Levels and variations in the quality of facility-based antenatal care in Kenya: evidence from the 2010 service provision assessment

Elizabeth Lee, Supriya Madhavan and Sebastian Bauhoff

Abstract

Quality of care is emerging as an important concern for low- and middle-income countries working to expand and improve coverage. However, there is limited systematic, large-scale empirical guidance to inform policy design. Our study operationalized indicators for six dimensions of quality of care that are captured in currently available, standardized Service Provision Assessments. We implemented these measures to assess the levels and heterogeneity of antenatal care in Kenya. Using our indicator mix, we find that performance is low overall and that there is substantial variation across provinces, management authority and facility type. Overall, facilities performed highest in the dimensions of efficiency and acceptability/patient-centeredness, and lowest on effectiveness and accessibility. Public facilities generally performed worse or similarly to private or faith-based facilities. We illustrate how these data and methods can provide readily-available, low-cost decision support for policy.

Key words: Global health, health care facilities, quality of care, variations

Introduction

As many low- and middle-income countries (LMIC) continue to make significant improvements in expanding access to health care services, policy-makers are increasingly cognizant of the need to improve the quality of these services. A growing body of evidence suggests that quality shortfalls are substantial even for basic health care services (Das and Gertler 2007; Das et al. 2008; Jha et al. 2013), prompting calls to measure and address these gaps. On a global level, the quality of health care services is featured in the Sustainable Development Goals, the successors of the Millennium Development Goals (LCSDSN 2015). Many countries are actively exploring how to ensure or raise the quality of care, with policies ranging from supporting access to private sector providers that may offer higher quality to providing financial rewards to providers for achieving quality targets.

However, there is currently little systematic documentation of the levels and variations in quality of care that could support the design and deployment of effective policies (Scott and Jha 2014).
One practical constraint to analysing the quality of care is the short-age of reliable facility data in developing LMIC (Glassman and Ezeh 2014).

In this article, we illustrate how data from large-scale facility sur-
veys currently available for 14 countries can be used to gauge the
quality of antenatal care (ANC) services and provide basic decision
support to policy makers. We focus on ANC services which have
relatively less complex standards of care than other aspects of health
care services. We derived 14 indicators in six dimensions from exist-
ing quality of care frameworks and applied these measures to the
2010 Kenya SPA to examine their empirical distributions across 
geography, facility type and management authority.

We discuss how such analyses can be used to inform the design
of quality improvement policies, such as Kenya’s nascent results-
based financing (RBF) initiative and the country’s recent drive to
provide free maternal care in public facilities. RBF schemes—
programmes that pay for pre-defined, measurable outputs as
opposed to inputs related to service delivery—have grown in scale
and scope across LMICs over the last decade. With support from
donors, countries have begun adapting RBF programmes to improve
quality of care in addition to increasing quantity of care. Mean-
while, demand-side interventions that could substantially in-
crease utilization could lead to deterioration in the quality of care.

This article illustrates how existing facility surveys can be used to
carry out low-cost quality of care analyses. Such analyses can help identify potential areas of concern that could be addressed in the
design of policies like Kenya’s free maternal care programme, or
that could be monitored more closely during implementation. For
informing programmes like RBF, baseline quality of care analyses
can help determine which issues to target and how to design incen-
tive structures to encourage overall improvements and address heter-
ogeneties across facilities. Throughout, we describe the potential
and challenges of using surveys like the SPA to evaluate the quality
of ANC services.

### Data and methods

To adequately capture the complexities of quality of ANC services, we reviewed quality frameworks proposed by Donabedian (1988),
the World Health Organization (2006), the Organisation for
Economic Co-operation and Development (OECD, Kelley and
Hurst 2006) and the Institute of Medicine (2001). We selected
the six dimensions of quality of care proposed by the WHO as the op-
erative conceptual framework for this study. These dimensions align
with the other frameworks and are the most appropriate for quality
measurement in the developing country context. WHO provides the
following definitions for the six dimensions: (1) effectiveness: ‘de-
delivering health care that is adherent to an evidence base and results
in improved health outcomes for individuals and communities,
based on need’; (2) efficiency: ‘delivering health care in a manner
which maximizes resource use and avoids waste’; (3) accessibility:
‘delivering health care that is timely, geographically reasonable, and
provided in a setting where skills and resources are appropriate to
medical need’; (4) acceptability/patient-centeredness: ‘delivering
health care which takes into account the preferences and aspirations
of individual service users and the cultures of their communities’; (5)
equity: ‘delivering health care which does not vary in quality be-
cause of personal characteristics such as gender, race, ethnicity, geo-
 graphical location or socioeconomic status’; and (6) safety: ‘delivering health care which minimizes risks and harm to service
users’.

### Indicators

We identified ANC quality indicators that can be constructed at the
facility level through a review of peer-reviewed and grey literature, as well as ANC-specific indicators used commonly in quality check-
lists of RBF programmes in LMICs, in some cases these overlapped.
We included the latter because indicators in RBF programmes are
clearly of interest to policy makers and have been operationalized
for LMICs. Together, these indicators were mapped into six quality
care dimensions using the WHO dimension definitions and to the
SPA survey. We selected candidate indicators in terms of reproduc-
ibility from SPA data and to ensure RBF indicators were represented.
The mapping and the final list of 14 indicators are presented in
Supplementary Appendix S1, which also describes how some indica-
tors were adapted to the Kenyan context and data.

### Data

We calculated the individual indicators using the 2010 Kenya SPA, a
facility-based cross-sectional survey. The SPA is a standardized sur-
vey supported by the Demographic and Health Surveys Programme
of USAID and routinely administered by ICF International and in-
country partners.

The 2010 Kenyan SPA consists of several survey instruments
that are administered concurrently and generally on the same day: a
facility audit, health care worker interviews and client observations
and exit interviews for ANC, family planning and sick child visits
(for comprehensive documentation and questionnaires see NCAPD
2011). The facility audit entails an interview with the in-charge of
the facility or the senior-most staff member; direct verification of the
presence of certain commodities, equipment and amenities; and veri-
fication of their use. The health worker interview asks about specific
training received and services offered by the health worker, as well
as workers’ opinions on the work conditions at the facility. The
client–provider observations are based on observation protocols spe-
cific to ANC, family planning services and sick child care. The exit
interviews capture clients immediately after a consultation and veri-
fy the services received and client opinions regarding those services.
The exit interviews are also specific to ANC, family planning and
sick child services.

Data from the Kenyan SPA are representative at the provincial,
facility type and management authority levels (NCAPD 2011). Seven
hundred and three facilities or 11% of all Kenyan facilities were
randomly selected, with 695 successfully surveyed. Voluntary
counselling and testing (VCT) centres, maternity and hospital facili-
ties were oversampled. Health workers were sampled to cover a
range of services provided; for observations and exit interviews cli-
ents were systematically sampled. We selected the Kenyan SPA as
the data are recent and there is large geographic variation in mater-
nal mortality rates. Figure 1 shows the distribution of public, private
for profit and faith-based facilities across Kenya. The SPA was im-
plemented between January and May 2010.

In this study, we included facilities if they reported providing
ANC services and also completed the ANC portion of the facility
audit, had non-missing data and were not an HIV VCT facility. We
excluded VCTs since they do not routinely provide ANC services.
We excluded NGO/private not-for-profit facilities from the final
analytic sample due to small sample sizes. Because many facilities
have missing data for one or more indicator, the analysis sample
consists of 144 out of 545 non-VCT and non-NGO/private not-for
profit facilities with ANC services with a completed questionnaire
(26%). Table 1 shows the basic characteristics of the facilities in the
analytic sample. Compared with other eligible facilities that had
missing data, the analytical sample has relatively more district or sub-district hospitals and more public facilities (see Table 1). In addition to data from the facility audit, we used client observations \((n = 654)\) and exit interviews \((n = 638)\) for facilities in the analytic sample with median counts of five per facility for both.

### Analysis

We first examined the individual indicators as well as dimension-specific indices constructed as equally weighted averages of the respective indicators. All indicators and indices were constructed at the facility level. The indicators were constructed to range from 0 to 1, with higher values representing better quality. We calculated medians and interquartile ranges (IQR) across all facilities and by facility type (national and provincial hospitals; district, sub-district and other hospitals; health centres and clinics; dispensaries and maternity clinics), management authority (public, private for profit, faith-based), and geography (eight provinces). For the comparisons across facilities, we focus on medians rather than means, as indicators and indices were either non-normally distributed or categorical.

To facilitate overall inter-dimension comparisons, we calculated coefficients of variation, defined as the ratio of the dimension standard deviation to its mean. Finally, we calculated an overall quality of care score for each facility by equally weighting and averaging five dimensions (all but equity) calculated on the facility level.
We address the sixth dimension—equity—separately from the five other dimensions and outside of the overall facility score due to data limitations. The SPA provides data on individual clients only in the observation and exit interviews, and information on client characteristics is limited. We therefore compared median scores for patients with low and high education, defined as no or primary schooling, and secondary school or above, respectively. We used three indicators which are collected in patient observations and which are linked to corresponding patient exit interviews, where education is recorded: ANC physical examination service score, patient satisfaction recorded after the ANC consultation, and whether the visit was conducted by qualified ANC provider. Because the number of observations/exit interviews at each facility is relatively small, we did not construct facility-level measures for education.

Limitations
The limitations of our study highlight practical challenges in using the SPA data to analyse quality of care. First, we focused on ANC, which has limited impact on mortality and may not be reflective of other aspects of quality maternal and neonatal health care. Similarly, our set of 14 indicators for quality ANC, though selected based on our review of quality frameworks and the ANC literature, may not adequately cover all aspects of quality. In particular, the selected indicators are mostly structure- and process-oriented as the SPA is not designed to capture outcome indicators. Further, our ability to measure the dimension of equity is impacted by the limited demographic data available in the SPA: age and education level derived from ANC exit interview data. We analysed equity at the exit interview level but analysed the remaining five quality dimensions at the facility level.

Low ANC observation/exit interview counts for many facilities may impact the stability of indicators derived from observations. SPA data collection procedures set a maximum limit of five observations and associated exit interviews per provider of ANC services, and 15 for any given facility (NCAPD 2011). In practice, the obtained range for ANC observations from each facility varied from 0 to 10. Where possible, we constructed indicators from the facility audit to avoid the instability issue; for some indicators this was not feasible, e.g. for patient satisfaction levels.

Design choices in the SPA may affect survey responses in several ways. For example, data based on exit interviews can be subject to ‘courtesy bias’, limiting variation in scores and introducing an upward bias. Observed visits and subsequent exit interviews may also induce the Hawthorne effect if providers know clients will be surveyed after the consultation and, as a result, lead to inflated client exit interview scores (Glick 2009).

We included 144 of 545 possible ANC facilities from the original SPA sample once data were restricted to eligible facilities with data for all component indicators. Data for the indicators of evidence-based maternal care, physical exam score, infection prevention score and whether a visit was conducted by a qualified ANC provider were missing somewhat more frequently than data for other indicators (see Supplementary Appendix Table S1.1). To address this issue, we performed a sensitivity analysis excluding the indicator for evidence-based maternal care and found qualitatively similar results for the unaffected dimensions, i.e. the dimensions effectiveness, efficiency and acceptability/patient-centeredness. We also addressed concerns about the analytic sample size by comparing results for component and composite indices, where we allowed the sample size to vary across quality dimensions, compared with the overall sample restricted within and across dimensions. These results, as well as coefficient of variation results, were qualitatively similar for the analytic sample compared with results of the sensitivity analysis.

We conducted two assessments of the missing data and the potential implications (see Supplementary Appendix Tables S3.1 and S3.2). First, we examined facility characteristics as predictors of missing data in multivariate regressions for all eligible facilities ($n = 545$) across each dimension and the overall quality score. The dependent variable was a binary indicator of whether the facility does or does not have data, and the independent variables were facility characteristics (region, managing authority, facility type and whether facility has: regular management meetings, back-up generator with fuel, catchment estimation data available, any first ANC visit observed). In general, we found that facilities with routine management meetings were consistently less likely to have missing data across dimensions, and that private-for-profit-managed facilities were more likely to have missing data compared with publicly managed, except for the accessibility dimension. Second, for each quality dimension, we used t-tests to compare mean values of the subset of facilities with quality measures for all domains vs the larger set of facilities with complete data for the specific dimension tested but missing data for alone or more of the other dimensions. We found no differences in quality between facilities with complete and incomplete data at the dimension level.

Our analytic sample may be subject to selection bias for several reasons. First, although the SPA’s sampling frame is the official Master Facility List it may not adequately represent all facilities in Kenya, e.g. smaller private providers. Second, as noted, there is non-random missing data across facilities, which determines which facilities are included in the analytic sample.

We constructed certain dimensions from overlapping indicators, primarily because several composite indicators included iron and/or folic acid tablets. Further, we constructed the overall quality of care score from dimensions comprised of several overlapping indicators. This overlap could introduce correlations among the affected dimensions and may give slightly more weight to the relevant indicators in the overall score. Finally, given we did not utilize regression adjustment in our analyses, some of the observed variation, e.g. across facility types, may be related to other observable as well as unobservable factors.

Results
Overall performance on quality of care indicators and dimensions
Table 2 shows the 14 quality indicators as components of five quality of care dimensions measured on the facility level (excluding equity, which we discuss below).

Overall and relative to the maximum score of 1.00, facilities performed well on most indicators. The two lowest performing indicators were: ANC physical exam score (median score of 0.23) and infection prevention score (0.50). Because we first constructed scores at the facility level, and these scores may be continuous, the figures in Table 2 should not be interpreted as the share of facilities meeting a certain standard but rather as the score of the median facility. For instance, the median score of 0.71 for the indicator ‘number of days per month ANC services provided’ implies that the median facility in our sample offered these services 71% of days in a 4-week (28 days) month.

Quality of care dimensions varied considerably in terms of median performance. Facilities performed highest in the areas of acceptability/patient-centeredness (median score of 1.00) and efficiency (0.92).
Conversely, performance was lowest for the effectiveness and accessibility dimensions, with respective median scores of 0.55 and 0.68. Safety had a middling performance (0.75). There was substantial variation across indicators within a dimension, e.g. the poor performance for the effectiveness dimension is primarily driven by a lack of adequate ANC physical exam services. The coefficients of variation for the effectiveness dimension is primarily driven by a lack of adequate ANC physical exam services. The coefficients of variation for the effectiveness dimension is primarily driven by a lack of adequate ANC physical exam services.

Variation across and within provinces
Quality of care varied substantially across provinces (Figure 2; for means and IQR see Supplementary Appendix Table S2.1). Six out of eight provinces performed relatively poorly or only moderately well in terms of effectiveness, with Central and Nairobi being exceptions. Almost all were high performers in the efficiency dimension (median scores above 0.8), with the limited sample for Nairobi scoring lowest (0.81). Five provinces scored 0.68 or 0.73 on accessibility and the remainder of provinces performed better overall in this dimension (scores larger than 0.79). Most provinces were moderate to high performers in terms of acceptability/patient-centeredness, except Northeastern which performed worse than all other provinces (median score 0.67). Six provinces scored 0.75 in the safety dimension; Nairobi and Central were higher with median safety scores of 0.88. We also found that provinces differed in their relative performance across quality of care dimensions. For instance, Central performed well overall, whereas Northeastern performed well in accessibility and efficiency, but poorly or only moderately well in other dimensions.

Variation across and within facility types
Figure 2 also shows scores across four facility types: national/provincial hospitals; district/sub-district/other hospitals; health centres/clinics; and dispensaries/thernities. All facility types consistently performed well in terms of efficiency (median scores above 0.90). Most performed poorly in terms of effectiveness (median scores of 0.75 or lower) with the exception of national/provincial hospitals (median score 0.83). Facility types varied in their performance for the other dimensions. For instance, district/sub-district/other hospitals and dispensaries/thernities performed comparatively poorly on effectiveness.

Performance also varied within facility types. Within district hospitals and lower-level facilities, effectiveness had the lowest scores; the highest scoring dimensions included efficiency and acceptability/patient-centeredness. Within national and provincial hospitals, the efficiency dimension had the highest median score; these facilities performed moderately well across all other dimensions.

Variation across and within management authorities
Finally, Figure 2 depicts quality dimensions grouped by three management authority types in Kenya. Public facilities performed worse than or about the same as private for profit or faith-based facilities. They performed poorly in the accessibility dimension, relative to other management types. Faith-based facilities performed better or about the same as other facilities. The dimensions of highest consistent performance across management authorities were acceptability/patient-centeredness and efficiency. Within management authorities, facilities run by faith-based organizations performed consistently well in terms of efficiency, accessibility, and acceptability/patient-centeredness and moderately well on effectiveness and safety. Inter-dimension variation was greatest for public facilities.

Variation by education level (equity dimension)
To approximate the equity dimension of the WHO framework, we calculated median scores by low/high education level for three indicators that are available in the ANC observation/patient exit interview data, where patients’ education is also recorded. We found that overall median scores for the two groups were similar for all three measures calculated on the ANC client level: ANC physical examination service score (0.75), patient satisfaction post-ANC...
consultation (1.00) and whether the visit was conducted by qualified ANC provider (1.00; detailed results not shown).

Discussion

Quality of health care is quickly emerging as a major concern in many LMIC, particularly as efforts to expand access to care are gaining traction. In this article, we constructed quality of care indicators from Kenyan facility data to explore the level and heterogeneity in ANC quality. Our findings indicate low overall performance (on our specific set of measures) in effectiveness, and comparatively high performance on the efficiency and acceptability/patient-centeredness dimensions. However, we also found substantial variation across Kenyan provinces, facility type and management authority, with public facilities generally underperforming relative to faith-based and private for profit facilities.

A possible explanation for the finding of good performance in the equity dimension is that the available indicators already performed well, so that there is little scope for variations. For instance, almost all patients reported being seen by an adequately trained provider. These findings from the SPA are supported by the 2008–09 household Demographic and Health Survey (KNBS, ICF Macro 2010), which suggests that the proportion of women ages 15–49 receiving ANC from a skilled provider differs little by education level. However, almost one-quarter of women with no education did not receive ANC services for the most recent birth, compared with only 3% of women with secondary education or better. One explanation could be that low- and high-education households have
different access to care but, once in the facility, receive comparable care from providers (as measured in the SPA). Thus, this finding also highlights the sensitivity of the results to the choice and availability of indicators.

Lessons from using existing facility surveys to measure the quality of ANC care

Our study illustrates the promises and challenges of operationalizing quality of care frameworks on standardized facility surveys, such as the SPA. On the one hand, these data are readily available (and more SPAs are planned) and can facilitate quality assessments and inform the design and scale-up of health policies. They can also serve as diagnostic tools and provide baseline measures against which to measure progress. On the other hand, we had to exclude or modify some accepted facility-based quality metrics in order to operationalize SPA data, and there was substantial missing data. This latter challenge suggests caution in interpreting or extrapolating our specific findings to all of Kenya. We also found variation across indicators within a particular dimension, indicating that the choice (and availability) of quality indicators matters for quality assessments. Similarly, the SPA does not cover several issues that are known to be important for quality, such as provider effort (Das and Gertler 2007; Das et al. 2008). Overall, our study therefore also suggests that existing assessment tools may benefit from harmonization and a redesign to rationalize and optimize tracking of meaningful measures that map to existing quality of care frameworks (Lee et al. 2015). This approach is endorsed in the Roadmap for the Measurement and Accountability for Health Summit held in June 2015 (USAID, World Bank, World Health Organization 2015). A harmonized instrument may also allow for more frequent and high-quality data collection, and could help track quality of care over time.

Implications for designing RBF programmes

The observed variations in quality of care have implications for designing interventions to improve quality, such as RBF which has emerged as a popular approach for increasing provider performance, especially for primary care. Kenya piloted an RBF scheme in Samburu County in 2011 with support from the World Bank, and is expanding to public facilities across 20 northern, rural counties, with the intent to explore eventual integration of private-side facilities including faith-based facilities (World Bank 2013).

In the design of RBF programmes, there are a number of central decisions for consideration which are related to the payout function; for example, what indicators to include and how to reward the rewarded indicators. Specific choices include whether to pay for exceeding thresholds or pay on a linear schedule, and whether to pay directly for quality or scale quantity payments by broader measures of quality (Basinga et al. 2011; Sherry et al. 2015).

Our methods and findings can help inform these decisions. First, programmes should address the quality as well as the quantity of care, as some dimensions of quality are consistently low. Second, the degree of inter-facility variation can provide guidance for determining the relative financial incentives, e.g. rewarding more generously those dimensions and indicators that perform very poorly (to encourage attention) or very highly (to defray potentially high marginal costs of further improvements). Third, baseline variations across facilities imply that it is challenging to set a threshold that simultaneously incentivizes high and low performers. A suitable payout function could involve graduated payments or only pay for improvements above facilities’ baseline performance. Fourth, although variation across provinces could be accommodated by a regionally differentiated RBF, there are substantial variations within each geographic area which also need to be addressed. Finally, our findings indicate scope for interventions to complement the RBF programme. We captured basic systemic quality problems—such as number of days ANC services are offered—which may be costly to rectify and for which RBF incentives may be too small to nudge providers into action. Similarly, the consistently low performance in the effectiveness dimension could be addressed in a larger, non-RBF effort.

Implications for demand-side interventions

Our findings can also contribute to designing demand-side interventions and tracking their effects on service quality. Kenya introduced free maternal care in public facilities in 2013 amid concerns that these facilities may find it challenging to adequately respond to the expected increase in demand (Bourbonnais 2014; Chuma and Maina 2014). Institutional delivery rates in Kenya have already increased significantly from 42.6% in Kenya Demographic and Health Survey 2008-09 to 61.2% in Kenya Demographic and Health Survey 2014, but there is a scope for further growth. In other settings, the combination of rapidly introduced demand-side interventions alongside stagnant supply-side conditions has led to decreases in quality (Chaturvedi et al. 2014). Our analysis of facility data collected prior to the start of this initiative indicates potential challenges and could be used to identify ‘hotspot’ areas such as effectiveness, which may need particular attention. Further, we identified groups of facilities which may struggle to maintain or increase quality—a particular concern for public facilities which are likely to experience the largest increase in demand and which already perform comparatively poorly on most quality dimensions.

Conclusion

Our findings suggest that policies need to address and account for heterogeneity in quality of ANC. In Kenya, the good performance of some facilities (for at least some of their patients) indicates scope for improvement in this context: raising all facilities to the level of best performance should be feasible and would lead to significant overall gains in quality. There is some evidence that changes in payment modalities could facilitate such gains (Das and Gertler 2007; Basinga et al. 2011), possibly in tandem with other interventions such as targeted training or investments in facility improvements. In general, there is a need for more systematic and harmonized data on the quality of care in LMIC.

Supplementary data

Supplementary data are available at HEAPOL online

Disclaimer

The views and opinions expressed in this paper are those of the authors and not necessarily the views and opinions of the United States Agency for International Development. The replication files for this analysis (StatA do-files) are available at: https://dataverse.harvard.edu/dataverse/cgdev.

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.

References

Basinga P, Gertler PJ, Binagwaho A, et al. 2011. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. The Lancet 377: 1421–1428.

Bourbonnais N. 2014. Implementing Free Maternal Health Care in Kenya: Challenges, Strategies and Recommendations [Internet]. Kenya National Commission on Human Rights. http://www.knchr.or.ke/Portals/0/Ecosoc Reports/Implementing%20Free%20Maternal%20Health%20Care%20in%20Kenya.pdf, accessed 9 April 2015.

Chaturvedi S, Randive B, Diwan V, De Costa A. 2014. Quality of obstetric referral services in India’s JSY cash transfer programme for institutional births: a study from Madhya Pradesh province. PLoS One 9: e96773.

Chuma J, Maina T. 2014. Free Maternal Care and Removal of User Fees at Primary-Level Facilities in Kenya: Monitoring the Implementation and Impact—Baseline Report. Washington, DC: Health Policy Project, Futures Group. http://www.healthpolicyproject.com/pubs/400_KenyaUserFeesBaselineReportFINAL.pdf, accessed 9 April 2015.

Das J, Gertler PJ. 2007. Variations in practice quality in five low-income countries: a conceptual overview. Health Affairs 26: w296–w309.

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

Donabedian A. 1988. The quality of care. How can it be assessed? JAMA 260: 1743–1748.

Glassman A, Ezeh A. 2014. Delivering on the Data Revolution in Sub-Saharan Africa. Washington, DC: Center for Global Development.

Glick P. 2009. How reliable are surveys of client satisfaction with healthcare services? Evidence from matched facility and household data in Madagascar. Social Science and Medicine 68: 368–379.

Institute of Medicine. 2001. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press.

Jha AK, Lartigoitia I, Audera-Lopez C et al. 2013. The global burden of unsafe medical care: analytic modelling of observational studies. BMJ Quality and Safety 22, 809–815. http://qualitysafety.bmj.com/content/early/2013/08/29/bmjqs-2012-001748.abstract, accessed 9 April 2015.

Kelley E, Hurst J. 2006. Health care quality indicators project conceptual framework paper. Organisation for Economic Co-operation and Development. Report No.: DELSA/HEA/WD/HWP(2006)3.

Kenya Demographic and Health Survey 2008–09. Calverton, MD: KNBS and ICF Macro. 2010. https://www.dhsprogram.com/publications/publication-FR229-DHS-Final-Reports.cfm, accessed 9 April 2015.

KNBS, ICF Macro. 2010. Kenya Demographic and Health Survey 2008–09. Calverton, Maryland: Kenya National Bureau of Statistics (KNBS) and IFC Macro.

LCSDSN. Indicators and a Monitoring Framework for the Sustainable Development Goals Launching a data revolution for the SDGs [Internet]. Leadership Council of the Sustainable Development Solutions Network; 2015 Feb. http://unsdsn.org/wp-content/uploads/2015/03/150218-SDSN-Indicator-Report-FEB-FINAL.pdf, accessed 12 March 2015.

Lee E, Madhavan S, Bauhoff S. Quality of Care Frameworks and Antenatal Care Indicators: a Systematic Selection Process. Working Paper, 2015.

NCAPD 2011. Kenya Service Provision Assessment Survey 2010. [Internet]. Nairobi, Kenya: National Coordinating Agency for Population and Development (NCAPD) [Kenya], Ministry of Medical Services (MOMS) [Kenya], Ministry of Public Health and Sanitation (MOPHS) [Kenya], Kenya National Bureau of Statistics (KNBS) [Kenya], ICF Macro. . http://dhsprogram.com/what-we-do/survey/survey-display-347.cfm, accessed 12 February 2015.

Scott KW, Jha AK. 2014. Putting quality on the global health agenda. New England Journal of Medicine 371: 3–5.

Sherry T, Bauhoff S, Mohanan M. 2015. Paying for performance in health care: results from the randomized roll-out of Rwanda’s National Program [Internet]. Report No.: Economic Research Initiatives at Duke (ERID) Working Paper No. 136. http://papers.ssrn.com/sol3/papers.cfm?abstract_id¼2170393.

USAID, World Bank, World Health Organization 2015. Health measurement and accountability post-2015: a common roadmap. Draft for Consultation, March 25, 2015 [Internet]. M4Health. http://m4health.b4success.org/docs/support-documents/common-roadmap-draft2-consultation-2015_03_25-edited.pdf?sfvrsn¼12, accessed 9 April 2015.

WHO 2006. Quality of Care: A Process for Making Strategic Choices in Health Systems. Geneva, Switzerland: World Health Organization.

World Bank 2013. Kenya—Health Sector Support Project : Restructuring and Additional Financing [Internet]. World Bank. http://documents.worldbank.org/curated/en/2013/12/18620311/kenya-health-sector-support-project-restructuring-additional-financing, accessed 4 October 2015.