.06	0.08	0.467
Maternal care during most recent pregnancy [b]							
Any ANC [c]	0.02	0.01	0.207		0.00	0.03	0.974
4+ antenatal consultations	0.04	0.05	0.435		0.00	0.05	0.927
ANC in first trimester	0.08***	0.02	0.001		0.08***	0.03	0.007
ANC with tetanus shot	0.03	0.03	0.206		-0.03	0.04	0.452
ANC with anti-malarial	0.01	0.04	0.784		-0.03	0.05	0.554
Institutional delivery	0.03	0.02	0.197		0.03	0.03	0.332
Received postnatal care	0.03	0.04	0.491		0.02	0.05	0.768
Family planning for women with birth in previous 2 years							
Use modern contraceptive	0.04**	0.02	0.022		0.04**	0.02	0.043
Ever discussed FP with health providers	0.08**	0.04	0.031		0.02	0.06	0.692

Note: Data from household surveys.
[a] Sample of children of women who were pregnant in the two years preceding the survey.
[b] Sample of women who had a pregnancy in the two years preceding the survey
[c] ANC = Antenatal care
Significance level: * 10%; ** 5%; *** 1%

Appendix Table 7: Comparison between estimation of the benchmark regression specification and a difference-in-differences specification, health center variables

	Estimation using Endline data only [c]				Difference-in-Differences specification [d]		
	PBF Impact	Standard Error	p-value		PBF Impact	Standard Error	p-value
Availability of equipment, medical supplies and protocols at health centers							
Basic equipment	0.04**	0.02	0.041		0.04*	0.02	0.054
Essential drugs	0.01	0.03	0.586		0.02	0.04	0.532
Vaccines	0.05	0.06	0.450		0.05	0.07	0.459
Family planning products	0.19***	0.04	0.000		0.17***	0.06	0.005
Clinical and administrative protocols	0.11**	0.04	0.010		0.10**	0.05	0.033
Infection prevention and control at health center							
Improved source of water	0.00	0.07	0.995		-0.03	0.08	0.681
Handwashing stations in consultation rooms	-0.07	0.07	0.351		-0.09	0.08	0.290
Incinerator	0.12*	0.06	0.061		0.08	0.07	0.233
Sterilization procedures posted	0.27***	0.06	0.000		0.25***	0.07	0.000
Disinfection procedures posted	0.12	0.09	0.165		0.13	0.09	0.157
Containers for sharps and needles	0.12**	0.05	0.016		0.21**	0.08	0.015
Delivery of reproductive and maternal health services at health centers							
Days offering ANC services per week	0.42	0.28	0.143		0.48	0.33	0.144
Offer family planning service	0.24***	0.06	0.000		0.20*	0.11	0.071
Delivery of reproductive and maternal health services at health centers							
Charge flat fees	0.21***	0.07	0.003		0.22*	0.12	0.067
Charge for lab tests	-0.37***	0.05	0.000		-0.36***	0.09	0.000
Charge for inpatient care	-0.28***	0.06	0.000		-0.25***	0.08	0.002
Charge for consumables	-0.33***	0.07	0.000		-0.29**	0.11	0.012
Charge for drugs	-0.39***	0.07	0.000		-0.38***	0.10	0.001
Fees posted	0.15**	0.06	0.011		0.19**	0.07	0.010
Fee exemptions for the poor	0.18***	0.05	0.000		0.19***	0.05	0.001
Delivery fee	-1914**	882	0.034		-2062*	1156	0.080
ANC fee	-735***	273	0.009		-323	286	0.263
Curative care fee	596*	336	0.081		1003**	469	0.037

Notes: Data from facility-based surveys. The list of items included in each index is presented in the appendix.
Significance level: * 10%; ** 5%; *** 1%

Appendix 3: Definition of indicators used in the analysis

Availability of medical supply indexes

1. List of items included in the basic equipment index and supplies: thermometer, adult stethoscope, child stethoscope, blood pressure cuff for adult, blood pressure cuff for child, otoscope, delivery room light, partograph, measuring tape, aspirator, delivery bed, obstetric stethoscope, absorbable suture, non-absorbable suture, sterile gauze, clamp, baby scale, adult scale, microscope, centrifuge, needle holder, protective apron, gloves, delivery kit/box, autoclave, electric sterilizer, steam sterilizer.
2. List of items included in the essential drugs index: paracetamol, ibuprofen, chlorpheniramine, oxytocin, mebendazole, iron sulfate, amoxicillin pill, amoxicillin injection, amoxicillin syrup, cotrimoxazole pill, cotrimoxazole syrup, doxycycline, sulfadoxine/pyrimethamine, oral rehydration salts, metronidazole, gloves, catheter 18g, catheter 22g, infusor, ringer solution, isotonic glucose, needle 5ml.
3. List of items included in the vaccines index: BCG, pentavalent, polio, yellow fever, measles, tetanus toxoid
4. List of items included in the family planning index: contraceptive pill, male condom, female condom, diaphragm, morning after pill, injectable contraceptive, implant, intrauterine device
5. List of items included in the protocols index: sterilization, malaria, acute respiratory infection, diarrhea, treatment directives, tuberculosis, health management information system, child vaccination, reproductive health, sexually transmitted infections, HIV, PMTCT, essential drugs, drug supplies, drug side effects, EMONC.

Content of care indexes

1. Child curative care – overall assessment score: greet patient, weigh child, ask about diarrhea, skinfold check, ask about cough, ask about stridor, check respiratory frequency, undress child, auscultation, ask about fever, asks about measles, check eyes, check nose, check skin, check throat, check ears, check neck, ask about ear pain, check palms, check feet, test for malaria.
2. Child curative care – verification of danger signs: inability to drink, vomiting, lethargic behavior, convulsions
3. Child curative care – essential checks: taking temperature, measuring respiratory rate, undressing child to inspect upper body, auscultating chest, skinfold test, checking arms and eyes, and checking feet and ankles and checking ear and neck.
4. First antenatal consultation – history taking index: ask about age, medication, last cycle, previous pregnancies, bleeding, fever, headache or blurry vision, swelling, fatigue,
5. First antenatal consultation – physical exams index: provider washes hands before touching pregnant woman, measure blood pressure, weight, check swelling, palpate abdomen
6. First antenatal consultation – diagnostic tests index: test for HIV, syphilis, hemoglobin level, urine test

7. First antenatal consultation – preventive care index: iron with or without folic acid, tetanus toxoid shot, malaria chemoprophylaxis
8. First antenatal consultation – counseling index: counseling on nutrition during pregnancy, birth preparation, pregnancy danger signs
9. Family planning consultation – respectful care score: provider addresses client respectfully, provider encourages questions, client satisfied with explanations, respect of client's privacy
10. Family planning consultation – method selection score: ask if client want any more pregnancies in the future, ask about preference for a method, ask about past contraceptive use, discuss multiple methods, doesn't show bias towards one method
11. Family planning consultation – effective use score: explain how to use, explain how the method works, explain side effects, explain how to manage side effects
12. Family planning consultation – continuity of use score: discuss when the method needs to be renewed, where to renew or restock, possibility of changing methods, suggest a follow up consultation
13. Labor and delivery – overall delivery score: provider explain sequence and objective of each procedure, ask about HIV status, ask about blurred vision during pregnancy, ask about bleeding during pregnancy, measure blood pressure, wash hand before physical exams, wear gloves for vaginal exam, disinfect hands before vaginal exam, use partograph, neonatal resuscitator laid out in delivery room, uterotonic injection, assess completeness of placenta and membranes, assess perineal and vaginal lacerations, keep baby warm, place newborn on mother's abdomen, cut umbilical cord, check for bleeding, blood pressure and pulse and palpate uterus in first hour after delivery, assist mother in initiating breastfeeding,
14. Labor and delivery – postpartum care score: assess bleeding, blood pressure and pulse and check uterus at least four times in the hour after birth and at least once in the following hour, assess bleeding in the third stage of delivery.
15. Labor and delivery – newborn care score: keep newborn warm, place newborn on mother's abdomen for skin-to-skin, cut umbilical cord, assist mother in initiating breastfeeding, take APGAR scores, measure newborn temperature, weigh newborn, clean newborn eyes

Health provider satisfaction index

Provider satisfaction index: satisfaction with respect to information received on own performance, autonomy, collaboration with local health development committee, support from supervisors, reward for effort, ability to use one's skills, training opportunities, security, work conditions, leave, work hours, team work, relationship with management, salary, regularity of salary, reward for performance, benefits, overall income, promotion opportunities.

User satisfaction

1. Satisfaction index in exit interviews: cleanliness, staff courteous and respectful, staff provide good explanations, fees are affordable, waiting time, staff respect confidentiality, staff spend sufficient time with users, hours of services, quality of care.

2. Satisfaction index in household survey: heads of households reported their perceptions on the nearest health facility. Respondents are defined to be satisfied if they respond that the following aspects are good or very good: hours of operation, cleanliness, availability of equipment, availability of drugs, efficacy of service, staff is welcoming, waiting time, respect by providers, staff competency, staff interaction with patients/users, providers encourage asking questions, cost of services, cost of drug.