Quality of self-assessment and medical record data in facilities with quality improvement teams: A cross-sectional study in India

Introduction

The USAID Applying Science to Strengthen and Improve Systems (ASSIST) Project is supporting 349 facilities in 27 districts in six states in India to improve reproductive, maternal and child health care. We are building capacity of teams of front-line workers to use quality improvement approaches to address priority problems in the facilities we support. As part of this approach, the teams collect and review data about the care that they are providing and use this data to make decisions about what to change and to learn whether the changes they are making are improving care. Some data are available from existing sources. Other data need to be collected by improvement teams.

To better understand the quality of the data teams are collecting and using to guide their improvement work, we undertook a study to assess the validity of the data compared to observation of actual practice. We compared three sources of data: 1) data from direct observation by external observers (observed data); 2) data recorded in individual patient records (recorded data); and 3) aggregated data used by improvement team members to track their performance (self-assessment data) that comes from individual patient records and other data sources in the facility.

The main questions we sought to answer were: 1) How well do recorded and self-assessment data compare to observed data? 2) What types of errors in data quality exist in self-assessment and recorded data? Our goal was to learn more about data quality to provide better guidance to front-line workers on which data to use for quality improvement as well as to identify data problems that need addressing.

Methods

To identify a 5% difference between the recorded and observed data, we planned to observe 444 deliveries (74 per state). Sample size was calculated using Ssize (Lwanga and Lemeshow 1998). We planned to observe 55 deliveries at one district hospital (DH) and 19 at one community health centre (CHC) in each state (3:1 ratio) and to observe half between the day shift (8am to 8pm) and the other half during the night shift. Twelve data collectors or observers, all having a medical background, were recruited for the study and received one day of training on what and how to observe, data entry, and ethical concerns.

The observers visited the facilities in November 2014. The facilities had been receiving support from the USAID ASSIST Project to carry out quality improvement activities for 6-9 months. The observers recorded data during deliveries and also collected data from the medical records. Data on the following activities were collected: 1) administration of oxytocin within one minute following delivery, 2) use of sterile cord clamp, 3) drying and wrapping of the newborn, 4) breastfeeding started within one hour of birth, 5) vitamin K given to infants within 6 hours of birth, 6) mother's condition assessed between 0-30 minutes after delivery, 7) mother's condition assessed between 31-60 minutes after delivery.

MAY 2015

This short report was prepared by University Research Co., LLC (URC) for review by the United States Agency for International Development (USAID) and authored by Enisha Sarin, Subir K. Kole, Rakesh Kumar, Edward Broughton, and Nigel Livesley of URC under the USAID Applying Science to Strengthen and Improve Systems (ASSIST) Project. USAID ASSIST is managed by URC under the terms of Cooperative Agreement Number AID-OAA-A-12-00101. URC's global partners for USAID ASSIST include: EnCompass LLC; FHI 360; Harvard University School of Public Health; HEALTHQUAL International; Initiatives Inc.; Institute for Healthcare Improvement; Initiatives Inc.; Johns Hopkins Center for Communication Programs; and WI-HER LLC. The contents of this report are the sole responsibility of URC and do not necessarily reflect the views of USAID or the United States Government. For more information on the work of the USAID ASSIST Project, please visit www.usaidassist.org or write assist-info@urc-chs.com.

Observed patients and the staff conducting the delivery were informed about the risks and benefits of participating in the study. Written consent was obtained, and the observer was present at the labor room for each delivery that was observed from the beginning to two hours after the delivery. During this period of observation, the observer had a checklist to record whether or not a particular practice was done and the time when it was done. The checklist also contained key demographic information on the staff conducting the delivery, which included years of experience and professional level. After one or two days (to give time to the staff to fill out registers and patient records), the observer checked the available hospital records and entered the data in a prepared Excel sheet which was sent to the central office for further analysis. All hard copies of the checklists were also sent to the office after data collection.

Ethical considerations

The study was approved by the University Research Co., LLC institutional review board. Official permission was also obtained from facility and district authorities. Written consent was obtained from all observed staff and patients. Male observers were only assigned to facilities in which male doctors already conducted deliveries. The checklists and other data collection tools did not carry any names but used unique identifying numbers for both patients and staff, thus precluding the possibility of identifying any patient or staff.

Data analysis

Results were entered into an Excel database and sent to the USAID ASSIST office in Delhi where the data was cleaned and analyzed using Stata/SE 13.1. Our main analysis was the difference between observed and recorded data and between observed and self-assessment data. The differences between these were tested using a chi-squared test. We also determined sensitivity (proportion of cases where practice was accurately reported) and specificity (proportion of cases where non-practice was accurately reported). We calculated the proportion of discordant results between observed and recorded data in which records show over- or under-estimation in reporting. We then conducted Fisher's exact test to determine whether or not this proportion is 50% as would be expected by chance if errors occurred randomly.

Results

A total of 461 observations were conducted; of these, 362 (78.5%) of observed deliveries occurred in district hospitals, 234 (50.2%) during the day, 324 (70.3%) were attended by staff nurses, and 130 (28.2%) of attending clinicians had between 2-5 years of experience (**Table 1**).

Table 1: Location, time, and staff characteristics of observed deliveries

Facility level	Percent of observed deliveries
• District hospital	362 (78.5%)
• CHC	99 (21.5%)
Shift	Percent of observed deliveries
• Day (8am – 8pm)	234 (50.8%)
• Night (8pm – 8am)	227 (49.2%)
Staff cadre	Percent of observed deliveries
• ANM	11 (2.4%)
• Staff nurse	324 (70.3%)
• Doctor	126 (27.3%)
Years of experience	Percent of observed deliveries
• < 2 years	124 (26.9%)
• 2 to <6 years	130 (28.2%)
• 6 to <10 years	98 (21.3%)
• 10 or more years	109 (23.6%)

Compliance with standards of care was high for oxytocin administration (81%), drying and wrapping (92%), cord care (98%), and vitamin K administration (89%); moderate for breastfeeding (64%); and low for maternal assessment at 0-30 minutes (35%) and 30-60 minutes (22%).

Comparison of observed, recorded, and self-assessment data

Self-assessment data were more accurate measures of performance than recorded data (**Table 2**). However, most self-assessment data were significantly different from observed data.

Table 2. Comparison of practice as measured through observed, self-assessment, and recorded data

Indicators	Observed data (%)	Self-assessment data (%)	Recorded data (%)
Oxytocin within one minute	81	98*	96*
Dry and wrap	92	98*	24*
Cord care	98	100	32*
Breastfeeding within one hour	64	96*	68
Vitamin K injection	89	99*	99*
Asphyxia managed	8	5*	3*

*Differs from observed data at $p<0.001$.

Accuracy of medical records

Medical record data were less accurate in measuring observed performance than self-assessment data. Observed practice was correctly recorded in medical records between 31-90% of the time. Sensitivity ranged from 35-100%, and specificity from 8-100% (**Table 3**). Errors of under-reporting were more common in drying and wrapping, cord care, and asphyxia management indicators, while over-reporting was more common with oxytocin administration, maternal assessment, and vitamin K administration. Both were equally common with early breastfeeding.

Table 3. Comparison of observed and recorded data

Indicators	Practice observed (%)	Sensitivity (%)	Specificity (%)	Correct recording (%)	Over-reporting (%)	Under-reporting (%)
Oxytocin	80.9	99.7	20.4	84.6	15.2*	3.9
Dry and wrap	92.2	99.1	9.94	31.0	0.2	68.8**
Cord care	98.0	31.4	55.6	31.9	0.8	67.2**
Breastfeeding	64.3	62.8	23.0	48.6	27.6	23.9
Maternal assess 0-30 min	35.0	83.0	62.2	69.4	23.2*	8.7
Maternal assess 30-60 min	22.3	51.0	67.9	64.0	24.9*	11
Vitamin K	89.3	99.3	8.2	89.6	9.8*	0.7
Management of asphyxia	8.0	35.0	100.0	89.4	0.0	5.2**

*Signifies self-enhancement or reporting when practice not observed.
**Signifies self-diminishment or not reporting practice though observed.

Sensitivity was high for oxytocin, vitamin K administration, and drying and wrapping the newborn. Specificity where non-practice was accurately reported was high for asphyxia management and moderately high for maternal assessment. For oxytocin, maternal assessment at 30 and 60 minutes, and vitamin K, there was statistically significant bias toward self-enhancement that made the health staff's performance appear better. On the other hand, drying and wrapping, cord care, and asphyxia management showed that although the practice was observed, it was not reported.

Discussion

There was good performance of routine delivery practices except for assessments of post-partum women and early breastfeeding. Self-assessment data more accurately reflected observed performance than data directly taken from patient records. The only exception was breastfeeding for which self-assessment data was substantially less accurate than recorded data. While self-assessment data were more accurate than recorded data, they over-estimated actual performance. In three of five indicators, self-assessed values were within 10% of the observed data. Breastfeeding was over-reported by more than 30%. One possible explanation for this is that data on early breast feeding is the only indicator we examined that is also recorded in the facility's health management information system. There may be incentives in the health system that lead to performance being over-reported through official data systems. Adherence to oxytocin administration within one minute of delivery was over-reported by 15%. Observers felt that this was because oxytocin was often given within 2-3 minutes but missed the one minute cut-off in their data collection tool. There is no evidence that giving oxytocin within 1 minute is any more effective than giving it within the first few minutes of delivery (Soltani et al. 2010).

Based on this study, self-assessment data in facilities with quality improvement teams is sufficiently accurate to guide quality improvement work when performance is below 80% adherence with standards but is not likely to be sufficient to guide decision making when performance is above that level. Using individual medical records to help improve data quality of self-assessments does not look like an effective strategy since recorded data are less accurate than self-reported data. Data quality for drying and wrapping and cord care indicators improved when teams focused on improving those indicators (data not reported in this paper). It is therefore possible that as teams continue to work on a specific goal, data quality related to that goal also improves. The need to improve data quality seems especially important in indicators such as early breastfeeding which are reported in official government data systems.

There are a number of limitations to this study. It was carried out in only 12 facilities and may therefore not be an accurate estimate of data validity in facilities in other parts of India. There is also the possibility that performance was higher due to the Hawthorne effect of observation (Leonard and Masatu 2006). Given that the observer effect would be expected to improve both performance and data recording, we do not believe that this influenced our results to any great degree. Another limitation is that observers only recorded activities as performed if the activities occurred in a specified time period. For example, oxytocin administration, early breastfeeding, and early and late maternal assessment were only scored as performed by the observers if they happened within one minute, 60 minutes, and 0-30 and 30-60 minutes, respectively. In addition, observers followed each delivery for two hours after birth, and vitamin K is supposed to be given within the first 24 hours. If it was given between 2-24 hours it would not have been observed. This limitation may have led to observation under-estimating actual performance.

In conclusion, this study demonstrates that self-assessment data leads to over-reporting of health worker performance but in most cases the degree of over-reporting was less than 10% compared to observations. We believe that this level of accuracy is sufficient to provide enough information on quality of care to guide improvement efforts when adherence with standards is below 80%. When teams are trying to improve performance from 80-100%, additional interventions to improve data quality are likely to be required. It will be important to test whether data quality improves as quality improvement teams are in place for a longer period of time.

References

1. Lwanga SK, Lemeshow S. 1998. *Sample size determination in health studies: A practical manual*. Geneva: World Health Organization. Software version 2.0 by KC Lun and Peter Chiam: National University of Singapore, Singapore
2. Soltani H, Hutchon DR, Poulose TA. 2010. Timing of prophylactic uterotonics for the third stage of labour after vaginal birth. *Cochrane Database of Systematic Reviews* 2010, Issue 8. Art. No.: CD006173. DOI: 10.1002/14651858.CD006173.pub2.
3. Leonard K, Masatu MC. 2006. Outpatient process quality evaluation and the Hawthorne Effect. *Soc Sci Med* 63(9),2330-40.

USAID Applying Science to Strengthen and Improve Systems

University Research Co., LLC • 7200 Wisconsin Avenue • Bethesda, Maryland 20814-4811 • USA

TEL 301-654-8338 • FAX 301-941-8427 • www.usaidassist.org • assist-info@urc-chs.com