How do performance-based financing programmes measure quality of care? A descriptive analysis of 68 quality checklists from 28 low- and middle-income countries

Erik Josephson,1 Jessica Gergen,2 Martha Coe,2 Samantha Ski,3 Supriya Madhavan4 and Sebastian Bauhoff5,*

1Independent Consultant, 1620 Corcoran St NW, Apt A, Washington, DC, USA, 2ThinkWell, Research & Evaluation Division, 1155 F Street N.W., Suite 1050, Washington, DC 20004, USA, 3University Research Co. (URC), Chevy Chase, MD, USA, 4US Agency for International Development (USAID), Washington, DC, USA and 5Center for Global Development (CGD), Washington, DC 20036, USA

*Corresponding author. Center for Global Development (CGD), Washington, DC, USA. E-mail: sbauhoff@cgdev.org

Accepted on 13 April 2017

Abstract

This paper seeks to systematically describe the length and content of quality checklists used in performance-based financing programmes, their similarities and differences, and how checklists have evolved over time. We compiled a list of supply-side, health facility-based performance-based financing (PBF) programmes in low- and lower middle-income countries based on a document review. We then solicited PBF manuals and quality checklists from implementers and donors of these PBF mechanisms. We entered each indicator from each quality checklist into a database verbatim in English, and translated into English from French where appropriate, and categorized each indicator according to the Donabedian framework and an author-derived categorization. We extracted 8,490 quality indicators from 68 quality checklists across 32 PBF implementations in 28 countries. On average, checklists contained 125 indicators; within the same program, checklists tend to grow as they are updated. Using the Donabedian framework, 80% of indicators were structure-type, 19% process-type, and less than 1% outcome-type. The author-derived categorization showed that 57% of indicators relate to availability of resources, 24% to managing the facility and 17% assess knowledge and effort. There is a high degree of similarity in a narrow set of indicators used in checklists for common service types such as maternal, neonatal and child health. We conclude that performance-based financing offers an appealing approach to targeting specific quality shortfalls and advancing toward the Sustainable Development Goals of high quality coverage. Currently most indicators focus on structural issues and resource availability. There is scope to rationalize and evolve the quality checklists of these programs to help achieve national and global goals to improve quality of care.

Keywords: Performance-based financing, quality of care, universal coverage
Introduction

As characterized by the third Sustainable Development Goal, universal health coverage requires access to health care services that are of high quality. A corollary to this statement is the immediate need to improve quality of care, often defined as the ‘degree to which health care services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge’ and as health care that is ‘safe, effective, timely, efficient, equitable and people-centered’ (Institute of Medicine [US] [IOM] Committee on Quality of Health Care in America 2001; Tuncalp et al. 2015). This general framework has been adapted to specific areas, such as maternal and newborn health (Tuncalp et al. 2015).

Improving health care quality is a pressing concern given rising evidence of low levels of basic quality—ranging from absenteeism and low levels of medical training to high rates of incorrect diagnosis and treatment (Chaudhury 2006; Das et al. 2008, 2015)—and the large variations in quality in low- and middle-income countries (LMICs) (Das and Gertler 2007; Lee et al. 2016). Shortfalls in quality exist overall and in specific areas of clinical care. For instance, an assessment of the quality of antenatal care in Kenya found low levels and substantial heterogeneity in the effectiveness, efficiency, accessibility, acceptability, equity and safety indicators (Lee et al. 2016).

Quality improvement efforts are not new to LMICs though the conventional methods to improve quality have yielded variable results. A recent systematic review—commissioned by the US IOM—of six major quality improvement modalities in which USAID has invested yielded few studies showing strong evidence of effectiveness of any one modality, and some evidence attesting to the effectiveness of a combination of two modalities (Institute of Medicine Board on Global Health; The National Academies of Sciences, Engineering, and Medicine 2015).

A newer approach to improving the quality (and quantity) of health care services in low- and lower-middle income countries is performance-based financing (PBF). PBF is a strategic purchasing mechanism which, in these settings, generally pays fee-for-service for a short list of services, while separately considering quality of care (Musgrove 2011).

More specifically, PBF programmes often incentivize quality of care directly (through inclusion in the payment formula) or through a composite quality multiplier which deflates or inflates the quantity-based bonus payment (Musgrove 2011; Health Results Innovation Trust Fund 2015). The multiplier is commonly constructed using a checklist of quality indicators that are scored and aggregated regularly, e.g. quarterly (Musgrove 2011). Verification not only ensure the accurate calculation of quality scores, but also affords providers opportunities to receive feedback, learn about quality improvement approaches, and request assistance from higher levels.

In practice, PBF programmes often combine financial incentives for quality with quality assurance, formative supervision, and continuous education (Fritsche et al. 2014). Some PBF programmes are devising approaches other than checklists to more directly measure those aspects of quality of care that are expected to be closely related to health outcomes, such as the use of medical simulations in a PBF programme in Kyrgyzstan (World Bank 2016). To date, however, quality checklists remain a central component of many PBF programmes that target quality (Fritsche et al. 2014).

Despite the widespread implementation of PBF in health, and the potential of PBF for quality improvement, there is little documentation of how operational PBF programmes measure and score quality of care. Most published research on PBF is concerned with the impact of the programmes (Rusa et al. 2009; Das et al. 2016), and only a few case studies exist on design considerations (Meessen et al. 2006; Musgrove 2011; Meessen 2014, Gergen et al. 2016). Indeed, an ongoing debate in the PBF community of practice relates to questions on whether and how to incorporate quality in these programmes.

In this study, we describe the content of PBF quality checklists, their similarities and differences, and how checklists have evolved over time within a programme.

The overall goal of this article is to complement guidance for designing PBF programmes with a descriptive analysis of what and how indicators are used in PBF implementations, and to initiate a broader reflection on how quality of care is approached within PBF designs.

Methods

We reviewed 68 checklists (at primary, secondary and tertiary levels) used in 32 PBF programmes across 28 countries. Our analysis is based on data extracted from operational manuals and checklists that are generally not publicly available. We also characterized areas for improving the PBF approach to quality, either through improvements across programmes or through learning between programmes, and reflect on the role of PBF in the broader agenda on quality of care.

We collected PBF manuals and quality checklists on supply-side PBF programmes that target health facilities in LMICs, and then extracted and classified quality indicators. In the first step, we compiled a list of known existing PBF programmes based on a document review of published analyses in both the peer-reviewed and grey literature and expert consultation with a number of key donor representatives from the World Bank, Kreditanstalt Für Wiederaufbau (KfW), Centres for Disease Control and Prevention and USAID. Each donor provided a list of their PBF programmes and a key contact person for each, if available. We solicited programmatic information from implementers and donors primarily through email. Our research team collected and organized programme manuals and accompanying tools used to measure quality performance for all facility levels (primary, secondary and tertiary).

All programmes identified were included for further analysis if sufficient programme information could be obtained. We considered all programmes irrespective of their year of implementation, size or phase (i.e. the programmes range from small-scale pilots to national implementations).

Key Messages

- Performance-based financing programmes often employ quality of care checklists.
- Checklists tend to be long and mostly focus on structural quality and resource availability.
- There is scope to rationalize and evolve checklists to better targeted specific quality shortfalls.
We solicited PBF quality checklists for all health facility levels, including current and previous versions, in the event that checklists had been iterated or had evolved. In four countries (DRC, Malawi, Rwanda and Tanzania) we obtained different quality checklists for the multiple concurrent PBF programme implementations.

In the second step we extracted and classified checklist indicators. For the purposes of this study we defined a quality indicator as any measure with an associated score in a PBF quality checklist. We excluded checklist indicators that did not have an associated score, i.e. these indicators have no impact on PBF payments to facilities. We considered each indicator in each checklist as a unique indicator, irrespective of whether it is worded exactly the same as, or very similarly to, an indicator in another PBF checklist. Thus, the same indicator may be used in the primary level (e.g. health center) and secondary level (e.g. hospital) checklists of the same PBF implementation, or similar indicators might be used in the PBF implementation of two countries. For the purpose of this analysis, these would count as two indicators in both cases.

We entered each indicator from each checklist into an Excel database verbatim. Indicators in a language other than English, primarily French, were translated and reviewed by a person who speaks the language to ensure a faithful, common parlance translation.

Choice of frameworks for analysis

In the absence of a fully developed framework for quality improvements in PBF, we used two quality of care frameworks to classify indicators. First, we employed the Donabedian (1988) framework, which defines the health-care triad of structure, process and outcome and most aptly speaks to the theory of change of PBF (Table 1). Although this has been refined for specific areas, such as maternal and newborn health (Tuncalp et al. 2015), it is also sufficiently broad to cover the wider range of areas captured in the checklists.

In addition, we developed a cross-cutting categorization which points more directly to the types of indicators used in the PBF mechanisms and which captures the types of structural and behavioural changes which PBF seeks to bring about (Musgrove 2011). We used four main categories that reflect focus areas for quality under supply side PBF: Resource Availability, Facility Management, Clinical Effort and Knowledge and Non-Clinical Effort (Table 1). Within Resource Availability we employed sub-categories such as Facility Equipment and Resources, Facility Infrastructure, Pharmaceutical Availability and Human Resources to further detail the types of resources PBFs focused on. Facility Management indicators included those on financial management, meeting minutes, management plans of various kinds, indicators looking at the correct implementation of PBF and others. We defined the category. Clinical Effort to include activities of clinicians in the interaction with patients, e.g. adherence to clinical protocols during the encounter or adequate record-keeping directly related to patient care (such as recording a patients’ blood pressure). We defined Knowledge and Non-Clinical Effort as knowledge (e.g. of protocols) and general activities, such as maintaining a current patient register.

We classified each indicator by both the Donabedian framework and the author-derived categorization. We included the point allocation for each indicator from the checklists and their weight (0–100%) within the checklists. The weight of each indicator within a checklist is not generally included in the checklist, nor are checklists generally scored out of 100. We used the maximum score of each indicator and, where needed, scaled the score so that the sum of all scores in the checklist would range from 0 to 100.

Three researchers entered and classified the indicators. Classifications were guided by agreed-upon definitions based on the Donabedian framework and the author-derived categorization. When a discrepancy about indicator classification arose, or if it indicator was not phrased specifically enough to be sure in which category it belonged, the three researchers discussed and made decisions based on consensus.

Analysis

We used pivot tables in Microsoft Excel to create the descriptive tables. We assessed indicators on their relative importance using weights normalized to range 0–100% of the total checklist. We generated descriptive statistics for all variables including indicator typology and quality of care dimensions for each country and cumulatively.

We also conducted a comparative analysis of the degree of similarity in checklists across countries for maternal, newborn and child care (MNCH)—which represent clinical areas present in almost all checklists. Researchers assessed indicators by keywords; segmenting the sample, like ante-natal care (ANC) visits and delivery room; and then manually coding indicators measuring the same construct. The similarity analysis was done on the most recent version of a PBF
programme’s checklist and with programmes as the denominator (rather than checklists) because of strong similarities across multiple versions of checklists within the same implementation. We focused on indicators, which were present in a very similar or identical form in at least 30% of the programmes. Finally, in each of the five countries for which we obtained multiple versions of the quality checklists, we compared across the multiple versions to determine the trends and evolution of quality of care checklists for PBF.

Results

Analytical sample
From an initial list of 51 PBF mechanisms compiled by the authors, we were able to collect 68 quality checklists across 32 PBF implementations that were initiated between 2008 and 2015, in 28 countries (Supplementary Appendix Figure 1). The coverage of PBF programmes was dominated by Sub-Saharan African (21), followed by Europe and Central Asia (3), East Asia and Pacific (2), South Asia (1) and Latin America and the Caribbean (1). We extracted a total of 8490 quality indicators across all checklists.

Length of checklists
On average checklists contained 125 indicators. 28 of 68 checklists were composed of >150 indicators (Table 2).

Table 2. Distribution of checklists by length

| Length of checklist | Checklist n |
|---------------------|-------------|
| 1–50 indicators     | 13          |
| 51–100 indicators   | 16          |
| 101–150 indicators  | 11          |
| 151–200 indicators  | 17          |
| 201–250 indicators  | 10          |
| 251–300 indicators  | 1           |
| Total               | 68          |

Donabedian typology
Using the Donabedian framework, 80% of the indicators were structure-type indicators, 19% process-type, and <1% outcome-type (Table 3; see Supplementary Appendix Table 1 for a breakdown by checklist). Outcome indicator predominantly focused on patient satisfaction or patient knowledge related to health education. Checklists which differed meaningfully from these proportions generally had indicators that were more evenly distributed across structure and process, or placed a heavier emphasis on process relative to structure; only a few checklists also placed any emphasis on outcome-type indicators, e.g. Uganda (58% structure; 35% process; 8% outcome), Kenya’s 2012 tertiary RBF (78% structure; 16% process; 6% outcome), and Malawi’s 2014 KfW-supported programme (53% structure; 41% process; 6% outcome). In six checklists, >90% of the indicators were of the structure type.

Author-derived categorization
Using the author-derived categorization showed that 57% of indicators were related to the availability of resources (Table 3). Within that category, nearly half of the indicators (24% across all indicators) were related to ensuring that equipment and other physical resources in the facility are available and functional. Categories related to the existence and appropriate/sanitary condition of facility infrastructure and pharmaceuticals are also large components of Resource Availability.

Nearly a quarter (24%) of indicators were related to Facility Management, e.g. on financial management of the facility, services available, management plans in place, performance monitoring in place, and staffing plans followed.

A smaller share (11%) of indicators related to Clinical Effort, and an even smaller share (6%) were related to Knowledge and Non-Clinical Effort.

Checklist evolution
Figure 1 details the evolution of quality checklists for different levels in five countries (Benin, Lesotho, Nigeria, Senegal and Zambia),
with a total of 20 checklists with multiple versions spanning 2010–15. In all countries except Lesotho, the number of indicators increased during the course of checklist revisions; on average by 12 indicators on the primary checklists, and 24 indicators on the secondary/tertiary checklists. Overall, it seems that there was no clear pattern for countries to revise up or down the number or weights or types of indicators on their checklists as suggested by the multiple conformations across the table. Benin (tertiary level) and Nigeria were the only two countries to have implemented an outcome indicator, albeit only one in each case. Lesotho’s secondary level checklist and Nigeria’s primary and secondary checklists all increased the proportion of process indicators.

**Commonly used MNCH indicators**

Table 4 lists the 16 most common PBF quality indicators focused on MNCH across the 32 PBF programmes, focusing on those indicators that appear in 30% or more of programmes. 11 of these common indicators related to maternal care and 5 related to newborn/child care.

**Discussion**

Our review of quality of care checklists from PBF programmes suggests several broad patterns. First, checklists used in the majority of supply side, health facility-based PBF mechanisms are relatively long

### Table 3. Indicator distribution by author-derived quality of care categorization and Donabedian dimensions

| Quality of Care Category               | Overall n (percent) | Donabedian Dimensions |
|---------------------------------------|--------------------|-----------------------|
|                                       |                    | Structure | Process | Outcome |
| Resource availability                | 4818 (57)          | 4818      |         |         |
| Facility equipment and resources     | 2014 (24)          |           | 2014    |         |
| Facility infrastructure             | 816 (10)           |           | 816     |         |
| Pharmaceutical availability      | 634 (7)             |           | 634     |         |
| Human resources                   | 414 (5)             |           | 414     |         |
| Other                               | 940 (11)            |           | 940     |         |
| Facility management                | 2,074 (24)         | 2,006     | 57      | 11      |
| Facility planning                  | 1,112 (13)         | 1,111     | 1       |         |
| Facility reports                   | 322 (4)             | 322       |         |         |
| Financial management               | 108 (1)            | 108       |         |         |
| Other                               | 532 (6)             | 465       | 56      | 11      |
| Clinical effort                    | 965 (11)           | 962       | 5       | 3       |
| Knowledge and non-clinical effort   | 513 (6)             | 513       |         |         |
| Other                               | 120 (1)             | 120       |         |         |
| Total                               | 8490 (100)         | 6824 (80) | 1652 (19)| 14 (<1) |

### Table 4. Common PBF quality indicators in MNCH across the 32 PBF programmes

| Global (common) indicators | Programmes n (%) |
|----------------------------|------------------|
| **Maternal health**        |                  |
| ‘Antenatal Care’           |                  |
| ANC register or form is correctly completed: 1) numbering (recording) the correct case 2) Fill all the information (column) required | 14 (44) |
| Weighing scale available and calibrated at zero (weight for ANC alone) | 12 (38) |
| ANC sheets for FOSA shows the administration of iron sulphate, folic acid and Mebendazole | 11 (34) |
| ANC performed by qualified personnel (verbatim) | 11 (34) |
| Book of the ANC (for mom) available—at least 10 (verbatim) | 10 (31) |
| ‘Maternity Care’           |                  |
| Analysis of 5–20 (randomly selected) partographs | 23 (72) |
| All deliveries are carried out by qualified personnel | 19 (59) |
| Sufficient water with antiseptic soap and liquid antiseptic in delivery room | 19 (59) |
| Presence of proper maternity equipment (e.g. sterile clamp, maternity beds, treated bed net, fetoscope, blood pressure machine, aspirator) | 19 (59) |
| Delivery room is in good condition (e.g. walls are made of solid material, are not cracked, and are plastered and painted, cement floor is not cracked; windows have glass and curtains, doors are in working condition, light 24/7, clean) | 14 (44) |
| Correct disposal of waste in the delivery room (e.g. dustbin, security box for sharps, placenta bucket, antiseptic solution) | 11 (34) |
| **Newborn and child health** |                  |
| Vaccination record available and completed (EPI/Immunization registry) | 18 (56) |
| Baby height and weight scale available and in working condition | 14 (44) |
| Integrated Management of Childhood Illness is applied correctly | 12 (38) |
| Under 5 (years of age) services (EPI, growth monitoring, curative care, health promotion) are available every day (at least 5 days a week) | 12 (38) |
| Available equipment and supplies for newborn care (1% Tetracycline eye ointment; Vitamin K) | 10 (31) |

*Indicators that appear in 30% or more of programmes.
with an average of 125 indicators and tend to grow in length over time. Second, the vast majority of indicators are related to the Donabedian structure type and very few are of the outcome type. The cross-tabulation of the Donabedian and author-derived categorization reveals that—as may be expected—indicators for resource availability are almost exclusively structural, while those for clinical effort and knowledge and non-clinical effort are of the Donabedian process or outcome type. Third, a majority of indicators assess resource availability rather than facility management or provider effort or knowledge. Fourth, checklists that evolved over time within the same implementation tended to increase in number of quality indicators. Fifth, there is a core set of MNCH indicators that is used across programmes that mostly focus on the antenatal, intrapartum and childhood periods with a lack of indicators for the post-partum period.

The overwhelming dominance of structural indicators (80% of all indicators) could be due to the costs of collecting process and outcome indicators, especially related to knowledge and effort-based indicators. The PBF quality checklists are also rather broad, especially given that the many PBF programmes focus the financial incentives on maternal and child health. The quality checklists capture a larger construct of facility quality, possibly in part to avoid adverse consequences of the targeted incentives for other service areas. Regarding the overall balance of capturing different dimensions of quality for different clinical areas and the trade-offs between the cost and relevance of data collection, our results suggest room for improvement and innovation as PBF programmes and quality assessments evolve. These findings and indicator sets may reflect the low baseline of infrastructure and resources in many settings where these programmes are implemented. They could also reflect common quality problems or policy goals. As our convenience sample was sourced from a small number of donors, this similarity could also reflect shared programme designs and/or learning across contexts. Overall these findings could indicate the need for better practical guidance on indicator selection and the appropriate length of checklists.

Our results also contribute to the debate on how much of the quality of care is being measured in PBF quality checklists. Rennams et al. (2016) point out that in Rwanda, every time the quality checklist was amended, the scores would drop relative to the previous quarter, potentially indicating that the quality checklists were not representative of overall quality. Similarly, the strong focus on structural indicators in the quality checklists in our sample, and the many process indicators which are management-focused and not directly related to the clinician-patient interaction, is indicative of a potentially ineffective blend of Donabedian indicator types to assess quality of care. This raises the question of how alternative blends of indicator types perform with regards to the validity in measuring quality of care at different levels of development of health systems. It could be argued that our quality checklist sample, with the dominance of structural indicators, represents the groundwork for quality of care.

A related question is whether these checklists are sufficiently flexible within and across countries. Our results indicate substantial homogeneity at least across MNCH indicators across countries and, by definition, an individual checklist applies to all facilities of a particular level in a country. The substantial heterogeneity in facilities implies that it will be very easy for facilities with good baseline performance to achieve ‘high’ quality without much effort—and vice versa for those facilities with lower baseline performance or in more challenging contexts. Adapting to this heterogeneity requires a balance of maintaining incentives and feasibility of improvements. There are clearly difficult administrative questions to tackle in developing more than one quality checklist for a country; however, with increased decentralization in many settings, allowing regional authorities more leeway in adapting to their context could make sense.

Our findings also indicate that a low share of indicators focus on effort related to the clinical encounter. This could reflect practical or financial constraints to routinely collect indicators that arise from doctor–patient interactions. As programmes evolve and shift away from structural measures there is scope for new technologies, such as phone surveys, to collect more clinical or patient-based metrics, including patient satisfaction. In comparison, the large share of Facility Management indicators most likely reflects the ease of verification of those indicators (e.g. meeting minutes, forms filled/filed etc.) as well as the desire of PBF practitioners to improve the operations of health facilities.

Our review did not establish whether indicators are appropriate and adequate to meet the intended policy outcomes in a specific context. In practice, PBF programmes must strike a balance between collecting detailed quality measures, and operational and financial constraints. The data collected by PBF are foremost used to execute payments rather than to comprehensively measure quality of care—an issue long recognized by PBF experts (Meessen et al. 2006; Rusa et al. 2009). However, our descriptive analysis suggests that there is scope to rationalize checklists and a need for further research to better understand how providers respond to quality checklists in order to improve programme design.

Limitations
Our analysis has several limitations. First, because of practical constraints our data are based on a convenience sample of checklists that was based on document reviews and communication with implementers and donors. The lists and collection of checklists remains incomplete both due to the sampling approach and from incomplete submission of checklists from country implementers and affiliated donors. As a result, our findings may not be representative of PBF programmes currently in operation. Second, in several instances (notably for Kenya and Zanzibar) we received multiple checklists that were undated but appeared to be revisions of the same document over time. In these cases, we used the version which appeared to be the most recent. Third, some indicators may straddle multiple Donabedian or author-derived categories, and coders may have categorized some indicators differently. We resolved such conflicts through discussion within the author team. We also recognize that the choice of framework for analysis—in this case, Donabedian—yields a distribution of indicators that may be categorized very differently under other frameworks. For example, under the WHO framework (Tunçalp et al. 2015), very few indicators would appear under Patient-centred Care, Equitable Care and perhaps Safe Care. Finally, while we report on the count and weight of indicators, there may be other factors that determine an indicators’ practical importance, such as the ease with which the facility can effect improvements.

Conclusions
Health facility-based PBF programmes offer an appealing approach to targeting specific quality shortfalls and advancing progress toward the SDGs related to high quality healthcare coverage. Our assessment of checklists from existing programmes highlights that currently most indicators focus on structural issues and resource availability. As our interpretation of ‘quality’ in the context of the SDGs evolves and there is increasingly robust evidence of the
potential of PBF to affect change, there is scope to rationalize and evolve the quality checklists of these programmes to help achieve this goal.

Supplementary Data

Supplementary data are available at Health Policy and Planning online.

Acknowledgements

We would like to thank the implementers and donor representatives who contributed PBF program manuals and quality checklists. From the World Bank Health Results Innovations Trust Fund, we thank Dinesh Nair and colleagues (Susanna Hayrapetyan, Toni Lee Kaguru, Ha Thi Hong Nguyen, Tayyeb Masud, Nadir Abdellatif Mohammed, Shunsuke Mabuchi, Maud Jiquoss, Omer Zang, Paul Jacob Robyn, Andrew Sunil Rajkumar, Hadia Samaha, Gyogy Bela Fritsche, Ayodeji Oluwole Oduotula, Itigak Asabe Sarah Dimka, Kanako Yamashita-Allen, Rifaat Hasan, Tekabe Ayalew Belay, Rousselie F. Lavado, Anessa Arur, Farai Sekeramayi-Noble). We also express appreciation for contributors from USAID, Beverly Johnston and Anna Coburn, in addition to Jean Kagubare, Isabel Bodemeyer and Fannie Kachale.

Disclaimer

The views and opinions expressed in this article are those of the authors and not necessarily the views and opinions of the United States Agency for International Development.

Funding

This research project is made possible through Translating Research into Action, TRAction, and is funded by United States Agency for International Development (USAID) under the cooperative agreement number GHS-A-00-09-00015-00.

Conflict of interest statement. None declared.

Notes

1. The modalities studied in the referenced review include accreditation, the COPE (client-oriented, provider efficient) methodology, Standards-based Management and Recognition (S-BMR), improvement collaboratives, supervision, and low- and high-intensity training.
2. The database of indicators we assembled is publicly available at: http://www.tractionproject.org/resources/results-based-management-performance-based-incentives-quality-care/multi-country.
3. In a few cases it was unclear whether an indicator referred to ‘clinical effort’ or not, e.g., comparing information in the register (non-clinical) with that on a health card (potentially related to clinical care). We assigned those cases to the ‘knowledge and non-clinical effort’ category.

References

Chaudhury N, Hammer J, Kremer M, Muralidharan K, Rogers FH. 2006. Missing in action: teacher and health worker absence in developing countries. The Journal of Economic Perspectives: A Journal of the American Economic Association 20: 91–116.

Das A, Gopalan SS, Chandramohan D. 2016. 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 16: 321.

Das J, Gerler PJ. 2007. Variations in practice quality in five low-income countries: a conceptual overview. Health Affairs (Project Hope) 26: w296–309.

Das J, Hammer J, Leonard K. 2008. The quality of medical advice in low-income countries. The Journal of Economic Perspectives: A Journal of the American Economic Association 22: 93–114.

Das J, Holla A, Mohpal A, Muralidharan K. (2013). Quality and accountability in healthcare delivery: audit evidence from primary care providers in India (Policy Research working paper; no. WPS 7334). Washington, DC: World Bank Group. http://documents.worldbank.org/curated/en/959771468000899235/Quality-and-accountability-in-healthcare-delivery-audit-evidence-from-primary-care-providers-in-India, accessed 5 December 2016.

Donabedian A. 1988. The quality of care. How can it be assessed? JAMA 260: 1743–8.

Fritsche GB, Soeters R, Meessen B. (2014). Performance-Based Financing Toolkit. Washington, DC: World Bank. License: Creative Commons Attribution CC BY 3.0.

Gergen J, Josephson E, Coo M et al. 2016. Incorporating quality of care in performance-based financing: How it is being done in 32 programs across 28 countries. Global Health: Science and Practice, Accepted.

Gergen J, Josephson E, Coo M, Shi, S, Madhavan S, Bauhoff S. 2017. Quality of Care in Performance-Based Financing: How It Is Incorporated in 32 Programs Across 28 Countries. Global Health: Science and Practice 5: 90–107.

Health Results Innovation Trust Fund. (2015). Achieving Results for Women’s and Children’s Health (Progress Report). Washington, DC: World Bank Health Results Innovation Trust Fund.

Institute of Medicine Board on Global Health; The National Academies of Sciences, Engineering, and Medicine. (2015). Improving Quality of Care in Low- and Middle-Income Countries: Workshop Summary. Washington, DC: National Academies Press (US).

Institute of Medicine (US) Committee on Quality of Health Care in America. (2001). Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press (US).

Lee E, Madhavan S, Bauhoff S. 2016. Levels and variations in the quality of facility-based antenatal care in Kenya: evidence from the 2010 service provision assessment. Health Policy and Planning 31: 777–84.

Meessen B. (2014). Results based financing “the Messy way”. http://www.health4africa.net/2014/07/results-based-financing-the-messy-way/, accessed 5 December 2016.

Meessen B, Musango L, Kashula J-P, Lemlin J. 2006. Reviewing institutions of rural health centres: the Performance Initiative in Butare, Rwanda. Tropical Medicine and International Health 11: 1303–17.

Musgrove P. (2011). Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and A Short Glossary. Washington, DC: World Bank.

Renmans D, Holvoet N, Orach CG, Criel B. 2016. Opening the “black box” of performance-based financing in low- and lower middle-income countries: a review of the literature. Health Policy and Planning 31: 1297–309.

Rusa L, Ngirabega J, de D, Janssen W et al. 2016. Levels and variations in the quality of facility-based antenatal care in Kenya: evidence from the 2010 service provision assessment. Health Policy and Planning 31: 777–84.

Tuno¨alp Q, Were WM, MacLennan C et al. 2015. Quality of care for pregnant women and newborns—the WHO vision. BJOG: An International Journal of Obstetrics and Gynaecology 122: 1045–9.

World Bank. (2016). Kyrgyz Republic: Performance Based Payments for Maternal and Neonatal Health: Quantitative Baseline Survey Report. https://www.rbfhealth.org/sites/rbf/files/Kyrgyz%20Republic%20Performance%20Based%20Payments%20Audit%20Maternal%20and%20Neonatal%20Health_0.pdf, accessed 5 December 2016.