

Understanding efficiency and the effect of pay-for-performance across health facilities in Tanzania

Peter Binyaruka ¹, Laura Anselmi ²

To cite: Binyaruka P, Anselmi L. Understanding efficiency and the effect of pay-for-performance across health facilities in Tanzania. *BMJ Global Health* 2020;**5**:e002326. doi:10.1136/bmjgh-2020-002326

Handling editor Valery Ridde

Received 20 January 2020

Revised 15 April 2020

Accepted 19 April 2020

ABSTRACT

Background Ensuring efficient use and allocation of limited resources is crucial to achieving the UHC goal. Performance-based financing that provides financial incentives for health providers reaching predefined targets would be expected to enhance technical efficiency across facilities by promoting an output-oriented payment system. However, there is no study which has systematically assessed efficiency scores across facilities before and after the introduction of pay-for-performance (P4P). This paper seeks to fill this knowledge gap.

Methods We used data of P4P evaluation related to healthcare inputs (staff, equipment, medicines) and outputs (outpatient consultations and institutional deliveries) from 75 health facilities implementing P4P in Pwani region, and 75 from comparison districts in Tanzania. We measured technical efficiency using Data Envelopment Analysis and obtained efficiency scores across facilities before and after P4P scheme. We analysed which factors influence technical efficiency by regressing the efficiency scores over a number of contextual factors. We also tested the impact of P4P on efficiency through a difference-in-differences regression analysis.

Results The overall technical efficiency scores ranged between 0.40 and 0.65 for hospitals and health centres, and around 0.20 for dispensaries. Only 21% of hospitals and health centres were efficient when outpatient consultations and deliveries were considered as output, and <3% out of all facilities were efficient when outpatient consultations only were considered as outputs. Higher efficiency scores were significantly associated with the level of care (hospital and health centre) and wealthier catchment populations. Despite no evidence of P4P effect on efficiency on average, P4P might have improved efficiency marginally among public facilities.

Conclusion Most facilities were not operating at their full capacity indicating potential for improving resource usage. A better understanding of the production process at the facility level and of how different healthcare financing reforms affects efficiency is needed. Effective reforms should improve inputs, outputs but also efficiency.

INTRODUCTION

Low/middle-income countries (LMICs), especially in sub-Saharan Africa, are constrained in achieving universal health coverage (UHC) by limited resources,^{1–3} double

Key questions

What is already known?

- ▶ There is a room for increasing efficiency in resource use across health facilities in low/middle-income countries.
- ▶ Healthcare financing reforms have the potential to stimulate efficiency in resource use including those focusing on output-based financing (eg, pay-for-performance (P4P)).

What are the new findings?

- ▶ The technical efficiency score from 150 facilities in Tanzania ranged between 40% and 65% for hospitals and health centres, and lowest around 20% for dispensaries.
- ▶ Higher efficiency scores were significantly associated with the level of care provided (hospital and health centre) and wealthier catchment populations.
- ▶ There is no evidence that P4P improved efficiency on average, but might marginally improved among public facilities only.

What do the new findings imply?

- ▶ Facilities can still improve efficiency, but the service delivery processes at the health facility level and the effect of incentive schemes need to be better understood to determine how to improve efficiency.
- ▶ Large share of technically inefficient facilities suggests that most facilities (especially dispensaries) are not operating in their full capacity or significantly wasting resources, perhaps due to poor healthcare seeking culture.



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Health System, Impact Evaluation and Policy, Ifakara Health Institute, Dar es Salaam, Tanzania

²Health Organisation, Policy and Economics, Centre for Primary Care and Health Services Research, University of Manchester, Manchester, UK

Correspondence to

Dr Peter Binyaruka, Ifakara Health Institute, Dar es Salaam, Tanzania; pbinyaruka@ihi.or.tz

burden of diseases (communicable and non-communicable diseases)^{4,5} and relatively poor health system performance.^{6,7} Despite the efforts to ensure more funding are allocated to the health sector, resources are still limited and ensuring they are used efficiently and equitably is crucial to achieve UHC goals.⁸ In 2010, the World Health Report estimated about 20%–40% of health resources are being wasted globally due to inefficient and inequitable use of resources.^{1,9} Indeed, countries' health systems need not only more resources for health but also efficient use of available

resources. Efficiency can either be technical efficiency (minimum amount of resources used for a given level of output or maximum amount of output produced for a given level of resources) or allocative efficiency (how resource inputs and their prices are combined to produce a mix of different outputs).⁹

Pay-for-performance (P4P) schemes, which provide financial incentives for health workers and/or facilities reaching predefined targets, are gaining popularity in LMICs as a means of improving facility performance and health system strengthening.^{10–13} Although there is a variation across schemes, P4P incentives are typically rewarded to facilities or healthcare providers after the achievement of quantity and quality targets. These payments are used for facility improvement as well as additional bonus to staff salary.^{14 15} So, while incentivising output improvement, bonus payments to facilities can also contribute to the improvement of infrastructure and to drugs and supplies availability, among others. Some P4P schemes also offer an initial payment to facilities enrolled in the scheme to expand their capacity for service provision. Since P4P is an output-oriented financing strategy aimed at increasing the provision of targeted services, one would expect P4P to enhance efficiency in service delivery. For instance, P4P can improve allocative efficiency by incentivising an optimal mix of services that maximises health improvements, and it can improve technical efficiency by increasing the quantity and quality of services produced for a given amount of inputs.¹¹

There is a growing body of evidence evaluating the impact of P4P in LMICs. However, most of these studies have focused on assessing the effects of P4P on service coverage,^{12 16} quality of care¹⁷, equity in service delivery and resource distribution,^{18–20} and on facility inputs, such as human resources^{21–23} and medical commodities.^{24 25} Many P4P evaluations indicate mixed results, with notable increases of incentivised service delivery indicators. The current literature, however, shows relatively little attention to efficiency which combines inputs and outputs. The little available evidence shows that P4P increased provider productivity in Rwanda,²⁶ increased efficiency of faith-based facilities in Uganda,²⁷ and efficiency improved over time among P4P implementing health centres in Rwanda.²⁸ However, these studies focused on a single type of facilities (eg, faith-based or health centres only^{27 28}), used only one input and output²⁶ or relied on a small sample of facilities and assessed trend over time without comparing them with control facilities.²⁸

We estimated the technical efficiency of public and private health facilities providing primary and/or secondary care, and we assessed the impact of a P4P scheme by comparing intervention and control facilities in Tanzania. Previous evaluations of P4P in Tanzania revealed an increase in service coverage (outputs) and in the availability of some healthcare inputs in the facilities where it was implemented.^{18 24} Assessing health facilities efficiency and how it changed under P4P is important

and timely, as it adds to a limited evidence on the association between P4P and efficiency, which we suggest policy-makers should consider when making decisions with the implementation of P4P in low-resource settings like Tanzania.

Study setting

This study took place in Tanzania, a low-income country in East Africa. Most facilities (70%) in Tanzania are government owned. The public health system has a hierarchical administration and is organised in a referral structure, with dispensaries and health centres providing primary care and hospitals providing secondary and tertiary care.²⁹ The health system is decentralised, whereby local managers are given power to plan and manage resources.³⁰ Like in other LMICs, most health facilities in Tanzania are facing shortage of staff,³¹ drugs and supplies.^{32 33} Health financing is highly fragmented with many sources including general taxation (26%), donor support (41%), out-of-pocket payments (23%) and health insurance contributions (8%).^{34 35} About 32% of Tanzanians are covered by health insurance—ie, 8% as public servants mainly through National Health Insurance Fund, 23% as informal workers through Community Health Fund (CHF) and 1% from private insurance.³⁶ In 2018/2019, about 9% of total government expenditure or total budget was allocated for health (below the Abuja Declaration Target of 15%).³⁶ Tanzania has introduced various healthcare financing reforms in an effort to move toward UHC, including fiscal decentralisation through direct health facility financing,^{37 38} improved CHF³⁹ and national roll-out of P4P, also called results-based financing programme.⁴⁰

In January 2011, Tanzania introduced a P4P pilot scheme in Pwani region before national roll-out, with the aim of improving maternal and child health (MCH) services.⁴¹ All health facilities providing MCH services in the region were eligible to implement the scheme. P4P performance targets for facilities were either related to the coverage of specific services (eg, institutional delivery) or to content of care (eg, uptake of antimalarials during antenatal care) as described in more detail elsewhere.^{18 41} Health managers were also eligible to receive performance payouts based on the performance of the facilities in their district or region. Performance data were compiled by facilities and verified every 6 months (one cycle) before P4P payments. The maximum payout if all targets were fully attained was US\$820 per cycle for dispensaries, US\$3220 for health centres and US\$6790 for hospitals. P4P payments were additional to funding for operational costs and salaries which are unrelated to performance. P4P payouts at the facility-level included bonuses to staff (equivalent to 10% of their monthly salary) and funds that could be used for facility improvement or demand creation initiatives (10% of the total in hospitals and 25% in lower level facilities). District and regional managers received bonus payments of up to US\$3000 per cycle.

The impact evaluation of P4P pilot showed a significant positive effect on two out of eight incentivised service indicators: institutional delivery rate and provision of antimalarial during antenatal care.¹⁸ The evaluation also documented various positive process changes, such as increased availability of drugs and supplies, increased supportive supervision, a reduced chance of patients paying user fees and greater provider kindness during institutional delivery.^{18 24 42 43}

METHODS

Data sources

Data were derived from two rounds of a repeated cross-sectional facility survey done in 150 health facilities from 11 districts in Pwani, Morogoro and Lindi region, as part of an impact evaluation of the P4P pilot programme in Pwani region.^{18 41} In Pwani region, all 6 hospitals and 16 health centres were eligible for inclusion, and a random sample of 53 dispensaries were included in the evaluation study. An equivalent number of facilities were randomly sampled by level of care in comparison districts from Morogoro and Lindi region. As a result, an equal number of facilities in the intervention and control study arms were included—ie, 75 facilities from Pwani region (P4P arm) and 75 facilities from Morogoro and Lindi (comparison arm). Further details on the sampling strategy of facilities are presented elsewhere.⁴¹

Data collection

Data were collected through a facility survey questionnaire adapted from the World Bank Impact Evaluation Toolkit.⁴⁴ The questionnaire had three main sections: (1) basic facility characteristics and management (staffing levels, opening hours, facility management and infrastructure), (2) availability of drugs, supplies and equipment and (3) monthly service utilisation data from facility register books. Trained enumerators administered a structured facility questionnaire to the facility in-charge or any knowledgeable health worker at each facility. The enumerators also measured inputs and contextual factors during data collection, while outputs (monthly utilisation data) were extracted from facility register books for 12 months prior to the interview. Data were collected in two rounds: January and February 2012 and 13 months later, reflecting before and after the introduction of P4P. Each round surveyed 150 facilities including 12 hospitals, 32 health centres and 106 dispensaries. Data collected have been used previously.^{19 24 42 43}

Input, output and contextual factor variables

This study used three healthcare inputs and two outputs to estimate efficiency scores across health facilities (table 1). Labour (clinical staff by cadre), capital (proxied by the number of beds) and medical commodities were included as key healthcare inputs to the production process for health facilities. Healthcare outputs included volume of health services utilisation, as intermediate outputs, since health status are difficult and costly to

measure.⁴⁵ As data were incomplete for some facilities, we imputed the number of maternity beds in seven facilities that missed in 2012 with the number of beds in 2013, assuming nothing has changed over time; and similarly, for seven facilities that provided inpatient services, but recorded zero bed in 2012. We replaced the missing values of outputs in 20 facilities with those from the previous or next year. Monthly mean numbers of outpatient visits and normal deliveries were computed annually, by summing consultations over the year and divide by 12 months. We also used data on a set of contextual indicators: facility ownership status, facility level of care, location (rural/urban), catchment population characteristics, availability of outreach services, availability of CHF insurance, distance from district headquarter and frequency of external supervision. The choice of inputs, outputs and contextual factors was guided by contextual consideration and previous efficiency studies in low-income settings, conditional to data availability.⁴⁶

Measuring efficiency: data envelopment analysis (DEA)

Efficiency analysis methods are based on the estimation of a production frontier, which represents the maximum output that a decision-making unit, health facilities in this study, can produce for a given input combination. Efficiency scores ranging from 0, completely inefficient, to 1, efficient, indicate for each decision-making unit the proportion of output produced with respect to a hypothetical identical one producing on the frontier (ie, efficient). Generally, two types of methods are used to estimate the frontier: stochastic frontier analysis, based on parametric techniques, and DEA, based on non-parametric techniques.^{9 45 47–49} Both have been widely used in the literature and present advantages and disadvantages. The choice ultimately relies on the analyst judgement based on the research question, study setting and data available.^{9 45–48} Thus, we used a DEA because of methodological and practical advantages. DEA accommodates multiple inputs and multiple outputs and it is a data-driven approach, which allow us to avoid specific assumptions about the functional form of the production function and the distribution of the error term. As such, DEA is also used for 'benchmarking' where efficient units may not necessarily form a production frontier, but instead a 'best-practice frontier'.⁴⁵

In this study, we assumed alternatively both constant returns to scale (CRS), that is, all units are operating at their optimal scale,⁵⁰ and variable returns to scale (VRS).⁵¹ Through VRS assumption, we were able to decompose the overall technical efficiency scores into *scale efficiency* and *pure technical efficiency*. We considered an output-oriented approach due to the hierarchical structure of the health system and the centralised procurement process, such that most public health facilities in low-income settings have fixed set of inputs at any given time with limited influence on their inputs. However, health facilities have the autonomy to use limited or fixed resources in an innovative way to maximise outputs. We estimated technical

Table 1 Definition and measurement of health facility input, output and contextual variables

Variables	Measurement
Facility inputs	
Clinical staff	Total number of doctors/clinicians, nurses/midwives and paramedics
Beds	Total number of maternity beds during the survey
Medical commodities	Total number of different drugs, vaccines, medical supplies and medical equipment available in the health facility during the 3 months prior to the survey <ul style="list-style-type: none"> ▶ Drugs and vaccine include 37 items: antimalarials (3), 6 antibiotics (6), antihypertensives (5), anti-diarrhoeals (2), oxytocics (3), ARVs (7), vaccines (5), vitamin A (1) and family planning medicines (5). ▶ Medical supplies include 11 items: male condom and female condom, disinfectant/iodine, sterile latex gloves, delivery kit, cotton wool, partograph, sutures, catheters, oxygen supply and gas supply (for vaccines). ▶ Equipment includes 19 items: time/watch, autoclave, delivery table, vacuum extractor, examination lamp/torch, blood pressure machine, stainless steel bowls, neonatal Ambu bag, incubator, cord ligatus/clamp, infant laryngoscope, mucus suction, weighing scale, MUAC measuring tape, test kit, delivery kit, reagents for test kit, thermometer and stethoscope.
Facility outputs	
Outpatient visits	Monthly mean number of outpatient (preventive and curative) visits including visits of children under 5 years, over 5 years and antenatal care.
Deliveries	Monthly mean number of facility-based normal deliveries.
Contextual variables	
Facility level of care	Binary variables: three levels of care (hospitals, health centres and dispensaries).
Facility ownership	1 for public-owned facility and 0 for non-public.
Facility location	1 for facility in urban district and 0 for rural district.
Population socioeconomic status	1 for non-poor population and 0 for poor catchment population.
Outreach services	Average number of outreaches in last 90 days.
Community Health Fund (CHF)	1 if facility has insurance scheme (CHF) and 0 no health insurance.
Distance (km)	Mean kilometre from facility to district headquarter.
Supportive supervision	1 if external supervision in the last 30 days and 0 for no external supervision in the last 30 days.

The items for medical commodities have similarly been used previously (Binyaruka and Borghi²⁴). ARVs, antiretrovirals; MUAC, mid-upper arm circumference.

efficiency for hospitals, health centres and dispensaries separately in 2012 and 2013, as inputs, outputs and healthcare delivery may be very different in each level of care. Allocative efficiency could not be considered in this paper due to the lack of reliable data on input prices. Moreover, due to the centralised procurement, the prices of staff and large equipment would not vary across facilities. We estimated DEA scores by using two models with different input and output combinations to test the relative importance of each input or output. Model 1 included four inputs (staff level, number of drugs and vaccines, medical supplies and equipment) and one output (outpatient visits) for all facilities. Model 2 included five inputs (staff level, number of drugs and vaccines, medical supplies, equipment and maternity beds) and two outputs (outpatient visits and normal deliveries) for hospital and health centres only. Dispensaries were not included in model 2 as they are not equipped to provide institutional deliveries. We estimated efficiency scores in each model by benchmarking health facilities

of the same type. We assumed CRS first then VRS. All health facilities with efficiency score of 1 were considered efficient, and thus we computed the percentage of efficient health facilities and we examined how that changes over time. Efficiency scores were computed in Stata by applying the DEA user-written command, with alternatively CRS and VRS, as required by each model, and with output orientation. Details are presented in online supplementary appendix.

Determinants of efficiency: regression analyses

To identify the determinants of technical efficiency, we regressed the efficiency scores over observed contextual factors. The efficiency scores are censored (0 to 1) and the correlation pattern among DEA efficiency scores is typically complex and unknown. We therefore used the approach developed by Simar and Wilson⁵² and recently applied by Moreno-Serra *et al.*,⁵³ as it accounts for bounded dependent variable, corrects the SEs by simulating the

unknown error correlation among efficiency scores and calculates bootstrapped SEs.⁵⁴

We pooled the efficiency scores estimated separately for dispensaries, health centres and hospitals in 2012 and in 2013 and we regressed them on a set of contextual and environmental factors (listed in table 1). We also included binary indicators to control for type of health facility, year and district. We used Stata *simarwilson* with externally estimated efficiency scores and 2000 replications. In model 1 we included dispensaries, health centres and hospitals, and in model 2 we included only health centre and hospitals. Details are presented in online supplementary appendix.

Effect of P4P on efficiency: difference-in-differences analysis

To test whether P4P affects efficiency scores, we first compared descriptively the changes in efficiency scores over time for P4P and non-P4P facilities. We then applied a linear difference-in-differences regression analysis based on a controlled before and after design.

We regressed efficiency scores pooled across health facilities and years over a dummy variable, taking the value 1 if a facility is exposed to P4P in 2013 and 0 if not. We controlled for time invariant determinants including facility fixed effects and for time-specific effects including year fixed effects. As average programme effect may mask important heterogeneous effects,⁵⁵ we further attempted to assess the heterogeneous effect of P4P by facility type, by including in the regression the interaction term between P4P in 2013 and the category considered for heterogeneity, along with the category itself. Details are presented in online supplementary appendix. As in the analysis of the determinants of efficiency, we applied the Simar and Wilson⁵² approach.

The effect of P4P on efficiency can be interpreted as causal on the assumption that trends in efficiency scores across facilities in intervention and comparison sites were parallel before the start of the programme. We were unable to formally test this assumption due to the lack of multiple observations of inputs and outputs, and hence efficiency scores, before the introduction of P4P. However, trends in facility outputs (eg, service utilisation rates) were parallel prior to the intervention.^{18 43} As trends in inputs were also likely to be parallel, due to centralised planning and procurement, we argue that the evidence lends some support to the assumption. All analyses were performed using STATA V.15.

Patient and public involvement

We used secondary data from the facility survey of the evaluation of P4P in Tanzania. Our data do not involve patient, but our findings will be shared with health system stakeholders.

RESULTS

Descriptive statistics

Table 2 presents descriptive statistics of baseline and endline characteristics of facilities including the levels of

inputs and outputs across study arms. The baseline levels of outputs and inputs were generally similar between intervention and comparison arms, although intervention facilities had marginally higher level of medical staff (especially paramedics) and lower availability of drugs and vaccines than comparison facilities. The other facility characteristics were fairly balanced between arms. However, most of the intervention facilities had CHF scheme and served poorer population than their counterpart facilities in comparison arm. Looking at changes over time, the availability of beds, of drugs and vaccines (in the comparison group) and of medical supplies slightly decreased. Outpatient consultations decreased in all facilities, while deliveries increased in P4P facilities and decreased in the control ones.

Efficiency scores

The technical efficiency scores calculated assuming CRS varied by facility type, both in model 1, which included four inputs and one output (outpatient visits) and in model 2, which included an additional input, maternity beds, and an additional output, normal deliveries. The average overall technical efficiency for hospitals and health centres was 0.46 and 0.40 in model 1, and 0.65 and 0.60 in model 2, respectively (table 3 and online supplementary appendix table 1). Dispensaries had relatively low average efficiency scores of 0.20, which contributed to lower the average efficiency scores across all health facilities in model 1. Model 1 resulted into an overall relatively smaller percentage of technically efficient facilities (2.4%) compared with model 2 (21.3%) (table 3). Since in model 1, the average technical efficiency score was 0.46, 0.40 and 0.20 for hospitals, health centres and dispensaries respectively, this suggests that, on average, they could respectively increase their outpatient visits by about 54%, 60% and 80% without reducing inputs.

Over time, the average technical efficiency scores of health centres and dispensaries decreased, as measured in model 1, while hospitals efficiency scores increased from 0.45 to 0.47 (table 3). Consistently, the proportion of efficient health centres (6.3%) and dispensaries (0.9%) in 2012 reduced to 0% for both in 2013. The proportion of efficient hospitals remained efficient throughout (16.7%). In model 2, overall technical efficiency score and share of efficient facilities slightly improved over time for hospitals and health centres. Furthermore, the overall technical efficiency scores were estimated assuming VRS assumption, in order to allow a decomposition into *pure technical efficiency* and *scale efficiency*. The average *pure technical efficiency* and *scale efficiency* remained almost similar in both models from 2012 to 2013, except for a slight increase for hospitals and decrease for health centres over time (online supplementary appendix table 2).

Determinants of efficiency

Table 4 presents results on factors associated with overall technical efficiency. Efficiency scores estimated

Table 2 Characteristics of health facilities over time by P4P arms

Characteristics	Description/units	All (n=300)	Baseline intervention (n=75)	Baseline comparison (n=75)	Baseline difference (p value)	Endline intervention (n=75)	Endline comparison (n=75)
Panel A: facility inputs							
Total medical staff	Total number, mean (SD)	15.8 (31.6)	19.7 (40.8)	11.1 (15.0)	8.6 (0.088)	20.6 (42.7)	11.8 (15.7)
Clinicians	Total number, mean (SD)	3.2 (5.3)	3.5 (6.2)	2.5 (3.2)	1.0 (0.213)	3.9 (7.1)	2.7 (3.4)
Nurses	Total number, mean (SD)	5.4 (14.9)	6.4 (19.5)	4.2 (7.4)	2.2 (0.354)	6.8 (19.9)	4.4 (7.5)
Paramedics	Total number, mean (SD)	7.2(14.2)	9.8(18.4)	4.4(6.3)	5.4 (0.018)	10 (19.2)	4.7 (6.5)
Maternity beds	Total number, mean (SD)*	5.9 (6.7)	7.9 (7.9)	4.4 (5.1)	1.0 (0.159)	7.4 (8.8)	3.9 (2.6)
Availability of drugs and vaccines	Mean no of items available in last 3 months (SD)	20.6 (5.7)	20.1 (5.5)	21.7 (4.5)	-1.6 (0.053)	21.1 (6.0)	19.6 (6.4)
Availability of medical supplies	Mean no of items available in last 3 months (SD)	7.5(2.0)	7.5(2.3)	7.9(1.9)	-0.4 (0.209)	7.3 (1.7)	7.3 (1.9)
Availability of equipment	Mean no of items available in last 3 months (SD)	15.4 (4.3)	12.2 (3.5)	11.4 (3.2)	0.8 (0.164)	19.0 (0.3)	19.0 (0.1)
Panel B: facility outputs							
Monthly outpatient visits	Monthly average of outpatient consultations	602.4 (894.0)	633.1 (1204.7)	587.7 (600.9)	45.3 (0.772)	605.9 (1048.0)	583.5 (574.1)
Monthly normal deliveries	Monthly average of normal deliveries*	49.9 (76.7)	60.7 (93.6)	32.3 (47.9)	8.2 (0.298)	68.8 (94.6)	31.1 (43.8)
Panel C: environment/contextual characteristics							
Facility level of care	1 for dispensaries (%)	0.7 (0.5)	70.7	70.7	0.0 (1.000)	0.7 (0.5)	0.7 (0.5)
Facility ownership	1 for public-owned facility (%)	0.8 (0.4)	84.0	82.7	1.3 (0.828)	0.8 (0.4)	0.8 (0.4)
Facility location	1 for facility in urban district (%)	0.2 (0.4)	21.1	16.0	5.3 (0.405)	0.2 (0.4)	0.2 (0.4)
Population socioeconomic status	1 for non-poor population (%)	0.5 (0.5)	36.0	64.0	-28.0 (0.001)	0.4 (0.5)	0.6 (0.5)
Outreach services	Total number in last 90 days, mean (SD)	2.1 (2.9)	1.9 (2.8)	2.0 (2.3)	-0.1 (0.871)	2.5 (3.3)	2.1 (3.1)
Community Health Fund (CHF)	1 if facility has insurance scheme (CHF) (%)	0.7 (0.5)	78.7	54.7	24.0 (0.002)	0.8 (0.4)	0.6 (0.5)
Distance (km)	Mean kilometres from facility to district headquarter	59.9 (40.2)	56.9 (38.8)	62.9 (41.8)	-6.0 (0.367)	56.9 (38.8)	62.9 (41.8)
Supportive supervision	1 if external supervision in last 30 days (%)	0.5 (0.5)	57.7	49.3	8.4 (0.312)	0.6 (0.5)	0.4 (0.5)

Drugs and vaccines (include 37 items); medical supplies (11); equipment (19); reference category in parentheses: public (vs non-public), dispensary (vs health centre and hospital), urban (vs rural district), non-poor population (poor population); two quantiles (high and low) for socioeconomic status of the facility's catchment population.

*Statistics on beds and deliveries available for hospitals (n=12) and health centres (n=32) only. P4P, pay-for-performance.

Table 3 Summary of overall technical efficiency scores and share of efficient facilities

Facility type	Model 1		Model 2	
	Mean (SD)	% efficient facilities	Mean (SD)	% efficient facilities
<i>Panel 1: Years 2012 and 2013</i>				
Level of care				
Hospitals (n=24)	0.46 (0.33)	16.7	0.64 (0.34)	33.3
Health centres (n=64)	0.40 (0.26)	3.1	0.60 (0.28)	17.9
Dispensaries (n=206)	0.20 (0.14)	0.5	–	–
All facilities (n=294)	0.26 (0.22)	2.4	0.62 (0.30)	21.3
<i>Panel 2: Years 2012</i>				
Level of care				
Hospitals (n=24)	0.45 (0.33)	16.7	0.64 (0.33)	25
Health centres (n=64)	0.42 (0.29)	6.3	0.60 (0.28)	14
Dispensaries (n=206)	0.21 (0.15)	0.9	–	–
All facilities (n=294)	0.27 (0.23)	3.5	0.60 (0.30)	20
<i>Panel 3: Years 2013</i>				
Level of care				
Hospitals (n=24)	0.47 (0.34)	16.7	0.65 (0.33)	33.3
Health centres (n=64)	0.38 (0.23)	0	0.61 (0.29)	21.1
Dispensaries (n=206)	0.19 (0.12)	0	–	–
All facilities (n=294)	0.25 (0.20)	1.3	0.61 (0.12)	25

Model 1: Inputs (total staff, drugs and vaccines, medical supplies, equipment); output (outpatient visits).

Model 2: Inputs (total staff, drugs and vaccines, medical supplies, equipment, beds); outputs (outpatient visits and normal deliveries).

in model 1 were significantly higher by 0.281 (CI: 0.118 to 0.407) for hospitals and by 0.294 (CI: 0.180 to 0.380) for health centres, compared with dispensaries.

Nevertheless, efficiency scores estimated in model 2 were significantly higher only for facilities with a relatively wealthier catchment population (0.266; CI: 0.036

Table 4 Determinants of efficiency

Variables	Model 1			Model 2		
	Coefficient	95% CI	P value	Coefficient	95% CI	P value
Hospital ***	0.281	0.118 to 0.407	0.000	–0.001	–0.172 to 0.181	0.989
Health centre ***	0.294	0.180 to 0.380	0.000			
Public facility	0.012	–0.106 to 0.125	0.834	0.060	–0.191 to 0.309	0.626
Urban facility	0.087	–0.099 to 0.258	0.332	–0.072	–0.376 to 0.225	0.637
Non-poor catchment population **	–0.067	–0.169 to 0.036	0.194	0.266	0.036 to 0.483	0.019
Number of outreach in 90 days	0.007	–0.005 to 0.019	0.229	0.012	–0.007 to 0.029	0.189
Health insurance (CHF) availability	0.024	–0.079 to 0.136	0.648	0.130	–0.028 to 0.301	0.126
Distance from facility to district HQ	0.000	–0.001 to 0.001	0.707	0.002	–0.001 to 0.005	0.111
External supervision within 30 days	–0.008	–0.083 to 0.064	0.821	0.089	–0.067 to 0.232	0.243
Time year (year 2013)	–0.003	–0.075 to 0.068	0.930	0.042	–0.079 to 0.164	0.496
Constant	0.064	–0.125 to 0.280	0.528	–0.067	–0.501 to 0.422	0.774
District controls	Yes			Yes		
Observations (n)	257			74		

Model 1 include all facilities over time and include four inputs and one output (outpatient visits). Model 2 excluded dispensaries over time and include five inputs and multiple outputs (outpatient visits and deliveries).

***Denotes significance at 1%, ** at 5% and * at 10% level.

CHF, Community Health Fund; HQ, headquarter.

Table 5 Estimated average and heterogeneous P4P effect

	Model 1			Model 2		
	Coefficient	95% CI	P value	Coefficient	95% CI	P value
<i>Average effect</i>						
P4P×Year 2013	−0.024	−0.058 to 0.010	0.175	0.044	−0.039 to 0.125	0.307
Year 2012	0.004	−0.019 to 0.029	0.741	0.013	−0.050 to 0.075	0.680
<i>Heterogeneous effects</i>						
Urban health facility						
P4P×Year 2013×urban	−0.002	−0.062 to 0.057	0.938	0.100	−0.165 to 0.386	0.469
P4P×Year 2013	−0.023	−0.059 to 0.013	0.201	0.039	−0.049 to 0.128	0.375
Urban	0.160	0.015 to 0.301	0.027	−0.328	−0.809 to −0.129	0.063
Year 2013	0.004	−0.020 to 0.028	0.732	0.013	−0.049 to 0.074	0.681
Public health facility						
P4P×Year 2013×public	0.060	−0.013 to 0.126	0.088	0.001	−0.174 to 0.170	0.987
P4P×Year 2013	−0.076	−0.140 to −0.003	0.028	0.043	−0.121 to 0.211	0.614
Public	0.479	0.289 to 0.650	0.000	0.267	0.075 to 0.789	0.174
Year 2013	0.004	−0.018 to 0.029	0.737	0.013	−0.046 to 0.074	0.673
Dispensary						
P4P×Year 2013×dispensary	−0.022	−0.072 to 0.032	0.427			
P4P×Year 2013	−0.009	−0.059 to 0.040	0.722			
Dispensary	−0.147	−0.288 to −0.001	0.044			
Year 2013	0.004	−0.019 to 0.027	0.739			
Low socioeconomic status						
P4P×Year 2013×low SES	0.013	−0.035 to 0.062	0.595	0.080	−0.034 to 0.199	0.176
P4P×Year 2013	−0.032	−0.079 to 0.018	0.180	0.009	−0.084 to 0.104	0.852
Low SES	−0.159	−0.300 to −0.021	0.024	−0.457	−0.867 to −0.248	0.009
Year 2013	0.004	−0.020 to 0.028	0.741	0.013	−0.051 to 0.075	0.672
Observation (n)	294			80		

Model 1 include all facilities over time and include four inputs and one output (outpatient visits); model 2 excluded dispensaries over time and include five inputs and multiple outputs (outpatient visits and deliveries); all regressions include facility fixed effects. P4P, pay-for-performance; SES, socioeconomic status.

to 0.483). These variations were observed over and above variations across districts. The other variables were not significantly associated with the efficiency scores.

Effect of P4P on efficiency

Results from the difference-in-differences regression presented in table 5 show that P4P was negatively but not significantly associated with technical efficiency in model 1 (−0.024; CI: −0.058 to 0.010) and positively but not significantly associated with efficiency in model 2 (0.044; CI: −0.039 to 0.125). P4P might have improved marginally the efficiency scores in public facilities by 0.060 (CI: −0.013 to 0.126) while it reduced it in non-public facilities by −0.076 (CI: −0.140 to −0.003) in Model 1. There was a positive but not significant increase of efficiency in both types of facilities in model 2. There was no evidence of significant differential effect of P4P for urban facilities, dispensaries or facilities with a poor catchment population in terms of socioeconomic status.

DISCUSSION

This study examined the effect of P4P on efficiency across health facilities in Tanzania. To the best of our knowledge, this is the first study to use multiple inputs and outputs to measure efficiency and the effect of P4P on efficiency by comparing P4P and non-P4P health facilities and by differentiating by facility characteristics (ie, level of care and ownership). We found overall technical efficiency scores ranging between 0.40 and 0.65 for hospitals and health centres, and lowest around 0.20 for dispensaries. Only 21% of hospitals and health centres were technically efficient when two outputs were considered, and less than 3% of all facilities (including dispensaries) were efficient when only one output was considered. In terms of determinants, higher efficiency scores were significantly associated with the level of care (hospital and health centre), and with wealthier catchment populations for hospitals and health centres only. There were also statistically significant variations in facilities' efficiency across

districts, but not over time. While there was no evidence of P4P effect on efficiency on average, P4P might have improved marginally the efficiency of public facilities.

The finding that majority of facilities were technically inefficient suggest that most health facilities (especially dispensaries) are not operating in their full capacity or are not on their efficiency frontier (ie, significant resources are being wasted). These inefficient facilities are either using inputs unnecessarily for production, which is unlikely given the low level of inputs available, or could increase service delivery with the available resources. In fact, the low technical efficiency observed is consistent with findings from other studies in sub-Saharan Africa,^{56–62} reporting efficiency scores typically between 0.40 and 0.75 with less than 50% of facilities being efficient. Few studies report efficiency scores above 0.75.^{63–66} While most studies in the region focused either on hospitals or health centres, we included all facility types, which, especially after including dispensaries, led to lower overall efficiency scores.

Higher levels of care (eg, hospitals and health centres) and for those with wealthier catchment population were significantly associated with higher efficiency levels in Tanzania. Low demand of services at facilities providing only basic primary healthcare services⁶⁴ and patients bypassing primary healthcare facilities^{67–69} may explain the association between higher level of care and higher efficiency. Indeed, more outreach visits were also associated, although not statistically significantly, with higher efficiency scores, which reflects the potential of outreach services to improve demand and utilisation of health services.^{70 71} The efficiency scores did not differ significantly between facilities in rural and urban districts, which is consistent with findings of a study from Zambia,⁵⁷ but not with findings from Ghana.⁵⁶ Characteristics of the catchment population can affect efficiency. Our study revealed that wealthier catchment populations were associated with higher efficiency levels for hospitals and health centres, when institutional deliveries were included in the output. However, according to studies in Ethiopia larger catchment populations were associated with higher efficiency levels.^{66 72} Although we found no variation in efficiency across ownership status, other studies have found either public facilities⁵⁶ or private facilities⁷³ to be more efficient than their counterparts in LMICs. In addition, a study that re-estimated country level efficiency on WHO panel data⁴⁹ revealed that public facilities could potentially be more efficient in health service delivery than their counterparts. Variations in efficiency across districts suggest that there may be district level factors still unexplored affecting efficiency, beyond rurality and catchment population characteristics.⁷⁴

The implementation of the P4P programme in Tanzania neither enhanced nor reduced efficiency across health facilities. The average efficiency scores remained almost unchanged over time. However, there was a slight increase in efficiency scores in intervention facilities and a slight decrease in control facilities, which may have

neutralised the effect. The lack of effect might also be due to the evaluation time frame of 1 year, which might be shorter to observe an effect on efficiency. Our finding of no P4P effect is in contrast to other three studies in SSA assessing P4P and efficiency. For example, Gertler and Vermeersch,²⁶ considering only one input and output, found the productivity of health workers on prenatal care improved in Rwanda after the introduction of P4P. Another study in Rwanda reported some changes in efficiency among P4P health centres;²⁸ however, they used a relatively small sample of 26 health centres and simply compared changes in efficiency over time without comparison facilities. In Uganda, evidence shows that efficiency increased among facilities implementing P4P,²⁷ but the analysis was limited to faith-based facilities only. To the best of our knowledge, ours is the only study which systematically used comparison facilities and combined facilities of different levels of care and ownership status to identify the effect of P4P on efficiency. Since the use of different set of variables generates different conclusions, we suggest that future studies should use multiple inputs and outputs and different facility type when assessing efficiency especially in the context of healthcare financing reforms.

Despite the lack of average effect of P4P on efficiency, we identified that P4P might have increased marginally the efficiency scores among public facilities and reduced efficiency in non-public facilities. This finding show that important heterogeneous effects may often be hidden when focusing on average effect only. These results may be driven by increased demand for service in public rather than non-public facilities. Public facilities in Tanzania are better able to respond to incentives as they can offer free MCH services (under the fee exemption policy) and have more financial autonomy than non-public providers like faith-based facilities.⁴² It is important to note that while we assess the effect that P4P may have on the efficiency with which health facilities operate, results do not say anything about cost-effectiveness of P4P itself, as the many costs of this reform are not accounted for.^{75 76}

Our study revealed that only less than a quarter of facilities were efficient. This indicates that only few facilities were more successful in converting inputs into outputs than others, despite being in a similar resource-constrained context. Additionally, it implies that the majority of facilities (especially primary healthcare facilities) can still increase their outputs with available levels of inputs to operate at the efficiency frontier. Facilities wishing to increase their level of outputs and eventually efficiency, may possibly need to increase the demand of health services. This can be achieved through different strategies including the expansion of health insurance coverage or voucher schemes which provide financial protection and reduce access barriers,¹⁷⁷ as well as through strategies aiming at improving quality of care, which may not only attract clients but also reduce the practice of bypassing primary healthcare facilities.⁷⁸ More targeted efforts and interventions toward some facilities and

location are also needed,^{79 80} since the inefficiencies were more prevalent among primary health facilities which are largely serving poorer populations. Furthermore, there is a need for more research to assess efficiency of primary healthcare facilities in LMICs. These assessments are currently very limited,⁴⁶ but primary healthcare facilities are the first point of contact and serves the majority of the population.^{81 82} Assessing the role of local health administrations, which in similar settings often manage the distribution of resources across facilities, would also be important to understand where improvements can be made.⁷⁴ Moreover, we encourage more research to assess whether performance-based provider payments (eg, P4P) can improve efficiency across facilities, as countries continue to implement health financing reforms toward UHC. In LMICs where staff and drugs are in shortage, it is important to improve efficiency by increasing output while maintaining or increasing inputs. Finally, we encourage the evaluation of reforms such as P4P in a comprehensive manner, in terms of changes in inputs, outputs and efficiency at the provider level, but also cost-effectiveness and efficiency of the reform itself.⁸³

This study has some limitations. First, our focus has been on technical efficiency and scale efficiency, as opposite to allocative efficiency due to the lack of data on input prices. However, this would have not been relevant in this study given the context of centralised purchasing and allocation of healthcare inputs. Second, as in the majority of efficiency studies in similar settings,⁴⁶ the outputs considered were intermediate outcomes, which are simply proxy measures of health outcomes, so that the interpretation of results relies on the assumption that reduction in service delivery was not driven by improvements in health. We were limited in the measurement of service delivery (not including inpatient use nor outreach activities) and recurrent expenditure as input especially in facilities benefiting P4P bonus. Moreover, quality is not explicitly accounted for. Third, because our study relied on DEA due to its advantages over other methods, it reflects only relative efficiency as opposed to overall performance. Thus, a high DEA score does not necessarily mean facilities are well managed with limited scope for improvement,⁶⁴ as absolute efficiency may be low. Fourth, although the assumption of parallel trends for facility outcomes prior to the intervention was confirmed, we were unable to formally test such assumption for healthcare inputs and efficiency itself due to the lack of historical data on inputs.

CONCLUSION

This study revealed inefficiencies in service outputs with available resources across facilities, and even more prevalent among primary healthcare facilities. Inefficiencies were also less likely among health facilities which were serving wealthier populations. To address the inefficiencies, efforts are needed especially toward increasing service outputs (healthcare demand) with available inputs. P4P programme had no effect on efficiency across

facilities on average, but might have marginally improved efficiency among public facilities than their counterparts, perhaps due to contextual setting. These findings still suggest a need for further assessment of the long-term effect. Although the inadequate resources are important constraint among health facilities in LMICs, the efficiency in resource utilisation is another challenge that needs attention from policy-makers, health managers and researchers.

Twitter Peter Binyaruka @peter_binyaruka

Acknowledgements We would like to thank all healthcare providers, health managers, and national officers that supported either the overall evaluation of P4P at the facility level, or in the planning of the evaluation. We also thank the whole P4P evaluation research team, including data collectors and field coordinators.

Contributors PB conceptualised this substudy and oversaw data collection. PB wrote the first draft of the manuscript. LA contributed to the conceptualisation of the study and together with PB involved in data analysis, interpretation, presentation and revision of the manuscript. All authors read and approved the final manuscript.

Funding The Government of Norway funded the data collection for the programme evaluation that was used in this paper (grant numbers: TAN-3108 and TAN 13/0005. <http://www.norad.no/en/>). The funding partially supported the first author during data analysis and writing of this paper. The UKRI University of Manchester (UoM) Research England Global Challenges Research Fund (GCRF) QA Allocation partially funded the first author and funded a visiting period to start the conceptualisation and data analysis of this paper. The funding bodies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The evaluation study received ethical approval from the Ifakara Health Institute institutional review board (approval number: 1B11RB/38) and the ethics committee of the London School of Hygiene & Tropical Medicine. Study participants provided written consent to participate in this study, requiring them to sign a written consent form that was read out to them by the interviewers. This consent form was reviewed and approved by the ethics committees prior to the start of the research.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The data have been uploaded into a data repository. The DOI URL for the dataset is <https://zenodo.org/record/21709#>.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Peter Binyaruka <http://orcid.org/0000-0002-1892-7985>

Laura Anselmi <http://orcid.org/0000-0002-2499-7656>

REFERENCES

- 1 World Health Organization. *The world health report: health systems financing: the path to universal coverage*. Geneva, Switzerland: World Health Organization, 2010.
- 2 Sachs JD. Achieving universal health coverage in low-income settings. *Lancet* 2012;380:944–7.
- 3 Wiysonge CS, Paulsen E, Lewin S, *et al*. Financial arrangements for health systems in low-income countries: an overview of systematic reviews. *Cochrane Database Syst Rev* 2017;9:CD011084.
- 4 Bygbjerg IC. Double burden of noncommunicable and infectious diseases in developing countries. *Science* 2012;337:1499–501.

- 5 Kushitor MK, Boatemaa S. The double burden of disease and the challenge of health access: evidence from access, bottlenecks, cost and equity facility survey in Ghana. *PLoS One* 2018;13:e0194677.
- 6 Bitton A, Fifield J, Ratcliffe H, et al. Primary healthcare system performance in low-income and middle-income countries: a scoping review of the evidence from 2010 to 2017. *BMJ Glob Health* 2019;4:e001551.
- 7 Tashobya CK, da Silveira VC, Ssengooba F, et al. Health systems performance assessment in low-income countries: learning from international experiences. *Glob Health* 2014;10:5.
- 8 Kutzin J. Health financing for universal coverage and health system performance: concepts and implications for policy. *Bull World Health Organ* 2013;91:602–11.
- 9 Cylus J, Papanicolas I, Smith PC. *Health system efficiency: how to make measurement matter for policy and management*. Copenhagen, Denmark: European Observatory on Health Systems and Policies, WHO, 2016.
- 10 Witter S, Toonen J, Meessen B, et al. Performance-based financing as a health system reform: mapping the key dimensions for monitoring and evaluation. *BMC Health Serv Res* 2013;13:367.
- 11 Meessen B, Soucat A, Sekabaraga C. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform? *Bull World Health Organ* 2011;89:153–6.
- 12 Renmans D, Holvoet N, Orach CG, et al. Opening the 'black box' of performance-based financing in low- and lower middle-income countries: a review of the literature. *Health Policy Plan* 2016;31:1297–309.
- 13 Paul E, Albert L, Bisala Badibanga N'Sambuka, et al. Performance-based financing in low-income and middle-income countries: isn't it time for a rethink? *BMJ Glob Health* 2018;3:e000664.
- 14 Renmans D, Holvoet N, Criel B, et al. Performance-based financing: the same is different. *Health Policy Plan* 2017;32:860–8.
- 15 Fritsche G, Soeters R, Meessen B, et al. *Performance-based financing toolkit*. Washington, DC: The World Bank, 2014.
- 16 Witter S, Fretheim A, Kessy FL, et al. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane Database Syst Rev* 2012;2:CD007899.
- 17 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.
- 18 Binyaruka P, Patouillard E, Powell-Jackson T, et al. 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:e0135013.
- 19 Binyaruka P, Robberstad B, Torsvik G, et al. Does payment for performance increase performance inequalities across health providers? A case study of Tanzania. *Health Policy Plan* 2018;33:1026–36.
- 20 Lannes L, Meessen B, Soucat A, et al. Can performance-based financing help reaching the poor with maternal and child health services? The experience of rural Rwanda. *Int J Health Plann Manage* 2016;31:309–48.
- 21 Shen GC, Nguyen HTH, Das A, et al. Incentives to change: effects of performance-based financing on health workers in Zambia. *Hum Resour Health* 2017;15:20.
- 22 Bhatnagar A, George AS. Motivating health workers up to a limit: partial effects of performance-based financing on working environments in Nigeria. *Health Policy Plan* 2016;31:868–77.
- 23 Paul F. *A qualitative study on health worker motivation and the Rwandan performance based finance initiative in district hospitals*. London: London School of Economics, 2009: 44. 1–35.
- 24 Binyaruka P, Borghi J. Improving quality of care through payment for performance: examining effects on the availability and stock-out of essential medical commodities in Tanzania. *Trop Med Int Health* 2017;22:92–102.
- 25 Brenner S, Wilhelm D, Lohmann J, et al. Implementation research to improve quality of maternal and newborn health care, Malawi. *Bull World Health Organ* 2017;95:491–502.
- 26 Gertler P, Vermeersch C. *Using performance incentives to improve health outcomes*. NBER working paper no 19046, 2013.
- 27 Duchoslav J, Cecchi F. Do incentives matter when working for God? The impact of Performance-based financing on faith-based healthcare in Uganda. *World Dev* 2019;113:309–19.
- 28 Zeng W, Rwiyerika AK, Amico PR, et al. Efficiency of HIV/AIDS health centers and effect of community-based health insurance and Performance-based financing on HIV/AIDS service delivery in Rwanda. *Am J Trop Med Hyg* 2014;90:740–6.
- 29 Ministry of Health and Social Welfare. *Tanzania health sector strategic plan (HSSP IV) 2015–2020*. Dar es Salaam: Ministry of Health and Social Welfare, 2015.
- 30 Frumence G, Nyamhanga T, Mwangi M, et al. Challenges to the implementation of health sector decentralization in Tanzania: experiences from Kongwa district Council. *Glob Health Action* 2013;6:20983.
- 31 Afnan-Holmes H, Magoma M, John T, et al. Tanzania's countdown to 2015: an analysis of two decades of progress and gaps for reproductive, maternal, newborn, and child health, to inform priorities for post-2015. *Lancet Glob Health* 2015;3:e396–409.
- 32 SIKIKA. *Medicines and medical supplies availability report. using absorbent gauze availability survey as an entry point. A case of 71 districts and 30 health facilities across mainland Tanzania*. Dar es Salaam, Tanzania: SIKIKA, 2011.
- 33 Ministry of Health and Social Welfare. *Tanzania service Savailability and Seadiness Ssessment (SARA) 2012*. Dar es Salaam: Ministry of Health and Social Welfare and Ifakara Health Institute, 2013.
- 34 Mtei G, Makawia S, Ally M, et al. Who pays and who benefits from health care? An assessment of equity in health care financing and benefit distribution in Tanzania. *Health Policy Plan* 2012;27(Suppl 1):i23–34.
- 35 Mtei G, Makawia S, Masanja H. Monitoring and evaluating progress towards universal health coverage in Tanzania. *PLoS Med* 2014;11:e1001698.
- 36 Ministry of Health Community Development, Gender, Elderly and Children. *Mid term review of the health sector strategic plan IV 2015 –2020: health finance technical report*. Tanzania: Ministry of Health Community Development, Gender, Elderly and Children, 2019.
- 37 Ministry of Health Community Development, Gender, Elderly and Children. *Decentralised direct facility financing concept note and roadmap*, 2017.
- 38 Kapologwe NA, Kalolo A, Kibusi SM, et al. Understanding the implementation of direct health facility financing and its effect on health system performance in Tanzania: a non-controlled before and after mixed method study protocol. *Health Res Policy Syst* 2019;17:11.
- 39 Ministry of Health Community Development, Gender, Elderly and Children. *Improved CHF design document: short term plan towards universal health coverage*. Tanzania: Ministry of Health Community Development, Gender, Elderly and Children, 2018.
- 40 Ministry of Health and Social Welfare. *Results based financing (RBF) design document*. Dar es Salaam: Ministry of Health and Social Welfare, 2015.
- 41 Borghi J, Mayumana I, Mashasi I, et al. Protocol for the evaluation of a pay for performance programme in Pwani region in Tanzania: a controlled before and after study. *Implement Sci* 2013;8:80.
- 42 Mayumana I, Borghi J, Anselmi L, et al. Effects of payment for performance on accountability mechanisms: evidence from Pwani, Tanzania. *Soc Sci Med* 2017;179:61–73.
- 43 Anselmi L, Binyaruka P, Borghi J. Understanding causal pathways within health systems policy evaluation through mediation analysis: an application to payment for performance (P4P) in Tanzania. *Implement Sci* 2017;12:10.
- 44 World Bank. *World bank impact evaluation toolkit*, 2018. Available: <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTH/EALTHNUTRITIONANDPOPULATION/EXTHSD/EXTIMPEVALTK/0,,contentMDK:23262154–pagePK:64168427–piPK:64168435–theSitePK:8811876,00.html> [Accessed 9 May 2018].
- 45 Jacobs R, Smith PC, Street A. *Measuring efficiency in health care: analytic techniques and health policy*. Cambridge University Press, 2006.
- 46 Hafidz F, Ensor T, Tubeuf S. Efficiency measurement in health facilities: a systematic review in low- and middle-income countries. *Appl Health Econ Health Policy* 2018;16:465–80.
- 47 Hollingsworth B, Parkin D. Efficiency and productivity change in the English National health service: can data envelopment analysis provide a robust and useful measure? *J Health Serv Res Policy* 2003;8:230–6.
- 48 Hollingsworth B. The measurement of efficiency and productivity of health care delivery. *Health Econ* 2008;17:1107–28.
- 49 Hollingsworth B, Wildman J. The efficiency of health production: re-estimating the WHO panel data using parametric and non-parametric approaches to provide additional information. *Health Econ* 2003;12:493–504.
- 50 Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *Eur J Oper Res* 1978;2:429–44.
- 51 Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Sci* 1984;30:1078–92.
- 52 Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. *J Econom* 2007;136:31–64.

- 53 Moreno-Serra R, Anaya-Montes M, Smith PC. Potential determinants of health system efficiency: evidence from Latin America and the Caribbean. *PLoS One* 2019;14:e0216620.
- 54 Badunenko O, Tauchmann H. Simar and Wilson two-stage efficiency analysis for Stata. *Stata J* 2019;19:950–88.
- 55 Markovitz AA, Ryan AM. Pay-For-Performance: disappointing results or masked heterogeneity? *Med Care Res Rev* 2017;74:3–78.
- 56 Alhassan RK, Nketiah-Amponsah E, Akazili J, et al. Efficiency of private and public primary health facilities accredited by the National health insurance authority in Ghana. *Cost Eff Resour Alloc* 2015;13:23.
- 57 Achoki T, Hovels A, Masiye F, et al. Technical and scale efficiency in the delivery of child health services in Zambia: results from data envelopment analysis. *BMJ Open* 2017;7:e012321.
- 58 Novignon J, Nonvignon J. Improving primary health care facility performance in Ghana: efficiency analysis and fiscal space implications. *BMC Health Serv Res* 2017;17:399.
- 59 Zere E, Mbeeli T, Shangula K, et al. Technical efficiency of district hospitals: evidence from Namibia using data envelopment analysis. *Cost Eff Resour Alloc* 2006;4:5.
- 60 Akazili J, Adjuik M, Jehu-Appiah C, et al. Using data envelopment analysis to measure the extent of technical efficiency of public health centres in Ghana. *BMC Int Health Hum Rights* 2008;8:11.
- 61 Masiye F. Investigating health system performance: an application of data envelopment analysis to Zambian hospitals. *BMC Health Serv Res* 2007;7:58.
- 62 Kirigia JM, Sambo LG, Scheel H. Technical efficiency of public clinics in KwaZulu-Natal Province of South Africa. *East Afr Med J* 2001;78:S1–13.
- 63 Mujasi PN, Asbu EZ, Puig-Junoy J. How efficient are referral hospitals in Uganda? A data envelopment analysis and tobit regression approach. *BMC Health Serv Res* 2016;16:230.
- 64 Marschall P, Flessa S. Efficiency of primary care in rural Burkina Faso. A two-stage DEA analysis. *Health Econ Rev* 2011;1:5.
- 65 Jat TR, Sebastian MS. Technical efficiency of public district hospitals in Madhya Pradesh, India: a data envelopment analysis. *Glob Health Action* 2013;6:21742.
- 66 Bobo FT, Woldie M, Wordofa MA, et al. Technical efficiency of public health centers in three districts in Ethiopia: two-stage data envelopment analysis. *BMC Res Notes* 2018;11:465.
- 67 Kruk ME, Hermosilla S, Larson E, et al. Bypassing primary care clinics for childbirth: a cross-sectional study in the Pwani region, United Republic of Tanzania. *Bull World Health Organ* 2014;92:246–53.
- 68 Kruk ME, Mbaruku G, McCord CW, et al. Bypassing primary care facilities for childbirth: a population-based study in rural Tanzania. *Health Policy Plan* 2009;24:279–88.
- 69 Yaffee AQ, Whiteside LK, Oteng RA, et al. Bypassing proximal health care facilities for acute care: a survey of patients in a Ghanaian accident and emergency centre. *Trop Med Int Health* 2012;17:775–81.
- 70 White JS, Speizer IS. Can family planning outreach bridge the urban-rural divide in Zambia? *BMC Health Serv Res* 2007;7:143.
- 71 Nxumalo N, Goudge J, Thomas L. Outreach services to improve access to health care in South Africa: lessons from three community health worker programmes. *Glob Health Action* 2013;6:19283.
- 72 Yitbarek K, Adamu A, Tsega G, et al. Technical efficiency of maternal and reproductive health services in public hospitals of Oromia regional state, Ethiopia. *Health Serv Insights* 2019;12:1178632919837630:117863291983763.
- 73 Jehu-Appiah C, Sekidde S, Adjuik M, et al. Ownership and technical efficiency of hospitals: evidence from Ghana using data envelopment analysis. *Cost Eff Resour Alloc* 2014;12:9.
- 74 Anselmi L, Lagarde M, Hanson K. The efficiency of the local health systems: investigating the roles of health administrations and health care providers. *Health Econ Policy Law* 2018;13:10–32.
- 75 Turcotte-Tremblay A-M, Spagnolo J, De Allegri M, et al. Does Performance-based financing increase value for money in low- and middle- income countries? A systematic review. *Health Econ Rev* 2016;6:30.
- 76 Borghi J, Little R, Binyaruka P, et al. In Tanzania, the many costs of pay-for-performance leave open to debate whether the strategy is cost-effective. *Health Aff* 2015;34:406–14.
- 77 Palmer N, Mueller DH, Gilson L, et al. Health financing to promote access in low income settings-how much do we know? *Lancet* 2004;364:1365–70.
- 78 Bhutta ZA, Salam RA, Lassi ZS, et al. Approaches to improve quality of care (QoC) for women and newborns: conclusions, evidence gaps and research priorities. *Reprod Health* 2014;11(Suppl 2):S5.
- 79 Atun R, de Jongh T, Secci F, et al. A systematic review of the evidence on integration of targeted health interventions into health systems. *Health Policy Plan* 2010;25:1–14.
- 80 Sambo LG, Kirigia JM. Investing in health systems for universal health coverage in Africa. *BMC Int Health Hum Rights* 2014;14:28.
- 81 Kruk ME, Porignon D, Rockers PC, et al. The contribution of primary care to health and health systems in low- and middle-income countries: a critical review of major primary care initiatives. *Soc Sci Med* 2010;70:904–11.
- 82 Walley J, Lawn JE, Tinker A, et al. Primary health care: making Alma-Ata a reality. *Lancet* 2008;372:1001–7.
- 83 Anselmi L, Borghi J, Brown GW, et al. Pay for performance: a reflection on how a global perspective could enhance policy and research. *Int J Health Pol Manage* 2020.