Impact of performance-based financing on effective coverage for curative child health services in Burkina Faso: Evidence from a quasi-experimental design

Authors: Jean-Louis Koulidiati[1], Manuela De Allegri[1], Julia Lohmann[2], Michael Hillebrecht[3], Joel Arthur Kiendrebeogo[4], Saidou Hamadou[5], Hervé Hien[6], Paul Jacob Robyn[5], Stephan Brenner[1]

[1] Heidelberg Institute of Global Health, University Hospital and Faculty of Medicine, Heidelberg, Germany
[2] Department of Global Health and Development, London School of Hygiene & Tropical Medicine, London, UK
[3] Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Germany
[4]Université Joseph Ki-Zerbo, Ouagadougou, Burkina Faso
[5] The World Bank, 1818 H Street, NW Washington, DC 20433 USA
[6] Centre MURAZ, 2054 Avenue Mamadou KONATE, 01 B.P. 390, Bobo-Dioulasso, 01, Burkina Faso.

Corresponding author: Jean-Louis Koulidiati

E-mail: jean-louis.koulidiati@uni-heidelberg.de Address: Heidelberg Institute of Global Health, Im Neuenheimer Feld 130.3, 69120 Heidelberg, Germany, Tel: (+49) 6221 56 35039 Fax: (+49) 6221 565948

Summary

Objective: To evaluate the impact of Performance-Based Financing (PBF) on effective coverage of child curative health services in primary healthcare facilities (PHC) in Burkina Faso.

Methods: An impact evaluation of a PBF pilot program, using an experiment nested within a quasi-experimental design, was carried out in 12 intervention and 12 comparison districts in six regions of Burkina Faso. Across the 24 districts, PHC facilities (537 both at baseline and endline) and households (baseline=7,978 endline=7,898) were surveyed. Within these households, 12,350 and 15,021 under-five-year-olds (U5YOs) caretakers were interviewed at baseline and endline respectively. Linking service quality to service utilization, we used difference-in-differences to estimate the impact of PBF on effective coverage of curative child health services.

Results: Our study failed to detect any effect of PBF on effective coverage. Looking specifically into quality of care indicators, we detected a positive effect of PBF on structural elements of quality of care related to general service readiness, but not on the overall facility quality score, capturing both service readiness and the content of childcare.

Conclusion: The current study makes a unique contribution to PBF literature, as this is the first study assessing PBF impact on effective coverage for curative child health services in low-income settings. The absence of any significant effects of PBF on effective coverage suggests that PBF programs require a stronger design focus on quality of care elements especially when implemented in a context of free healthcare policy.

Introduction

Performance-based financing (PBF) in the health sector refers to the provision of financial incentives to healthcare providers tied to the achievement of specific quantity and/or quality indicators [1]. Over the last two decades, PBF has received considerable attention , particularly in Africa [2], where PBF schemes have been mainly rolled out, with a specific focus on maternal and child health services [2].

As yet, a handful of studies have evaluated PBF effects on the utilization and quality of care (QoC) of maternal and child health services [3-9]. Looking specifically at the evidence related to PBF and child services, there are a relative paucity of studies, particularly in low-income settings. Nevertheless, we note the impact evaluation in Rwanda, which found an increase in the number of institutional deliveries and preventive care visits by children due to PBF [4]. A cluster-randomized controlled trial in the Philippines, showed that PBF improved the health status of children after hospital stay [6]. Other studies revealed that quality of services (e.g. the number of qualified health professionals, patient satisfaction) increased after PBF implementation [5, 10, 11].

Some authors argue that PBF can likely improve health system performance [12], thus supporting systems moving towards Universal Health Coverage (UHC). However, despite these findings, more research on PBF impact is still needed [13, 14], particularly in sub-Saharan African countries that are working towards achieving UHC. One essential objective of UHC is effective coverage (EC) [12], which refers to the proportion of a population in need of a service that had a positive health outcome from the service [15]. Thus, corresponding EC metrics, are highly relevant indicators of intervention effectiveness in regards to UHC [12]. Based on the existing literature outlining the PBF theory of change [16-18], we postulated that PBF intervention would act on both the supply- and the demand-side to improve both quality and utilization of child health services. We expected that the combination of the following factors, namely i) increased revenues (due to PBF incentives); ii) improved facility managerial autonomy; and iii) regular supervision and verification, would enable healthcare professionals to improve quality of service provision and motivate them to stimulate communities to utilize services. In turn, we postulated that changes on the supply-side would induce changes on the demand, resulting in an increase in health service utilization. As effective coverage is driven by both utilization and quality of service provision, we expected that PBF would increase effective coverage of CHS.

To the best of our knowledge, we note no study assessing the impact of PBF on effective coverage for child health services. To fill this knowledge gap, this study evaluates the effect of PBF on effective coverage of children in need of curative healthcare in Burkina Faso.

Methods

PBF intervention in Burkina Faso

In 2011, the Ministry of Health (MoH) of Burkina Faso, with the support of the World Bank, implemented a PBF pre-pilot in three districts to improve the utilization and the quality of primary healthcare (PHC) services. In 2014, the MoH decided to expand this program to an additional 12 districts across six regions [19] and implement specific additional components intended to enhance equity. As a result, four different PBF approaches (PBF1-4) were conceptualized and implemented. Three of them (PBF2, PBF3, PBF4) had specific equity components. Details of these PBF components are reported elsewhere [20, 21].

Each of these components included the same performance indicators related to CHS, such as under-five curative care, child vaccination, and growth monitoring visits [20]. In 2016 (two years after the start of the implementation), the government of Burkina Faso introduced a free healthcare (FHC) policy removing user fees for under-five children for the whole country [22]. This affected the PBF implementation in so far as this policy resulted in some adjustments to PBF rates for those fee exempt services. Unit prices for all services included in the FHC policy benefit package were adjusted to match PBF1 prices in PBF2, PBF3, and PBF4 (by removing any additional subsidy). As PBF design significantly changed after two years due to the FHC policy, we only estimate in this study the overall PBF effects, and do not look at the additional benefit of each component.

Study design

Our study relied on a quasi-experimental design with independent controls. Within each of the six study regions, two intervention and two control districts were identified. Across the 24 districts, facility and household surveys were conducted in 2013 before PBF implementation (baseline survey) and in 2017 (endline survey). Details of the study design are reported elsewhere [21].

Samples, data sources and data collection

Both baseline and endline surveys collected the same set of primary data obtained from the following four sources:

a) In the facility sample, a facility inventory was conducted assessing the availability of staff, infrastructure, equipment, drugs, supplies, and consumables.

b) Across all sampled facilities, under-five-year-old children (U5YO) visiting to the outpatient department on the day of the study visit were identified, but only first-time cases (i.e. not follow-up visits) were included in the U5YO case sample. For each sampled U5YO case, patient-provider interactions were directly observed to assess providers' adherence to clinical standards outlined by the Integrated Management of Childhood Illness (IMCI) framework.

c) To evaluate providers' theoretical familiarity with IMCI guidelines related to the management of relatively common, but severe or life-threatening pediatric case presentations (i.e. dehydration, fever, and respiratory distress), a vignette-based knowledge assessment of three different case scenarios was conducted with each sampled clinical provider.

d) For the household survey, households were included following a two-stage sampling technique. First, one village was randomly selected from all villages located within a surveyed facility's catchment area. Second, in each selected village those households with at least one currently pregnant woman or at least one woman who gave birth or had a miscarriage within the previous two years were identified. After listing all eligible households in a village, 15 of them were randomly selected to be surveyed. Structured interviews were conducted with the caregivers of each U5YO child living in a sampled household to collect information on any illness episodes and. care-seeking behavior during the month preceding the survey date and on child and caregiver characteristics (i.e. age, sex, education, socio-economic status, etc.).

Outcome measures

In line with a published conceptual work [23], the primary outcome measure was effective coverage for CHS defined as those U5YO identified as crudely covered (i.e. service use by any U5YO conditional upon reporting an illness episode in the preceding month) who sought care at a facility meeting our definition on 'high quality' based on QoC scores. These QoC scores were created from indicators grouped into three quality dimensions based on the validated indices of Gouws'work and Donabedian framework[24, 25]: a) Management of common childhood diseases (MCCD) based on indicators related to routine clinical management taken from the case observations; b) Management of severe childhood diseases (MSCD) based on indicators related to acute clinical management taken from the provider knowledge assessment, and c) General service readiness based on indicators related to service readiness taken from the facility inventory. Based on this score, facilities were grouped into 'high quality' and 'low quality' providing facilities (see Appendix 1). As components of EC and for the completeness of the analysis, we also used crude coverage and facility quality as secondary outcome measures.

Co-variates

Based on a previous work on determinants of crude and effective coverage [26], to improve the precision of our impact estimates, we adjusted our estimation models by the following covariates: *i) Health worker characteristics* (i.e. years in service), *ii) U5YO children characteristics* (i.e. sex, age, number of symptoms as a proxy of illness severity)*, iii) Household characteristics* (i.e. distance to nearest health facility , household socio-economic status , household dependency ratio).

Analytical approach

We used a difference-in-difference (DiD) multivariate regression model [27] to estimate the effect of PBF vis à vis standard service provision:

$$Y_{dfit} = \alpha_f + \beta \cdot 2017_t + \lambda \cdot PBF_d + \delta \cdot [PBF_d * 2017_t] + \phi \cdot X_{it} + \epsilon_{dfit},$$

where Y_{dfit} is the outcome variable for individual *i* from catchment area *f* in district *d* at time *t =[2013, 2017]*; 2017_t is a dummy variable designating observations at endline; PBF_d represents a dummy variable that equals 1 for individuals for PBF districts and 0 for control districts; X_{it} stands for a set of additional control variables, and ϵ_{dfit} the error term. Coefficient α_f is facility/catchment area fixed effects capturing time-invariant unobserved differences across facilities/catchment areas, coefficient β represents the time fixed effect, and coefficient δ serves as the DiD estimate for the effect of being in a PBF district when compared with non-PBF districts. We clustered standard errors at the district level, since the treatment was assigned at this level. However, the relatively low number of clusters (n=24) might have led to somewhat biased standard errors and an over-rejection of the null hypothesis [28, 29]. Therefore, to correct potential biases, we applied 'wild bootstrapping', which leads to substantially lower over-rejection of the H_0 [30, 31].

To test the common trend assumption [30], we inspected trends of selected PBF incentivized routine child health indicators taken from the Health Management and Information System (HMIS) data for the period (11 months) preceding our baseline survey. Findings supported the assumption of parallel trends between intervention and control facilities before the intervention launch (see Appendix 2, Figure 1). The FHC policy, implemented in both intervention and control districts at the same time, did not affect our estimates because this policy was acting as a group invariant factor [30, 32].

A first inspection of descriptive statistics also found in two control districts an important increase in facility quality score. Hence, we conducted a sensitivity analysis by removing these control districts from the analysis to estimate the PBF effects.

Ethical considerations

This study obtained clearance from the authors institutes. Written informed consent was obtained from all study participants.

Results

Table 1 presents the sample characteristics at baseline and endline surveys. In total, 537 primary healthcare facilities (420 intervention vs 117 control) were surveyed both at baseline and endline. In these facilities, 1,361 health providers were interviewed and 2,046 U5YO cases were observed at baseline, and at endline 1,769 health providers were interviewed and 2,649 U5YO cases observed. Within the catchment areas of these facilities, 7,978 and 7,898 households were surveyed at baseline and endline respectively. Within these households, 12,350 and 15,021 U5YOs caretakers were interviewed at baseline and endline respectively. At baseline, 650 U5YOs reported an illness episode in the prior month whereas 996 did so at endline. We used a two-sample t-test to test for differences in outcome measures and covariates at baseline, finding no statistical difference between intervention and control groups for most indicators (See Appendix 2, Table 1).

Table 2 shows the estimated effects in percentage points (pp) of PBF on EC for curative CHS for both the unadjusted and adjusted models. The comparison of PBF versus *status quo* shows that PBF in Burkina Faso had no significant effect on effective coverage compared with controls. Considering exclusively the crude coverage outcome, no effects were found when comparing PBF with *status quo*.

Looking at the effects of PBF on the dimension-specific quality scores and the overall facility quality score (Table 3), we detected a positive effect only for the general service readiness score (0.30 score point increase on a scale from 0 to 5; CI 95% [-0 .11   0.68]) after model adjustment.

Tables 4 show the findings of the sensitivity analysis that removed from the analysis two control districts as outliers. We now found a positive but statistically non-significant PBF effect on effective coverage, as well as on the overall quality score.

Discussion

The current study contributes to the existing literature on the impact of PBF in sub-Saharan Africa. Our study failed to detect any significant effects of PBF on EC for curative CHS in Burkina Faso, nor on crude coverage or overall facility-specific quality. As EC is defined by these latter elements, our discussion focuses on potential reasons that might have prevented stronger PBF-related effects on both crude coverage and QoC to explain the observed lack of impact on effective coverage.

Based on Tanahashi's framework of health service coverage [33], crude coverage itself is conditional on the availability, accessibility, and affordability of services. To increase utilization and QoC, the PBF intervention in Burkina Faso focused on both improving service provision to and accessibility of the most destitute population sub-groups.

Given our study context with about 60% of households in both the PBF and control samples located within 5 km from their catchment facility (Table 1 of Appendix 1), we assume that challenges related to service availability are secondary to our study context. With respect to accessibility of CHS, however, the selective focus of the Burkina PBF on expanding service coverage for the most destitute households might have had only a limited effect on service use of U5YOs. Interestingly, an impact evaluation in Rwanda assessing PBF's equity effect on child health service utilization found that children of poorer households were more likely to receive medical treatment once seen at a facility (i.e. effective care), but were not more likely to seek care (i.e. accessibility) compared with children from less poor households [3]. This might indicate that pro-poor targeting elements of PBF schemes play a more limited role for child health service use compared with the use of other health services.

With respect to affordability of CHS, the introduction of the FHC policy for child curative services nationwide might have contributed to the absence of a strong PBF impact on crude coverage. A recent study on the FHC policy in Burkina showed that prior to this policy, there were significant positive effects of the PBF on maternal and child service utilization. After the introduction of the FHC policy, however, PBF no longer impacted service utilization in intervention compared with control facilities [34]. Given these observed changes in service utilization at the facility level in response to the FHC policy, it might be likely a similar trend pattern occurred with respect to service use at the level of households in the respective catchment areas. In fact, as shown in Table

2 of the Appendix 1, the rather parallel upward trends in observed crude coverage between baseline and endline further supports this assumption. This pronounced secular effect of the FHC on service accessibility might therefore have obscured the effect of the PBF's targeting component.

Regarding the effectiveness of services based on our quality scores, the implemented PBF schemes seemed not to have been able to produce any substantial change. Especially in two of the control districts we found a substantial improvement in our quality scores for the case management dimensions, indicating that with respect to CHS quality health system strategies other than the currently implemented PBF design might have been more effective. Considering these two districts as outliers to be removed from the analysis, our sensitivity analysis showed more positive effect sizes for EC and the different quality dimensions (Table 4). However, these changes might statistically still have occurred by chance and thus be unrelated to the PBF.

As PBF intervention increased revenues of facilities and provided them managerial means, QoC is expected to improve. However, our study did not find any significant effects on QoC. This might be explained by the incentive structure of the PBF scheme. As this scheme followed a "carrot and carrot" design, in which quantity (i.e. each new U5YO case) and quality performance were verified and reimbursed separately [20], quantity performance are reimbursed by the same unit rate regardless of the service quality provided, which might have reduced health workers' attention to the QoC aspects of child consultation services [35]., . In addition, with respect to quality performance, providers' observance of IMCI treatment protocols (reflected in our case management scores) contributed only about 10% to facilities' overall quality bonus, compared with a roughly 25% contribution of quality indicators directly or indirectly related to service readiness [20]. PBF programs largely (over 90%) measure quality of service delivery in relation to structural elements of quality such as infrastructure, equipment, supplies etc. [36]. Further, our quality score's stronger focus and weighting of elements related to case management over service readiness elements, especially compared with the PBF implementation focus and applied weights. The difference in incentive weights in combination with our quality evaluation focus (heavier on case management elements) might have contributed to our observed effect patterns with mainly negative effects on the case management dimensions, and a rather significant positive effect on readiness (Table 3).

.

An additional explanation for the absence of strong effects on CHS quality could be an over-proportional increase of workload for service providers at PBF facilities, especially with respect to administrative tasks during patient consultations, which might have taken providers' focus away from clinical protocol adherence [35]. Also, in response to the introduction of the FHC policy, the PBF program modified the unit prices used for reimbursing quantity performance, thus reducing the subsidy component applied to more remote and less accessible facilities were substantially reduced [35]. This might have resulted in limited financial capacity to invest in quality improvement strategies for affected facilities. Furthermore, during PBF implementation, implementers faced several issues in the fidelity of implementation of the PBF program, especially delayed subsidy and bonus payments to facilities, which in turn negatively affected the link between performance and payment, thus limiting health workers' motivation and buy-in [35, 37, 38].

As policy implications, our findings invite policy makers to re-focus the PBF scheme in Burkina Faso more closely on quality of care, particularly on content of child health services. Also, in Burkina Faso, one of the biggest challenges of the FHC policy is how to efficiently purchase health services [39]. Specifically, linking quality to purchasing of health services remains an unsolved issue. Lessons learned from the PBF program, particularly on PBF design to improve quality of care and further its implementation could be an asset for an efficient implementation of FHC policy as it was the case in Burundi [40].

Methodological considerations

This study has noteworthy strengths and limitations. The common trend assumption has been tested before DiD estimation. In addition, as the number of clusters was low, a correction through wild bootstrapping was applied to get unbiased standard errors. An additional strength is that the EC measures are valid for the technical quality we define and assess. First, content of score variables are taken from international or national standard protocols, thus face validity exists. Further, the combination of indicator measures into indices follows a previously validated approach [24]. To this extent, given the systematic approach to derive these measures they are reliable. As limitations, the sensitivity analysis revealed changes in PBF effects when removing outlier control districts from the analysis, this calls for caution in DiD approach because in a real-life setting important changes for external reasons in control districts can seldom be fully excluded. Another methodological weakness is the lack of randomization of intervention and control within the same districts. As the impact evaluation engaged different stakeholders (MoH, the World Bank and the research team), this study design had to accommodate both interests [21]. Also, there is a slight misalignment between the quality indicators (particularly content of care) of this study and those incentivized by the PBF program. While the former was based on direct observation of childcare (i.e. gold standard) and health workers' knowledge, the PBF program quality indicators mainly relied on document review. Also, in the absence of a validated empirically defined threshold to identify the minimum quality in the context of primary care delivery in low-income countries, we categorized health facilities to high quality and low quality using an arbitrary threshold. Nonetheless, since our categorization was informed by prior literature [41, 42], we are confident that our threshold is valid.

Conclusion

Our findings make a unique contribution to PBF literature, as this is the first study assessing PBF impact on EC for curative CHS in low-income settings. The absence of any significant effects of PBF on effective coverage suggests that effective curative care for U5YOs seems to be more complex and requires a stronger design focus on QoC elements when designing a PBF scheme. Our study demonstrates that, especially in a context where other policy tools (e.g. user fee removal as in our case) are in use to increase the utilization se of child health services, PBF schemes should be adjusted to more explicitly address shortcomings in quality rather than quantity of care.

References

1. Eichler R, Levine R: **Performance incentives for global health. Potential and pitalls**

In. Washington DC: Center for Global Development; 2009.

2. György Bèla Fritsche, Soeters Robert, Bruno. M: **Performance-Based Financing Toolkit.** Washington, DC:; 2014.
3. Skiles MP, Curtis SL, Basinga P, Angeles G, Thirumurthy H: **The effect of performance-based financing on illness, care-seeking and treatment among children: an impact evaluation in Rwanda**. *BMC Health Serv Res* 2015, **15**:375.
4. Basinga P, Gertler PJ, Binagwaho A, Soucat ALB, Sturdy J, Vermeersch CMJ: **Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation**. *The Lancet* 2011, **377**(9775):1421-1428.
5. Bonfrer I, Van de Poel E, Van Doorslaer E: **The effects of performance incentives on the utilization and quality of maternal and child care in Burundi**. *Soc Sci Med* 2014, **123**:96-104.
6. Peabody JW, Shimkhada R, Quimbo S, Solon O, Javier X, McCulloch C: **The impact of performance incentives on child health outcomes: results from a cluster randomized controlled trial in the Philippines**. *Health Policy and Planning* 2014, **29**:615–621.
7. Huntington D, Zaky HHM, Shawky S, Fattah FA, El-Hadary E: **Impact of a Service Provider Incentive Payment Scheme on Quality of Reproductive and Child-health Services in Egypt**. *J HEALTH POPUL NUTR* 2010, **28**(3):273-280.
8. Binyaruka P, Patouillard E, Powell-Jackson T, Greco G, Maestad O, Borghi J: **Effect of Paying for Performance on Utilisation, Quality, and User Costs of Health Services in Tanzania: A Controlled Before and After Study**. *PLoS One* 2015, **10**(8):e0135013.
9. Binagwaho A, Condo J, Wagner C, Ngabo F, Karema C, Kanters S, Forrest JI, Bizimana JdD: **Impact of implementing performance-based financing on childhood malnutrition in Rwanda**. *BMC Public Health* 2014, **14**(1132).
10. Rusa L, Ngirabega JdD, Janssen W, Bastelaere SV, Porignon D, Vandenbulcke W: **Performance-based financing for better quality of services in Rwandan health centres: 3-year experience**. *Trop Med Int Health* 2009, **14**(7):830-837.
11. Soeters R, Peerenboom P, Mushagalusa P, Kimanuka C: **Performance-based financing experiment improved health care in the Democratic Republic of Congo.** . *Health Affairs* 2011, **30**:1518–1527.
12. WHO: **Tracking universal health coverage: First global monitoring report. Geneva, Switzerland**. In. Geneva; 2015.
13. Witter S, Fretheim A, Kessy FL, Lindahl AK: **Paying for performance to improve the delivery of health interventions in low- and middle-income countries**. *Cochrane Database Syst Rev* 2012(2):CD007899.
14. Das A, Gopalan SS, Chandramohan D: **Effect of pay for performance to improve quality of maternal and child care in low-and middle-income countries: a systematic review**. *BMC Public Health* 2016, **16**:321.
15. Marsh AD, Muzigaba M, Diaz T, Requejo J, Jackson D, Chou D, Cresswell JA, Guthold R, Moran AC, Strong KL *et al*: **Effective coverage measurement in maternal, newborn, child, and adolescent health and nutrition: progress, future prospects, and implications for quality health systems.** . *Lancet Glob Health* 2020, **8**(5):e730-e736. doi:710.1016/S2214-1109X(1020)30104-30102.

16. Renmans D, Holvoet N, Orach CG, Criel B: **Opening the 'black box' of performance-based financing in low- and lower middle-income countries: a review of the literature**. *Health Policy Plan* 2016.
17. Lohmann J, Souares A, Tiendrebéogo J, Houlfort N, Robyn PJ, Somda SMA, Allegri MD: **Measuring health workers' motivation composition: Validation of a scale based on Self-Determination Theory in Burkina Faso**. *Human Resources for Health* 2017, **15**:33 DOI 10.1186/s12960-12017-10208-12961.
18. Savedoff WD: **Basic economics of results-based financing in health. MaineUSA: Social Insight Bath. 2010. http://www.focusintl.com/RBM082-RBF%20Economics.pdf. Accessed 26 Apr 2020**. 2010.
19. Steenland M, Robyn PJ, Compaore P, Kabore M, Tapsoba B, Zongo A, Haidara OD, Fink G: **Performance-based financing to increase utilization of maternal health services: Evidence from Burkina Faso**. *SSM - Population Health* 2017, **3**:179–184.
20. Ministère de la Santé: **Guide de mise en oeuvre du financement basé sur les résultats dans le secteur de la Santé.** In*.* Ouagadougou: Burkina Faso; 2013.
21. DeAllegri M, Lohmann J, Souares A, Robyn PJ, Schleicher M: **Responding to policy makers' evaluation needs: combining experimental and quasiexperimental approaches to estimate the impact of performance based financing in Burkina Faso. In:To be submitted.** *BMC Health Services Research* 2019, **19** 733 https://doi.org/710.1186/s12913-12019-14558-12913.
22. Service d'information du gouvernement: **Gratuité des soins pour les enfants de moins de cinq ans : Une réalité au Burkina Faso à partir du 02 avril 2016**. In*.* Ougadougou, Burkina Faso; 2016.
23. Koulidiati J-L, Nesbitt RC, Ouedraogo N, Hien H, Robyn PJ, Compaoré P, Souares A, Brenner S: **Measuring effective coverage of curative child health services in rural Burkina Faso: a cross-sectional study**. *BMJ Open* 2018, **8**:e020423. doi:020410.021136/bmjopen-022017-020423.
24. Gouws E, Bryce J, Pariyo G, Armstrong Schellenberg J, Amaral J, Habicht JP: **Measuring the quality of child health care at first-level facilities**. *Soc Sci Med* 2005, **61**(3):613-625.
25. Donabedian A: **The quality of care: How can it be assessed?** . *JAMA* 1988, **260**(12):1743-1748.
26. Koulidiati J-L, DeAllegri M, Souares A, Ouedraogo S, Hien H, Robyn PJ, Brenner S: **Factors associated with effective coverage of child health services in Burkina Faso**. *Tropical Medicine and International Health* 2018, **23**(11):1188-1199.
27. Bertrand M, Duflo E, Mullainathan S: **How much should we trust differences-in-differences estimates?.** *The Quarterly journal of economics* 2004, **119**(1):249-275.
28. Carter AV, Schnepel KT, Steigerwald DG: **Asymptotic behavior of a *t* test robust to cluster heterogeneity**. *Review of Economics and Statistics* 2017, **99**(4):698-709.
29. Imbens GW, Kolesar M: **Robust standard errors in small samples: Some practical advice**. *Review of Economics and Statistics* 2016, **98**(4):701.
30. Angrist DA, Pischke J-S: **Mostly Harmless Econometrics: An empiricist's companion.** : Princeton University Press.; 2009.
31. Cameron AC, Miller DL: **A practitioner's guide to cluster-robust inference**. *Journal of Human Resources* 2015, **50**(2):317-372.
32. Wing C, Simon K, Bello-Gomez RA: **Designing Difference in Difference Studies: Best Practices for Public Health Policy Research.** . *Annu Rev Public Health* 2018, **39**:453-469. doi:410.1146/annurev-publhealth-040617-013507.
33. Tanahashi T: **Health Service Coverage and its Evaluation.** . *Bull World Health Organ* 1978, **56**(2):295-303.
34. Kuunibe N, Lohmann J, Hillebrecht M, Nguyen H, G T, De Allegri M: **What Happens When Performance Based Financing Meets A Free Healthcare Policy? Evidence From An Interrupted**

Time-Series Analysis With Independent Controls. *Health Policy and Planning DOI: 101093/heapol/czaa062 In Press* 2020.

35. DeAllegri M, Lohmann J, Koulidiati J-L, Somé P-A: **Qualitative Research to Explain and Unpack Quantitative Findings from the Performance-Based Financing Impact Evaluation in Burkina Faso. Results report.** . In.; 2019.
36. Gergen J, Josephson E, Coe M, Ski S, Madhavan S, Bauhoff S: **Quality of Care in Performance-Based Financing: HowIt Is Incorporated in 32 Programs Across 28 Countries**. *Global Health: Science and Practice* 2017, **5**(1).
37. Ridde V, Yaogo M, Zongo S, Somé PA, Turcotte-Tremblay AM: **Twelve months of implementation of health care performance-based financing in Burkina Faso: A qualitative multiple case study**. *Int J Health Plann Mgmt* 2017:1–15.
38. Lohmann J, Koulidiati J-L, Somda SM, De Allegri M: **“It Depends on What They Experience in Each Health Facility. Some Are Satisfied, Others Are Not.” A MixedMethods Exploration of Health Workers’ Attitudes Towards Performance-Based Financing in Burkina Faso**. *International Journal of Health Policy and Management* 2020:-.
39. Ministère de la santé: **Stratégie nationale de mise en œuvre de la gratuité des soins et des services de planification familiale. Février 2019** In. Ouagadougou. Burkina Faso.
40. Falisse JB: **Beyond the traveling model? Strategic usage and hybridisation of performance-based financing (PBF) in health in Burundi.** . *Politique africaine* 2019(4): pp.83-100.
41. Nesbitt RC, Lohela TJ, Manu A, Vesel L, Okyere E, Edmond K, Owusu-Agyei S, Kirkwood BR, Gabrysch S: **Quality along the continuum: a health facility assessment of intrapartum and postnatal care in Ghana**. *PLoS One* 2013, **8**(11):e81089.
42. Larson E, Vail D, Mbaruku GM, Mbatia R, Kruk ME: **Beyond utilization: measuring effective coverage of obstetric care along the quality cascade**. *International Journal for Quality in Health Care* 2017, **29**(1):104–110.

Tables

Table 1. Study sample characteristics

	Baseline			Endline		
	Total	**PBF**	**Control**	**Total**	**PBF**	**Control**
Number of primary level health facilities	537	420	117	537	420	117
Number of health workers interviewed	1,361	1,076	285	1,769	1,410	359
Number of U5YO cases observed	2,046	1,687	359	2,649	2,070	579
Number of households surveyed	7,978	6,224	1,754	7,898	6,152	1,746
Number of U5YOs' caretakers interviewed	12,350	9,715	2,635	15,021	11,622	3,399
Number of U5YOs who reported illness	650	466	184	996	793	203

Table 2: PBF impact on effective and crude coverage

	PBF vs control					
	Unadjusted			Adjusted		
	Effect size (% points)	CI% 95$^{\theta}$	p-value	Effect size (% points)	CI% 95$^{\theta}$	p-value
Effective coverage[¥]	- 4.1	[-25.4 27.8]	0.935	-1.8	[-21,1 27.4]	0.893
Crude coverage[§]	-5.6	[-18.8 .7.7]	0.335	-3.9	[-15.5 10.8]	0.526

¥ We adjusted effective coverage with the following control variables: child sex, child age, household wealth quintile, distance between household and health facility, household dependency ratio, number of symptoms reported and Health workers seniority.

§ We adjusted crude coverage with the following control variables: child sex, child age, household wealth quintile, distance between household and health facility, household dependency ratio, number of symptoms reported,

$^{\theta}$ CI% 95 after wild bootstrapping test

Table 3: PBF impact on facility quality scores

	PBF vs control					
	Unadjusted			Adjusted$^{\theta}$		
	Effect size (score points)	CI% 95[§]	p-value	Effect size (score points)	CI% 95[§]	p-value
Management common child-hood diseases (score range 0 to 9)	-0.56	[-1.99 -0.58]	0.451	-0.57	[-1.98 -0.58]	0.463
Management severe child-hood diseases (score range 0 to 11)	-0.18	[-1.07 0.69]	0.686	-0.17	[--1.05 0.67]	0.668
General service readiness (score range 0 to 5)	0.29	[-0.14 0.72]	0.139	0.30	[-0.11 0.68]	0.133
Overall facility quality (score range 0 to 25)	-0.35	[-1.94 1.14]	0.653	-0.33	[-1.91 1.16]	0.663

[θ] We adjusted facility quality scores with the following control variables: Health worker seniority, under-five patient age

§ CI% 95 after wild bootstrapping test

Table 4: Sensitivity analysis[a] : PBF impact on crude, effective coverage, and facility quality scores

	PBF versus control		
	Adjusted		
	Effect size (% points)	CI% 95§	p-value
Effective coverage	1.5	[-21 30.9]	0.876
Crude coverage	-3.6	[-16.81 10.6]	0.378
	Effect size (score points)	CI% 95§	p-value
Management common childhood diseases (score range 0 to 9)	0.13	[-0.52 0.74]	0.688
Management severe childhood diseases (score range 0 to 11)	-0.12	[-1.02 0.95]	0.810
General service readiness (score range 0 to 5)	0.24	[-0.18 0.66]	0.257
Overall facility quality (score range 0 to 25)	0.36	[-0.98 1.36]	0.55

[a]We removed from the analysis two control districts (Barsalgho and Ziniare) where we observed a huge increase of QoC between baseline and endline.

§ CI% 95 after wild bootstrapping test