New methods of improving the quality of medical services

Abstract
This article draws together examples of new ways to improve quality of health services at scale in Performance-based financing programs. Universal Health Coverage is one of the Sustainable Development Goals being pursued by many countries. Coverage without commensurate quality health services provides limited effective benefits to the national population. With strategic purchase of services, performance-based financing programs combine coverage and quality and therefore measure and reward quality. We recognized that strengthened administrative accountability and the widespread presence of performance-based financing programs, some at a large scale, were an opportunity to introduce and test new approaches at scale for measuring and improving quality. The approaches we pursued looked at structural quality and process measures of quality - which are closer to ‘real quality’ - as well as some outcome measures, such as patient satisfaction. Three types of tools are described: vignettes, competency tests and patient satisfaction surveys, and their application is illustrated by country examples Kyrgyzstan.

Keywords: universal health coverage, quality of care, performance-based financing, pay-for-performance, scaling-up, LMIC, vignettes, competency tests, client satisfaction surveys

Introduction
The World Health Report 2010 promotes investing in Universal Health Coverage (UHC). The path to UHC involves securing resources, reducing the reliance on direct payments for health services, and improving efficiency and equity. Lack of resources to finance health services, high financial barriers to access services and, in the poorest countries unavailability of services hamper efforts to move towards universal health coverage. In September 2015, the United Nations General Assembly adopted the Sustainable Development Goals, of which Goal 3, sub goal 3.8 is achieving Universal Health Coverage. Moving towards UHC would involve defining a benefit package, financing this benefit package in such a way that the poorest can also access these services, and embarking on strategic purchasing to ensure that this benefit package is delivered. Absence of quality or even services of lesser quality will yield lesser health benefits and attenuate the benefits of an expanded benefit package. Poor quality is well documented, regardless the country. A service of lesser quality would yield lesser health benefits and therefore would be less effective. The World Health Organization estimates that between 20%-40% of all health spending is wasted due to inefficiencies and poor quality. Technical progress in health interventions in poor countries has an effect on under-five mortality. Quality is important for the post-MDG agenda. Traditional approaches such as provider training and supervision have at best a limited impact, requiring a better or different way of doing things. Hence the focus on effective coverage and effective benefits. This article is organized as follows: first, we will provide a background by describing quality of care as a concept, further we will discuss Performance-based financing and discuss the experience measuring quality in Performance-based financing programs. Thereafter, we will describe a theory of change underlying the use of quality measures in the context of Performance-based financing, after which we will describe the novel quality measurement methods. Finally, we will provide four country examples of the application of these tools. In the discussion section, we will reflect on some challenges and possible caveats.

Background
Quality of care
Quality of Health Services in lower- and middle income countries is low and variable everywhere. Through the World Bank’s Service Delivery Indicator Surveys carried out in various Sub-Saharan African countries, the poor state of service delivering ability for delivering quality health services is apparent. An established framework for understanding quality of health services is Donabedian’s which a distinction is made between structural, process and outcome measures of quality. Structural measures are the inputs necessary for providing quality services such as equipment, drugs and trained providers, process measures are the actions and activities of health providers and outcomes are the results of medical action such as patient satisfaction, improved health, disability or death. Process measures of quality, i.e. what happens between the patient and the provider also referred to as provider effort are commonly associated more closely with health outcomes. Frequently, structural and process quality are present conjointly to lead to desired outcomes. As an example, in the case of treating pulmonary tuberculosis: tuberculosis drug availability (structural element) is proximate to making the right diagnosis, the right explanations to the patient, and the delivery of effective direct observed treatment (process elements) and should all be present to yield the highest possible chance for successful cure (outcome). The strength of this framework, for example in the case of moderately severe diarrhea in a young child, is that, appropriate care can be effectively ‘provided under a palm tree’ by a ‘knowledgeable person’ (structural) if that provider make the right diagnosis, and provides correct advice using integrated management of childhood illnesses.
New methods of improving the quality of medical services

DOI: 10.15406/jpnc.2018.08.00329

Performance-based financing and strategic purchasing

Key features of these PBF approaches are:

a) Defining basic and complementary health packages for strategic purchasing.

b) Strategic purchasing adjusted through measures of service quality.

c) Purchasing from public, private and quasi-public providers.

d) Involvement of national and sub-national health administration in stewardship, supervision, quality measurements and other technical and managerial supportive tasks.

e) Community client satisfaction surveys to strengthen community voice.

f) Rigorous internal and external verification mechanisms.

g) Creation of fiscal space through linking to national health financing strategies assisted through the Global Financing Facility.

PBF approaches are systemic approaches, and can best be conceptualized as a leveraging mechanism and as a health system strengthening mechanism. These approaches work best when embedded in overall health reforms in which various system pillars are simultaneously addressed such as human resources, health information systems, pharmaceutical sector and fiscal space for health.

PBF and quality of care

Quality adjusted strategic purchasing of essential health services is an essential aspect of PBF approaches

PBF approaches thus include the use of a quantified quality checklist applied quarterly to each contracted health facility. Such checklists differ depending on level (health center/community versus hospital), context, and typically involve over 120 data elements. Health facilities in poor countries, due to chronic underfunding, lack many structural quality features that health facilities in better resourced contexts display as a matter of routine. Autonomy is introduced in health facilities through PBF, business plans are developed and negotiated, and start-up funds are provided to facilities based on approved business plans. There is a reason why quality checklists have preponderance for structural quality elements, namely: these structural quality elements are linked—through business plans—to the use of startup funds and income gained through PBF and other cash sources. Conjointly with a stipulation in the performance contract that, for instance, at the least 50% of income ought to be invested in structural quality, health financing can be directed effectively and efficiently to the most obvious defined needs. These needs are defined bottom-up by health facility staff in collaboration with their communities and governing boards. Therefore, the preponderance of structural quality measures in PBF approaches is part and parcel of the output-based health financing approach and should be assessed likewise. Notwithstanding, there is a general recognition that there is a paucity of process measures and accordingly, the various quality checklists have increased weights for process measures of quality. These process measures are mostly based on file or record reviews and sometimes on observations. They are known to be difficult to ascertain or to counter-verify. For instance, a commonly used process measure like ‘well-filled partograph’ is frequently filled after delivery took place instead and when missing partograph elements can be filled by knowledgeable providers before a verification would take place. In a similar vein, criterion-based medical audits although perhaps working in different contexts also break down when used in PBF approaches due to a similar propensity for gaming the right answer instead of the actual practice. The veracity of these process measures is thus near impossible to counter-verify by third party audit after financing has taken place. In contexts where the health administration is weak, barely functional or absent performance contracting of the health administration—the so-called ‘internal contracting’—is an important element for health system strengthening. The local health administration is contracted to carry out various health administrative tasks as part of an output-based contract. The health administration performs so-called ‘ex-ante quality verification’; these verifications constitute the composite quality score that is used by the purchasing agent to determine the facility financing. This ex-ante verification is an internal verification exercise. However, the health administration is kept accountable through independent third party counter-verification exercises, with consequential penalties in the case of gaming. PBF approaches have expanded and now reached scale in many countries (e.g., in SSA, Central and South-East Asia and the Caribbean) making it possible to look at the impact of strategic purchasing for UHC and any novel or promising quality interventions. A project in Kyrgyzstan covers all 63 Rayon (district) hospitals.

A theory of change

Well-designed and well-implemented PBF programs targets the health system at various levels and therefore intended results have a system-wide scope. Stewardship, for example, is strengthened at central level while administrative accountability is strengthened at the sub national level. The supply of essential health services at the health facilities across countries are expanded, financial accessibility of services is augmented targeting the poorest populations. Various other interventions at the community level strengthen voice and amplify the demand for services. Let us turn our attention to some of the novel quality tools that are being made operational.

Description of the novel quality interventions

Traditionally, much of our quality assessments have been about judging service availability which is a structural quality of health facilities. Such assessments are useful.

Below, we reflect on our modest armamentarium of methods to measure and document quality processes and outcomes. Our intent is to operationalize these methods in such a manner that they are effective along three dimensions:

- The remaining 50% is distributed as performance bonuses to health staff, using customized tools. A key element of the PBF approach is making health personnel see their health facility as a cooperative with a social mandate, in which a significant part of income gained is theirs.
1. Efficient: can be added to an existing PBF program at minor incremental cost.

2. Responsive: able to document an improvement or lead to advancement on the previous state of affairs.

3. Scalable: can be done at the population level.

The operationalization of these schemes needed to be done in a way that they can strengthen indices for health care quality and work in an environment of a Performance-based financing scheme. Such scheme, introduced in the PBF context, will be validated by regular and rigorous assessments (backed by third-party verifications), performance feedback, elements of supportive supervision and public benchmarking. We postulate that such environments offer a fertile context for behavior change as they mix various behavioral triggers and repercussions. Benchmarking; inexpensive and readily available; documentation is highly variable; gaming is not practical.

Vignettes are structured written or online case simulations (Table 1) Vignettes are structured written or online case simulations. By standardizing patients, vignettes offer the possibility of examining physician has the skills/knowledge required to care for the patient. To simulate a range of medical conditions to evaluate whether a physician has the skills/knowledge required to care for the patient. This also encompasses some patient-focused quality measures such as provider responsiveness and attitudes; health facility cleanliness; drug availability and affordability. The objective, therefore, is to add measures to the quality index that flesh out quality measures and directly reward process and outcome quality. We looked at five methods used for assessing medical process quality:

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Scope (efficiency; responsiveness and scalability)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chart review</td>
<td>Readily available; currently practiced</td>
<td>Documentation is highly variable; Gaining is easy and Case mix is uncontrolled</td>
<td>Inefficient but responsive and scalable</td>
</tr>
<tr>
<td>Vignettes</td>
<td>Cases are standardized for benchmarking; inexpensive and readily scalable; can also be used for rare conditions</td>
<td>Limited experience in a PBF environment</td>
<td>Efficient, responsive and scalable</td>
</tr>
<tr>
<td>Direct Observation</td>
<td>Assesses competency</td>
<td>Limited experience in a PBF environment, difficult to scale</td>
<td>Difficult to scale</td>
</tr>
<tr>
<td>Mystery Patient</td>
<td>Avoiding Hawthorne effect ('Gold Standard')</td>
<td>No experience in a PBF environment; limited range of conditions can be simulated; training and inter-rater reliability a challenge</td>
<td>Not efficient nor scalable</td>
</tr>
<tr>
<td>Exit interview</td>
<td>Patient perspective on the care provided can be quantified providing information on effort</td>
<td>No experience in a PBF environment Laborious review and analysis of data</td>
<td>Theoretically possible, but probably not practical</td>
</tr>
<tr>
<td>Client Satisfaction survey</td>
<td>Information on patient opinion and appreciation Information on out of pocket payments</td>
<td>Recall is a problem Design and testing of the instrument is crucial</td>
<td>Yes, adjusted</td>
</tr>
</tbody>
</table>

Chart reviews abstract practice documentation from the patient’s chart using explicit evidence based criteria. Vignettes are standardized medical cases, where providers care for the same patients facilitate comparison of practices; they are done in various ways from full cases akin to role play, or as a short answer, affecting a written exam. Medical care can be observed directly by the examiner or a mystery or standardized patient. Mystery patients are trained actors who present with certain symptoms and are trained to observe and report on medical actions performed on them. Finally, patients can be interviewed on what has been done to them when leaving a consultation room. These methods are variably used for research and for audits, for accreditation or teaching. Considering a pre-existing PBF program, efficiency, responsiveness and scalability, the following options to measure process and outcome quality are available to us: (Table 1) Vignettes are structured written or online case simulations that have been used in a wide variety of clinical [Peabody et al 2000, Wainwright et al 2010] and non-clinical [Baudson and Preckel 2013, Wallander20 Wason et al 2002] settings. As described by Alexander and Becker (1978), vignettes are “short descriptions of a person or a social situation which contain precise references to what are thought to be the most important factors in the decision-making or judgment-making process of respondents”. Clinical vignettes are designed to simulate a range of medical conditions to evaluate whether a physician has the skills/knowledge required to care for the patient. By standardizing patients, vignettes offer the possibility of examining different providers’ clinical interpretations and directly comparing their responses.
These vignettes are customized to the local treatment guidelines. The same vignettes are used at both health center and first level referral hospitals. Vignettes will be applied by trained supervisors from the district health teams (for health centers) and trained peers and provincial health staff (for hospitals). Vignettes are utilized in the same environment where the provider usually sees her patients; i.e. in a consultation room or in the delivery room. The supervisor brings a ‘case’ for instance presents himself as the mother of a three-year-old child with diarrhea. The provider then should start his consultation and to each question asked a programmed answer is provided. There are six categories of questions: history taking, physical examination, laboratory exams, diagnosis, treatment, and explanations provided to the patient. Tablet based software will be used which will enable feedback after the vignettes session (in-service training component), and automated upload to a cloud based database for merging into the quality index. Staff will be randomly selected from the duty roster (hat method), feedback on performance is confidential and direct to the individuals concerned. District and provincial health offices are under a performance contract which inter alia finances timely and correct execution of the quality measures. Counter-verification of results will be done through a mixed systematic random and risk-based protocol, and included in the existing counter-verification mechanism carried out by a third party. Fair play is encouraged through financial rewards, temporary exclusion from benefits for the offending supervisors and public benchmarking. Direct observation has been modified to a competency test using the Mama Natalie and Neo Natalie tools. The Mama Natalie will be used for a competency test for post-partum hemorrhage which is one of the leading causes of maternal mortality; the Neo Natalie will be used for a competency test for neonatal resuscitation which is another key competency for providers. Neonatal resuscitation can decrease neonatal mortality by 30%. Both competency tests have been quantified and weighted to score 100 points when 100% correctly executed.

In the table below, adjustments to these methods are described to make the methods dovetail with PBF approaches: Patient satisfaction surveys in Kyrgyzstan are carried out after a systematic random sampling of the inpatient register of women who delivered in the hospital. Ten patients are selected randomly in each of the 63 district hospitals, once per quarter. The patients are contacted through mobile phone by trained supervisors, and after obtaining consent are asked five questions. Four of the five questions have responses on a five point Likert’s scale, while one final question is binary seeking to clarify whether the patient paid providers informally. If the last question is answered ‘yes’, then it voids the points for the entire interview. The ten interviews are weighted 10% of the total quarterly performance score for each hospital.2

The Kyrgyzstan health results-based financing project

Kyrgyzstan is a central Asian country with a 2016 population of 6 million. Its nominal 2015 GDP is $1,103 per capita; its 2016 human development index is 120/187. The Kyrgyz Health Results-based Financing Project, a $12M three-year project funded through a Health Results Innovation Trust fund grant (financed through Norway and financing Project, a $12M three-year project funded through a Health Results-based protocol, and included in the existing counter-verification process) was conceived to study its effects. An impact evaluation was designed with three arms: Group 1 Hospitals received a balanced scorecard with feedback but no payment and a third group where no intervention took place.3 The value of the performance bonus budget was about 15% of the annual hospital income, approximately $1 per capita per year. Each of the groups had 21 hospitals. A baseline survey was carried out in 2013, and a follow-up survey will be done in the third quarter of 2017. Results over a two-year period showed impressive improvements in both intervention arms whereas the control arm remained the same. An average score of 9.8% was found for 63 Rayon hospitals in July 2014. Group 1 increased its performance over a period of two years from 9.3% to 79% (double difference with Group 2 of 60.8% (52.2 percentage point increase; double difference with Group 3 of 67.5 percentage points). Group 2 increased its performance from 8.6% to 60.8% (52.2 percentage point increase; double difference with 49.3%). These increases occurred notwithstanding changes to weightings after one year and introduction of a multitude of skill and competency based tests that put the performance bar progressively higher. Group 3 performance barely increased. Two rounds of balanced scorecard application was done to the Group 3 hospitals in the first half of 2016, and their performance remained low at 13.7%, representing a 2.2% increase from their July 2014 baseline of 11.5%. Kyrgyzstan RBF portal http://rbf.med.kg/


Discussion

New methods to strengthen service delivery quality are being applied in various contexts, in PBF programs and at scale. From field observations and from routine data collection it appears that quick improvements in key quality measures are obtained. In all four country examples, scientific evidence will be acquired through rigorous impact evaluations, and these will shed light on the degree of impact through quantitative analysis enriched with a battery of qualitative research. Results of these studies will be available in a few years from now. While awaiting these impact evaluation results we can observe results obtained through routine data, and we can reflect on lessons learned. For now, there are a set of theoretical and actual challenges with these novel methods, and these problems can be classed at the micro level (the methods, tools and reforms) and at the macro (performance improvement) level. For the micro level, we have challenges with the know-do gap, with design and implementation of these novel quality tools in general, and design and implementation of PBF approaches in specific. The know-do gap (3.4.2) is the observed gap between what providers know and what they do with that knowledge. This know-do gap is measured through impact evaluations. This know-do gap is a real problem for using knowledge tests as a proxy for process quality, namely: if providers systematically skimp on actual care notwithstanding a high score on the vignettes and the competency tests, then we would have a problem. We know very little about this phenomenon in PBF contexts. A few theoretical arguments can be brought forward that would mitigate this know-do gap. First, provider effort is enhanced by structured formative supervision and this is a key component of PBF approaches; second, the change in provider payment mechanism (moving to a case-based payment for curative care) would align incentives to prescribe drugs more rationally i.e. adhere to protocols (taught through the vignettes), third, perceived enhanced quality by patients (observed structural elements; more time spent in consultation) will lead to a positive virtuous cycle, and fourth, observation with positive feedback, and benchmarking of performance are all strong enablers for enhancing performance. Nevertheless, weak provider capacity is due to generally weak pre-service training and this phenomenon needs to be tackled at that level too. Introducing competency based training in pre-service curricula and competency based tests in in-service trainings would go well together and would mutually reinforce each other.9 Design and implementation of these novel quality tools are a challenge. Field observations from Kyrgyzstan teach us that continued high quality mentoring and support are necessary to reach desired effects. There are challenges in ensuring bottom-up design of these tools, and this bottom-up design is necessary to strengthen local acceptance, gather inputs and ensure adaptations. There will also be a need for validating instruments.4 All this costs time and needs continued handholding. Investments need to be made in creating capacity of a group of national trainers who will be able to continue implementing and developing these approaches.

Design and implementation of effective PBF approaches is not easy.20,24 A further challenge in these approaches is the use of local administrative structures for assessments. Whereas engaging the local health administration is a powerful strategy for strengthening local health systems, using PBF approaches for tagging on these methods necessitates very good counter-verification measures to make all adhere to the new rules of the game. Finally, there are political challenges. PBF approaches frequently challenge those that stand to lose out due to these reforms. For instance, decentralizing drug procurement and decentralizing budgets to health facilities all lead to lesser income for those that previously used to manage these resources at higher levels. Countries that do not invest in PBF, for instance through creating a line item for PBF in the Ministry of Finance budget and paying for PBF do pose a sustainability risk to these reforms.36–60 Also, leadership changes pose challenges. A new team might have different priorities and could derail these reforms. Finally, development partner buy in is crucial in many contexts. Development partners operating in countries where external funding is a large part of the fiscal space for health frequently do not buy into PBF reforms. Fragmented health systems in the poorest countries, emerging from conflict or hit by socio economic crises would benefit from strong leadership from the host country, and from development partners willing to join efforts in financing and thus enabling such reforms to happen.

Acknowledgements

None.

Conflict of interest

The author declares there is no conflict of interest.

References

5. Janisson Dear T, Shane M Murphy, Martin E Sandbu. Why has under-5 mortality decreased at such different rates in different countries?. Journal of Health Economics. 2016;48:16–25.
New methods of improving the quality of medical services


17. Editorial Can we Expect Results-Based Financing to Improve Quality of Care?. Glob Health Sci Pract. 28;5(1):1–3.


New methods of improving the quality of medical services


50. MSP. Indice de disponibilité et de capacité operationnelle des services de santé (SARA RDC); 2014.


