

Box 9	Cross-tabulation: Stillbirth delivery	39
Table 7	Indicators that Met Both Validation Criteria	40

Appendices

Annex A, Table 1	List of Indicators & Measured Coverage (Matched Data)
Annex B, Table 1	Validation Results for Indicators With at Least 5 Counts per Cell

Study Background

Global monitoring of the percentage of women who have received quality maternal health services is crucial to inform efforts to guide the scale-up and allocation of resources to address preventable maternal deaths. Given difficulties in measuring maternal deaths, the proportion of births attended by skilled health personnel and the proportion of births delivered in health facilities have become the most widely used indicators for measuring progress towards maternal health goals. Coverage rates of ‘skilled attendance’ and ‘institutional deliveries’ have become benchmarks for quality of maternal health care routinely tracked by national and international agencies.

Such reliance on these indicators requires the assumption that women delivering in an institution with the assistance of a skilled attendant will also have access to comprehensive essential services, such as emergency obstetric care and lifesaving commodities including uterotonics, magnesium sulfate, and antibiotics.^{1,2} Given discrepancies in the quality of care between providers and facilities, however, specifying the actual interventions that a woman receives is warranted to provide a more informative and accurate assessment of the coverage of key interventions.

Little previous research has been conducted on this topic. To our knowledge, the two most widely used proxy indicators – skilled attendance at birth and institutional delivery – have not been empirically validated or systematically evaluated. In addition, there have been few attempts, until very recently, to test the feasibility of collecting data on specific elements of the care received by women during labor and delivery.³⁻⁷

In response to a call to increase reliable maternal health information in the *Lancet “Manifesto for Maternal Health”*, a 2013 PLOS Medicine special issue, reported in partnership with the Child Health Epidemiology Reference Group (CHERG), includes three quantitative studies in this area. These studies examine the validity of women’s reports of: 1) the indications for cesarean sections in Ghana and the Dominican Republic⁸, 2) indicators of care received by women and their newborns during labor, delivery, and the postnatal period in Mozambique⁹ and (3) indicators of care received by women and newborns in rural China.¹⁰ In these studies, women’s reporting of events during their labor and delivery is compared against a reference standard, either medical records or observation in a health facility. In addition, a few small qualitative studies have examined whether specific events during labor and delivery (e.g., cord cutting) were understood and recalled by women, whether women were able to recall their sequence and timing, and the terms used to describe them.¹¹⁻¹³

The present study extends this research by comparing women's self-reports of maternal and newborn service provision during the intrapartum and early postnatal periods prior to discharge from a hospital facility to third party observation at the time of delivery. The study also seeks to provide insight into factors (e.g., participant characteristics, type of delivery, instances of complications or other events) that may influence the accuracy of recall. The results of the study will inform the recommendation of a select number of indicators that have the potential for valid and reliable measurement as well as integration into routine population-based and facility-based data collection systems.

Kenya and Mexico were chosen as study sites given the variations between the two countries in the status of maternal health and the coverage and organization of maternal health services. We chose two countries with very different levels of skilled attendance and institutional delivery so that we would be able to validate indicators in a wide range of situations. This report presents results from the Kenya study site.

Study Overview

OBJECTIVE

The goal of the study is to improve monitoring of the quality of maternal health care through identifying, developing, and validating maternal health indicators that can be practically applied in population-based surveys. The main question addressed by this research is: Can accurate information on the quality and content of maternal health care received by women during labor and delivery be self-reported by women in a survey format? The two specific objectives of the study are as follows:

1. To assess the validity of women's reports of skilled birth attendance; and
2. To assess the validity of women's reports of indicators of the quality of routine obstetric and immediate postnatal service delivery.

INDICATOR SELECTION

To identify quality care indicators for maternal health to be validated, a landscaping scan was conducted from April to July 2012. The scan focused on indicators currently in use or proposed for use, including both population-based and facility-based indicators. Indicators were identified by performing a key word search of electronic databases including: PUBMED, POPLINE, JSTOR and EMBASE. Additional searches were conducted of publications of organizations known for their involvement in measuring maternal health care, such as WHO, UNICEF, UNFPA, MCHIP, AMDD and IMMPACT, and by searching reference lists of identified papers and reports. Key search terms included: maternal health, safe motherhood, quality of care, indicator, valid, skilled attendant, neonatal, perinatal, obstetric, and intrapartum. No studies were excluded because of language or date of publication.

An indicator matrix was developed to organize findings. From an identified 2,505 documents, 71 provided information on indicators for assessing quality in maternal healthcare. This listing was used to select a set of indicators for validity testing (see Annex A). These indicators were considered the most commonly used or critical variables for assessing the quality and coverage of maternal care. The observation and interview questionnaires were translated into the appropriate local dialects and underwent minor modifications to improve local understanding and clarity for participants.

STUDY DESIGN

To accomplish stated objectives, the validation study employs a facility-based design with comparisons against a gold standard. Specifically, women's reports on indicators of the quality of maternal health care they received are compared against third party observations of the care provided at the time of labor and delivery using a structured checklist (gold standard).

Third party observations were chosen as the reference standard since they are likely to reflect all facets of the care-giving process. In the event that additional information or clarification was needed, medical and facility records were also checked. Women's self-reports of the services they received at the time of labor and delivery were gathered via exit interviews prior to their discharge from participating hospital facilities, *New Nyanza Provincial General Hospital*, Kisumu District, Kenya, now known as *Jaramogi Oginga Odinga Teaching and Referral Hospital* (JOOTRH), and *Kiambu County Hospital*, Kiambu District, Kenya.

ETHICAL CLEARANCE

The protocol was approved by the Population Council's Institutional Review Board in May, 2013 (IRB Protocol 594), and the Kenya Medical Research Institute in July, 2013 (KEMRI Non-SCC 395). No participants were enrolled into the study until ethical approval was obtained from both ethics committees.

Kenya Medical Research Institute (KEMRI)
KEMRI/National Ethics and Review Committee
P.O. Box 54840-00200
Nairobi, Kenya
Division of Reproductive Health,
Ministry of Health
Tel: +254-02-2722541 / +254-02-2713349

Methods

RESEARCH SITE & STUDY POPULATION

Study population: The study population consists of women whose births were documented by study data collectors at participating facilities between the months of July and September, 2013. Women aged between 15-49 years who were admitted for labor at study facilities and who consented to study participation were eligible for inclusion. The study population also includes the providers who attended participating women in labor and delivery and whose labor and delivery care was observed by study data collectors.

Study locations: *New Nyanza Provincial General Hospital, now known as Jaramogi Oginga Odinga Teaching and Referral Hospital (JOOTRH), Kisumu District, Kenya; and Kiambu County Hospital, Kiambu District, Central Province, Kenya.*¹

¹ Kenya was formerly divided into 8 administrative provinces and subdivided into districts. In August 2010, a new constitution was enacted which ordered the replacement of the provinces with 47 semi-autonomous counties, now considered 'sub-counties'. At the time the study was conducted, use of the new terminology had not yet begun. Described geographical areas are comparable to the 2008-2009 DHS report for Kenya.

In Kenya women were recruited from the above hospitals, both of which are public governmental facilities providing comprehensive obstetric care. Obstetric patients in these hospitals are women with normal pregnancies who are self-referred for admission as well as women with high-risk pregnancies who are referred from other public hospitals and health centers. The New Nyanza Provincial General Hospital, now JOOTRH, delivered 5099 women in 2013 and the Kiambu County Hospital delivered 6851 women in 2013.

Obstetric patients who visit the two hospitals are largely from the surrounding Kiambu and Kisumu districts. In Kiambu district, formerly classified as within Central Province, 73% of births are facility-based, while in Kisumu district, formerly classified as within Nyanza Province, 44% of births are facility-based.¹⁴ This compares to a national rate of 43% facility-based delivery.¹⁴ In Kenya, education level is correlated with facility-based delivery, and nationally the vast majority (84%) of women with no education deliver at home.¹⁴ The highest level of education attained by women in Central Province and Nyanza Province is approximately equivalent. Specifically, in Central Province 64% of women attended at least some or completed primary school and 11% have no education, while in Nyanza Province 67% of women have attended at least some or completed primary school and 13% have no formal education.

RECRUITMENT PROCEDURES

Data Collectors

Two types of data collectors were involved in this study:

- 1) Researchers selected to observe labor and delivery consultations were drawn from a pool of registered nurse/midwives with a minimum of three years' experience in a maternal, newborn and child health (MNCH)/maternity unit and who had previous experience conducting similar Council research (observers had recently participated in the *Promoting Dignified Care During Childbirth* project). Observers were carefully selected to work under the supervision of the Kenya Population Council study team. Observers received training on how to observe activities in the labor and delivery ward and to check any inpatient notes, records or partographs for further information or clarification if necessary. Observers wore uniforms with name badges to identify themselves as researchers.

- 2) Data collectors conducting the client exit interview were diploma/degree holders in a social science. Data collectors were carefully selected to work under the supervision and support of the Population Council study team.

Data collectors were asked to work where they were not known personally, to respect privacy and ensure respondents' anonymity. All data collectors were fluent in the local language spoken in the labor ward (Kiswahili, Dholuo or Kikuyu) and received training on appropriate procedures for conducting ethical research as detailed below.

Data Collector Training

Training of data collectors took place over 4 days and consisted of an introduction to the research objectives, procedures for interviewing/observing women, and appropriate practices for conducting ethical research. Specifically, training included a detailed review of the study protocol and each element in the questionnaire and the observation checklist so that all team members fully understood the research aims and methods, meanings of the instrument components and how to record responses and observations. All data collectors were trained on ethical research practices, including focused sessions regarding the meaning and process of informed consent, as well as the importance of participant privacy and procedures for maintaining confidentiality of all collected information. Study interviewers were trained to listen and observe intently, without displaying any judgmental attitudes towards information they received. Interviewers also received information on how and where to refer clients in the event they required additional support. Observers received training on procedures for being unobtrusive and on locating themselves toward the head of the client rather than the foot when possible. The medically trained observers recruited for the study were also accustomed to being in a facility and followed appropriate professional conduct for the setting. Time was spent on 'role-playing' both the interviews and observations, with the trainees taking turns acting as interviewer or observer and client or provider. In addition, the team pre-tested study tools in order to practice scenarios expected in the actual field work. The team also pre-tested the programmed exit tool to confirm that the paper tool was programmed as it appeared on the PDA.

Recruitment and Informed Consent Procedures

All women who met study eligibility criteria and were admitted to the antenatal ward were asked by a health provider if they were willing to meet with a researcher. Due to Ministry of Health guidelines, women who were more than six centimeters dilated were not approached for consent. Women who agreed to speak with a researcher were provided with a description of the study, procedures, benefits

and risks. Specifically, women were asked for permission to allow a researcher to stand in the corner of the room during their labor and delivery to record the actions of their provider(s) using a checklist. Women were informed they had the right not to be observed and could stop the observation at any time. Women were also asked if they were willing to be interviewed by a study data collector about their labor and delivery experience prior to hospital discharge. At this point in the recruitment process, written informed consent was obtained in the woman's native language (Kiswahili, Dholuo, or Kikuyu). In Kenya, pregnant adolescents between ages 15-17 years are considered "emancipated minors" because of their pregnancy status (by having assumed adult responsibility for procreation). As such, informed consent was also obtained for participants younger than 18 years. Verbal consent was obtained from women again at the time of exit interview prior to conducting the interview. If the woman agreed, the provider indicated to the data collector whether the woman's medical status was acceptable for her to be interviewed.

Informed consent was also obtained from providers at participating health facilities. Council staff first introduced the study to the facility director in-charge and to all participating staff in group meetings. All staff who provide care during labor and delivery, and who were identified by the hospital/obstetrics and gynaecology (OBGYN) director as such, were asked to take part in the study. Written informed consent was obtained from all providers prior to observation by study data collectors.

Data Collection Procedures

Data collection in Kenya took place between July and September, 2013. One observer and one interviewer constituted a data collection pair. Seven pairs worked as a data collection team in each facility location. The team worked in two shifts (day and night). The day shift had 3 to 4 observers and a similar number of interviewers. The night shift had only observers (3 or 4). As data collection progressed, the number of women admitted for labor was lower than anticipated and the number of interviewers was reduced to 3 per facility.

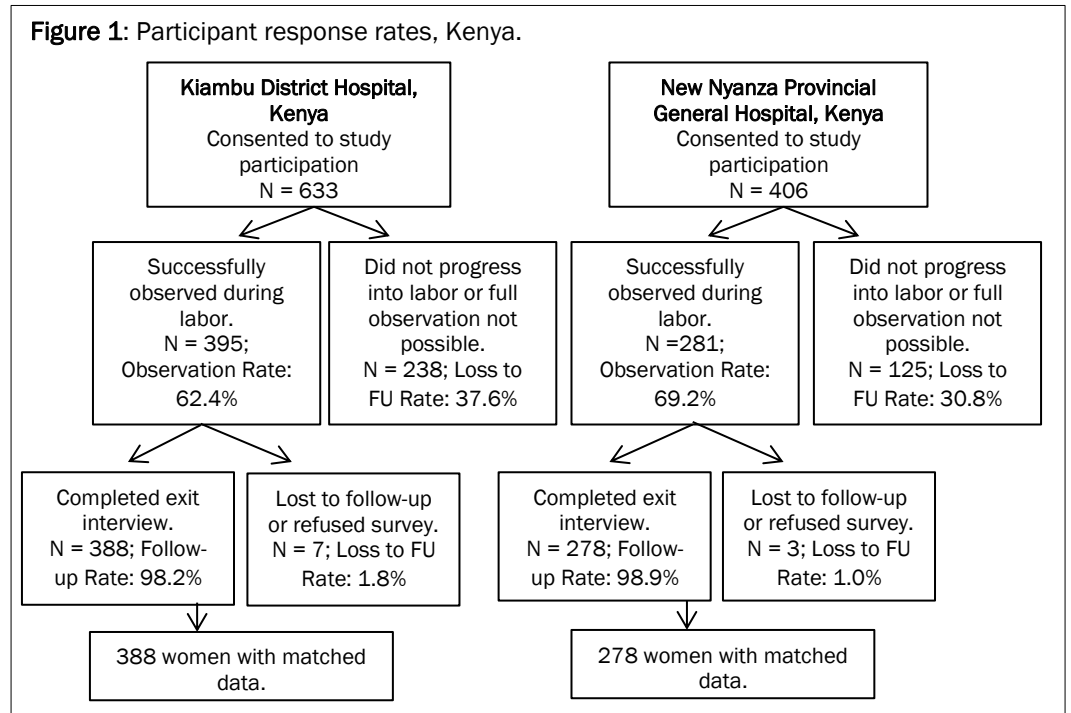
Providers were observed monitoring women during labor and delivery by third party study observers. The trained researchers observed from the maternity admission room and labor and delivery room. Observations included the birth itself and interactions between the client and the provider before, during and up to 1 hour after delivery. Observers used a structured checklist to record their observations of the actions taken by the provider. Exit interviews with participating women used a structured, in-person questionnaire. In each facility, a private place was identified to ensure tht women felt comfortable being interviewed.

The target sample size for participants in Kenya was 600 women (approximately 300 per facility). This was calculated to detect at least a 10 percentage point difference between the direct observations and women's reports, regardless of where the 'true' prevalence lies. A 20% increase was also added to account for anticipated loss to follow-up for a separate study to re-interview women approximately 12 months following delivery.

Results

STUDY SAMPLE

A total of 1,039 women admitted for labor at participating study facilities were recruited to participate. Of those who consented to participate, a total of 676 women were successfully observed during labor (Kiambu n=395, Kisumu=281). As only women in early labor could be



approached for consent, data collectors sought to consent all eligible women in the antenatal ward admission room, which at times included women admitted for purposes other than delivery. As a result, approximately one-third of women (38% in Kiambu and 31% in Kisumu) were not observed either because they did not progress into labor or they progressed rapidly into labor and full observation was not possible. Among those whose births were observed, 666 women (Kiambu n=388; Kisumu n=278) completed an exit interview. A small percentage of women observed during labor refused participation in the exit interview (2% in Kiambu and 1% in Kisumu) at the time of interview. A total of 666 observer reports and client exit interviews were accurately matched and could be analyzed (N=388 Kiambu; N=278 Kisumu). See **Figure 1** for a detailed flow chart of participant enrollment.

SAMPLE CHARACTERISTICS

All participants were recruited from the surrounding Kiambu and Kisumu districts. Approximately 7.9% of women whose births were documented as part of the study were referred from another facility.

Descriptive statistics on the sample's socio-demographic and delivery characteristics are presented in **Table 1**. The mean age of women was 24 years (Std dev. \pm 5.1) and ranged between 15 and 42 years of age. The vast majority of women were currently married (78%), and less than half of participants had secondary education or higher (45%). Approximately half of women had no prior births (50%), with six prior births being the highest prior parity. Women who delivered in the Kisumu facility were on average 0.9 years younger in age (mean difference: -0.9 ± 0.4 (margin of error), $p < 0.05$) and had on average 0.2 mean units greater educational attainment (mean difference: -0.2 ± 0.1 (margin of error) $p < 0.01$, assessed on four point scale where 1=no education and 4= greater than secondary school). There were no differences among other key demographic indicators, including previous births, marital status and type of delivery (vaginal or cesarean section).

TABLE 1: Sample Background Characteristics by Facility Location

	Total Matched %	Kiambu %	Kisumu %	P-Value
Age				p=0.021
15-19	14.5	11.9	18.1	
20-24	42.0	43.1	40.6	
25-29	28.9	29.4	28.6	
30-34	8.5	8.8	9.1	
35-39	5.7	6.2	4.7	
40+	0.4	0.5	0.4	
Prior Parity (Total # Live Births)				p=0.408
0	50.2	49.7	51.3	
1	26.5	28.8	22.9	
2	14.0	13.2	14.9	
3	6.0	5.2	7.3	
4 or more	3.3	3.1	3.6	
Educational level				p=0.007
None	10.2	10.3	10.2	
Primary	43.6	45.9	41.2	
Secondary	29.2	33.2	24.1	
Higher	16.2	10.6	24.5	
Marital status				p=0.128
Single, never married	14.7	9.8	21.5	
Married	78.1	80.2	75.2	
Living together	5.3	7.5	2.2	
Separated	1.8	2.6	0.7	
Widowed	0.2	0.0	0.4	
Type of delivery				p=0.679
Vaginal	86.6	87.0	85.9	
C/S	13.4	13.0	14.1	

Maternal & Newborn Outcomes

Nearly all deliveries observed were singleton newborns, with 22 sets of twins and 2 sets of triplets. A total of 1% of babies (n=9) were stillborn, and <1% (n=3) babies died within the first hour of birth.

The majority of women whose births were observed came to the facility because they had planned to do so (92%), while 8% of women reported coming to the facility because they had a problem.

Complication(s) were observed to occur among 11% of women. The majority of complications occurred

before delivery (52%), with approximately one-quarter occurring at the time of delivery (24%), or during the post-partum interim (24%). The most common type of complication experienced was hemorrhage (APH+PHH) (5% of women), followed by prolonged labor (>12 hours), (4% of women). 1% of women with complications experienced more than one.

The majority of observed births occurred by vaginal delivery (87%), while 13% were by cesarean operation. Women who underwent cesareans did not differ from those who had vaginal delivery by age, prior parity, education level or marital status (Table 2). However, experiencing a complication was associated with the type of delivery. Specifically, of women who had a cesarean operation, 38% had a complication, while only 7% of women with vaginal delivery had a complication.

Service delivery coverage

The list of indicators selected for validity testing is presented in Annex A. The table also describes the matched prevalence of each indicator by women’s self-report (reported prevalence) and observer report (‘true’ prevalence), excluding “Don’t Know” responses. The observed prevalence

TABLE 2: Sample Background Characteristics by Delivery Type

Variable	Vaginal	C/S	P-value
Age (Mean)	24.2± 0.41	25.3±1.0	0.064
Parity (Mean)	0.87± 0.09	0.79±0.20	0.498
Education level (Mean)†	2.51±0.07	2.54±0.20	0.826
Marital status, %			
Single, never married	14.4	16.1	
Married	78.5	75.9	0.573
Living together	5.5	4.6	
Separated/ Divorced	1.4	3.4	
Widowed	0.2	0.0	
Complication (Y/N), %	6.7	38.2	<0.01

Note: Margin of error reported for all mean values; † Highest school attended: 1=no education, 2=primary, 3=secondary, 4= higher than secondary school.

TABLE 3: Near Universal and Extremely Rare Indicators, Unmatched Data

Indicator	Observed Prevalence
Near Universal practices (>95%), matched data	
Woman’s HIV status checked	94.7
Fetal heart rate checked with fetoscope/doppler ultrasound	99.9
Provider wore high-level disinfected or sterile gloves for vaginal examination	99.9
Uterotonic administered within few minutes of delivery (via injection, IV medication, or oral/rectal tablets)	98.8
Controlled cord traction applied	98.9
Uterine massage performed after delivery of placenta	98.6
Position of mother at birth- on back	99.9
Health provider wore gloves during delivery of baby	99.8
Newborn immediately dried with towel/cloth following delivery	99.5
Newborn placed directly on mother’s skin and covered with dry towel on mothers’ abdomen	100.0
Newborn not placed skin-to-skin, was wrapped in dry towel	97.4
Baby weighed after delivery	100.0
Of those who received cesarean section, provider made decision	100.0
Rare practices (<5%), matched data	
Offered HIV test	1.7
Takes urine sample	1.4
Induces labor with a uterotonic	4.7
Support person/ companion present during birth	4.8
Something other than breastmilk given to baby within first hour of birth	1.1
Baby bathed within first hour after birth	0.1

of indicators ranged from extremely rare (<5%) to near universal (>95%). The broad range reflects that some preventive interventions were almost always implemented, while other harmful practices rarely, if ever, occurred, reflecting a generally high standard of care. Indicators meeting these criteria are listed in **Table 3**. Rare practices were likely to be harmful (e.g. baby bathed within first hour after birth, something other than breastmilk given to baby within first hour of birth). Of note is that an HIV test was rarely offered because the practice of first checking the woman’s HIV status by either asking the woman or consulting her record was near universal.

For all indicators, women and observers were given the option to respond “I Don’t Know”. The proportion of women who responded “I Don’t Know” to indicators was generally minimal (<4%). Indicators where the proportion of women who responded “Don’t Know” exceeded 5% are reported in **Table 4**.

TABLE 4: Indicators with High ‘Don’t Know’ Responses, Unmatched Data

Indicator	% “Don’t Know”
Woman Self-Report (>5% DK)	
Did the health provider(s) wash his/her hands with soap and water or use antiseptic before examining you?	29.5
Was your baby wrapped in a towel or cloth immediately after birth?	20.6
Was your baby dried off with a towel or cloth immediately after his/her birth?	8.4
In your first physical examination after delivery, did a health provider do a perineal exam?	9.8
Observer Report (>1% DK)	
Was anything besides breast milk given to the baby to drink within the first hour after birth?	5.1
Was baby bathed within first hour after birth?	2.4

The greatest proportion of women responded “Don’t Know” to the indicator, “*Did the health provider(s) wash his/her hands with soap and water or use antiseptic before examining you?*” A potential explanation for this lack of knowledge is that hand washing or antiseptic use may have occurred outside of the

woman’s view. Approximately 21% of women reported “Don’t Know” to the indicator “*Was your baby wrapped in a towel or cloth immediately after birth?*” (for babies not placed skin-to-skin on mothers’ chest). Receiving a cesarean section was significantly associated with not knowing if the baby was wrapped in a towel immediately after birth. Specifically, women who responded “Don’t Know” to this indicator were nearly 3 times as likely to have received a cesarean section relative to a vaginal delivery (OR: 2.7 ± 1.1, p<0.01).

Receiving a cesarean section was also significantly associated with not knowing if the baby was immediately dried following birth (8% ‘Don’t Know’ self-reported prevalence). Women who responded “Don’t Know” to this indicator were 15 times as likely to have received a cesarean section (OR: 15.3 ± 4.8, p<0.01). Potential explanations for the greater lack of knowledge are that some women who

underwent cesarean section in the Kenya facilities were given general anesthesia or had delayed recovery and may not have been aware of care received at this time, or that the baby was taken out of view of the mother. Cesarean delivery was not associated with not knowing whether a perineal exam was performed ($p=0.349$). It is possible that this indicator had a high ‘Don’t Know’ prevalence among women due to lack of understanding of the term “perineal”. Two indicators approached 5% “Don’t Know” prevalence by women: “Were you allowed to drink liquids or eat any food while in labor?” (5%) and “Did you experience any of the following complications during or after delivery?”(5%).

Observer responses of “Don’t Know” were minimal, and where they occurred were generally less than 1% of responses (**Table 4**). Two indicators had an observer report of “Don’t Know” higher than 1%. Both of these indicators related to immediate post-partum care for the newborn (if the baby was fed anything other than breastmilk in the first hour after birth and if the baby was bathed in the first hour of birth). The higher percentage of “Don’t Know” responses in part reflects hospital practices of taking the newborn outside of the room (and away from the view of the observer) if the mother had a complication and required care, or if the newborn required resuscitation.

VALIDITY OF QUALITY OF CARE INDICATORS

Analysis Approach

The general approach for the validity study relied on comparing women’s self-report of each quality of care indicator to its “true” classification according to the observer report (e.g., reference standard) at the time of facility-based delivery. Using this data, sensitivity and specificity analysis was conducted by first constructing two-by-two tables for each quality of care indicator. An indicator’s sensitivity is the proportion of actual positives (women who received care) who correctly identify as such (i.e. “true positive rate”). An indicator’s specificity measures the proportion of negatives (women who did not receive care) who correctly identify as such (i.e. “true negative rate”). For indicators where two-by-two tables had at least five counts per cell, the area under the receiver operating curve (AUC) and the inflation factor (IF) was also analyzed.

Receiver operating curve analysis is a valuable method to describe the accuracy of diagnostic tools by plotting the tradeoff between sensitivity (true positive rate) and its false positive rate (or 1- specificity). In practice, the area under the curve (AUC) represents the “average accuracy of a diagnostic test” and summarizes the accuracy of a test (sensitivity and specificity) by a single number.¹⁵⁻¹⁷ An AUC of 1.0 can be interpreted as a test with perfect accuracy, while an AUC of 0.5 represents a random guess.¹⁸ To assess the population-based validity of indicators, we also estimated each indicator’s inflation factor

(IF). Using an equation by Vecchio, each indicator’s estimated sensitivity, specificity is applied to its true prevalence (or observer report) to estimate the prevalence of an indicator that would be obtained using a population-based survey.¹⁹ By comparing the ratio of the estimated survey-based prevalence to its true population prevalence (observer report or reference standard), we will estimate the degree to which each indicator would be over- or under-estimated if assessed using a population-based survey.²⁰

A priori benchmark criteria for ‘valid’ indicators were an AUC>0.6 and IF between <1.25 and >0.75 and were informed by criteria previously used in the literature.^{17,18} Use of the two methods provides complementary data on indicator use in population-based surveys. Indicators with sufficient variation to allow for validity analysis are presented in **Annex B**. For indicators listed in the text, prevalence data refers to matched indicator data. Throughout the text, the AUC and its margin of error are also reported.

Tables throughout the text report varying N’s for the matched sample depending on whether the woman was observed during the indicator-specific period and whether the woman and observer gave a valid response for the indicator.

Facility Arrival & Initial Assessment

Type of facility

Of interest is whether women can accurately report on the type of facility where they delivered. Since both study facilities were public government hospitals (100% observed prevalence), the present study was not designed to assess whether women can accurately report on this indicator. However, cross-tabulation results show that of the women who gave a valid response to this indicator (i.e., excluding those who said “Don’t Know”), 98% correctly classified the type of facility as a public/ government hospital (**Box 1A, 1B**). These results should be interpreted with care as variation in collected data did not allow for robust analysis.

Box 1A. Descriptive frequencies: Type of facility.

Can you tell me the type of facility where you gave birth to your baby? (Self-report)

	Number	Percent
Public Sector		
Govt. hospital	638	98.00
Govt. clinic/health center	8	1.23
Govt. health dispensary	1	0.15
Other public sector		
Private Sector		
Private hospital	2	0.31
Private clinic	2	0.31
Private maternity home	0	0
Other private sector	0	0
Total	651	100

Box 1B. Cross-tabulation: Type of facility.

Can you tell me the type of facility where you gave birth to your baby?

		Observer Report (Number)		Total
		Other facility	Gov. Hospital	
Self-Report (Number)	Other facility	-	13	13
	Gov. Hospital	-	638	638
	Total	-	651	651

Initial assessment practices

Three indicators of the initial client assessment phase could be assessed. Other practices were near universal or rarely occurred and did not have enough variation for robust analysis. Assessed indicators in the initial client assessment were: (1) woman referred from another facility because of a problem, (2) provider washed his/her hands or used antiseptic before examination of woman and (3) woman's blood pressure was taken. No indicators met both study validity criteria. High rates of false reporting (1 - indicator specificity) among women occurred for provider hand washing/antiseptic use (67% of women whose provider did not sanitize hands falsely reported that they done so; 77% of women who did not have blood pressure taken on intake indicated they had received this care) (**Annex B**). Women's self-reports of coming to the facility because of a problem (rather than originally intending to come) had low sensitivity (25%), indicating not all referral cases were reported by women.

INITIAL ASSESSMENT: SUMMARY

No assessed indicators of facility-based care received at initial client assessment met study acceptability criteria for valid *measurement*.

Respectful Care

Five indicators reflected aspects of women-centered care (a proxy for respectful care), including whether the woman was: (1) allowed to drink liquids or eat, (2) encouraged or assisted to ambulate during labor, (3) encouraged or assisted to assume different positions in labor, (4) allowed to have a support person/ companion present during labor and delivery, and (5) whether a support person/ companion was actually present during labor and delivery.

Women tended to overreport the following indicators of respectful care: 1) allowed to drink liquids or eat and 2) encouraged/assisted to ambulate during labor. These indicators had low specificity (37% and 24%), indicating that a high proportion of women reported receiving this care when they had not (i.e., high false positive rate). In contrast, indicators on whether the woman received assistance assuming different positions in labor (sensitivity: 19%) and whether they were able to have a support companion present during labor and delivery (sensitivity: 24%) had low sensitivity, indicating not all women who received this care reported it.

Only one indicator whether a support person was actually present during the birth, met both acceptability criteria (AUC: 0.7346 ± 0.07, IF: 0.77).

PROVIDER RESPECTFUL CARE: SUMMARY

Only 1 of 5 assessed indicators, whether a support person was actually present during the birth, met study validity criteria. Results suggest women may be more able to accurately report observable indicators rather than rules or policies (which may be less explicitly stated or understood).

Induction / Augmentation of Labor

Induction of labor

Approximately 13% of all women reported “someone did something” to bring on their labor. 11% of all women reported they received an injection, IV line, or were given tablets to bring on their labor, while 5% of observers reported that women received a uterotonic to induce labor. The methods by which women indicated something was done are listed in **Box 2A**. We were unable to assess the

Box 2A. Descriptive frequencies: Induction of labor, method of administration.				
	Self-Report		Observer	
	(Of women who reported something done to bring on labor), what was done to bring on your labor?		(Of women with labor induced by uterotonic), route of administration.	
	Number	Percent	Number	Percent
Injection in thigh/buttocks	4	4.44	0	0.00
IV line inserted in arm	44	48.89	3	10.00
Given tablets	28	31.11	27	90.00
Membranes ruptured	7	7.78	NA	NA
Total	79	100	30	100

validity of receiving membrane rupture for induction of labor, as observer reports did not record the indication for receiving this intervention (i.e., whether performed for induction or augmentation of labor).

A two-item indicator in which women who reported “*something was done to bring on labor*” and who reported receiving a uterotonic (e.g., an injection, IV line, or an oral/vaginal/rectal tablet) to bring on labor was compared to observer report of the woman receiving a uterotonic to induce labor. The two-item induction of labor indicator did not meet both study validity criteria (AUC: 0.8049 ± 0.03, IF: 2.35). Given the relatively low overall frequency of women who received a uterotonic, we were unable to assess the accuracy of women’s report on the method through which the uterotonic was received.

Augmentation of labor

A greater proportion (45%) of women reported that “something was done to speed up or strengthen” their labor. 39% of all women reported that they received an injection, IV line, or were given tablets to strengthen their labor, while 22% of observers reported that women received a uterotonic for labor augmentation.

The most common method of uterotonic administration reported for augmentation of labor was an IV line inserted in the arm (according to both observer and women’s self-report). Given the relatively low observed prevalence of women whose labor was augmented by a uterotonic (22%), it was not possible to assess the accuracy of women’s report on the method of how the uterotonic was received. However, by combining all methods women reported for how uterotonic for labor augmentation was received (IM injection, IV line in arm, or tablets), we were able to compare the accuracy of women’s self-report of receiving a uterotonic for labor augmentation to the observer report (**Box 3A**). This indicator met the AUC only, suggesting it may be useful for individual-level classification but would be overreported at an aggregate level (AUC: 0.7169± 0.04, IF: 1.75).

Box 3A. Cross-tabulation: Augmentation of labor by uterotonic.
Injection in thigh/buttocks, IV line in arm, or tablets were received to speed up/ strengthen labor.

		Observer Report (Number)		Total
		No	Yes	
Self-Report (Number)	No	342	38	380
	Yes	143	102	245
Total		485	140	625

Uterotonic for induction or augmentation of labor

To investigate whether women could accurately report on receiving a uterotonic for induction or augmentation of labor, we constructed a combined indicator. This indicator met criteria for the AUC only (AUC: 0.7347 ± 0.04, IF: 1.61), suggesting the indicator may be useful for individual-level classification, but would overestimate the prevalence at an aggregate level.

Membrane rupture

3% of all women reported receiving membrane rupture (either for induction or augmentation of labor), compared to 42% of observers. Given the low sensitivity of the indicator (4%), the indicator did not meet either study validity criteria (AUC: 0.5080 ±0.04, IF: 0.07), and was particularly underestimated. Taken together, these results suggest women are not able to accurately report on receiving membrane rupture in general.

INDUCTION & AUGMENTATION OF LABOR: SUMMARY

No indicators on whether a uterotonic was received for labor induction, augmentation, or either indication met both validation criteria. In particular, indicators on uterotonic use met criteria for acceptable individual-level classification, but would be overestimated at an aggregate level. An indicator on receiving membrane rupture in general (for induction or augmentation of labor) did not meet validation criteria. This indicator was particularly underestimated by women.

Uterotonic for Prevention of Post-Partum Hemorrhage (PPH)

Nearly all women received a uterotonic for the prevention of PPH immediately after the birth of the baby. The most common method was by IM injection (**Box 4A**). Women had the option to report more than one method received. To compare against the observer report (one option reported), women who reported receiving IM injection and IV line were coded as IV line (n=1); women who reported receiving IV line and tablets were coded as receiving tablets (n=2). Give that this indicator was near universal, there was not sufficient variation to conduct robust analysis. However, cross-tabulation results (**Box 4B**) show that most women who report receiving either an IM injection, IV medication, or tablets post-delivery were observed to have been given a uterotonic for the prevention of PPH (97% self-reported prevalence; 99% observed prevalence).

Box 4A. Cross-tabulation: Method of prophylactic uterotonic administration.

In the first few minutes after the delivery of your baby, did anyone give you...(1) an injection in the thigh or buttock? (2) medication intravenously through a tube in your arm? (3) tablets to swallow or hold in your mouth or placed in your rectum?

		Observer Report (Number)			Total
		IM Inj.	IV medication	Tablets	
Self-Report (Number)	IM Inj.	504	26	1	531
	IV line	11	1	0	12
	Tablet	2	1	0	3
	Total	517	28	1	546

Box 4B. Cross-tabulation: Received prophylactic uterotonic.

In the first few minutes after the delivery of your baby, did anyone give you...(1) an injection in the thigh or buttock? (2) medication intravenously through a tube in your arm? (3) tablets to swallow or hold in your mouth or placed in your rectum?

		Observer Report (True Prevalence)		Total
		No	Yes	
Self-Report (Number)	No	1	17	18
	Yes	6	538	544
	Total	7	555	562

Route of uterotonic administration

There was not sufficient variation to assess the accuracy of women’s reporting by delivery method. However, cross-tabulation results show that of the women who received prophylactic uterotonic through either an IM injection and/or IV line (n=562), 97% reported having received medication through these routes within the first few minutes of delivery (**Box 4C**).

Timing of uterotonic delivery

To assess the accuracy of women’s reporting with regard to the timing of when the uterotonic was administered, women were asked whether *“Just after the delivery of your baby, in the first few minutes after delivery.”* anyone gave her an injection, IV medication, or tablets. Observers recorded whether the uterotonic was administered 1) at the delivery of the anterior shoulder of the infant during delivery, 2) within 1 minute post-delivery, 3) within 3 minutes post-delivery, 4) more than 3 minutes after delivery and before the delivery of the placenta, or 5) more than 3 minutes after delivery and after delivery of the placenta. Responses of within 1 minute or 3 minutes post-delivery were compared against women’s report of receiving a uterotonic within “the first few minutes after delivery”.

Cross-tabulation results are presented in **Box 4D**. The timing of nearly all uterotonic administration occurred within 3 minutes after birth (82% observed prevalence), and there is

insufficient variation to conduct full validity analysis. Cross-tabulation results suggest that of the women who received a uterotonic within 3 minutes of delivery (n=450), 436 women correctly reported it (97% sensitivity) (**Box 4D**). However, a fairly high proportion of women who reported receiving the intervention (n=99) were *not* observed to have received it within 3 minutes, reflecting a very high false positive rate

Box 4C. Cross-tabulation: Prophylactic uterotonic received by IM injection or IV medication.

In the first few minutes after the delivery of your baby, someone gave you an injection in the thigh or buttock and/or or medication intravenously through a tube in arm.

		Observer Report (Number)		Total
		No	Yes	
Self-Report (Number)	No	0	19	19
	Yes	1	543	544
	Total	1	562	563

Box 4D. Cross-tabulation: Timing of prophylactic uterotonic (1-3 mins post-delivery).

In the first few minutes after the delivery of your baby, someone gave you an injection, medication intravenously or tablets.

		Observer Report (Number)		Total
		No	Yes	
Self-Report (Number)	No	3	14	17
	Yes	99	436	535
	Total	102	450	552

by women (97%). These results should be interpreted with care as full validity analysis was not possible given the insufficient variation in indicator prevalence.

Women were asked whether anyone gave an injection in the thigh or tablets to take or hold in the mouth “*after the delivery of the placenta*”. More than half (59%) of women responded “Yes” to these questions. In contrast, receiving a prophylactic uterotonic after the delivery of the placenta had an observed prevalence of 2%. This indicator had low sensitivity and specificity (54%, 41%, respectively). In particular, a large proportion of women (60%) reported receiving the intervention when they were not observed to have done so. The indicator did not meet either study validity criteria (**Annex B**).

PROPHYLACTIC UTEROTONIC: SUMMARY

Because all women received a prophylactic uterotonic via IM injection within 3 minutes of delivery, there was insufficient variation to rigorously assess the validity of this indicator. The results suggest, however, that women are able to report on receipt of medication or an injection (proxy for prophylactic uterotonic) after delivery, but cannot accurately assess the specific timing of the intervention.

Skilled Birth Attendance

Main provider who assisted in delivery

Women’s self-reports of skilled attendance at birth (defined as primary provider during delivery a doctor, medical resident, or nurse/midwife) was 94%. Self-reported skilled birth attendance during delivery had high sensitivity (95%) and low specificity (15%), reflecting a high false-positive rate. This indicator did not meet both study acceptability criteria ($AUC=0.5510 \pm 0.04$, $IF=1.02$). These results suggest that while the indicator is not recommended for use at the individual level (i.e., classifying individual women who did or did not receive skilled birth attendance), the indicator did meet IF criteria, suggesting the indicator would generate an acceptable estimate of skilled birth attendance coverage at the population level.

The vast majority (81%) of births were attended by nurse/midwives as the primary provider, followed by medical residents (9%), student nurses (5%), doctors (3%), and other providers such as medical interns (<1%) and clinical officers (<1%) (observed prevalence). Two types of providers during delivery had sufficient variation for robust analysis: nurse/midwife and student nurse. Validity results indicate that women were able to accurately report that a nurse/midwife was their primary provider during delivery

via exit interview prior to hospital discharge (AUC=0.7958± 0.03, IF=0.93). However, women were not able to accurately report the presence of a student nurse as their main provider during delivery (AUC=0.5733± 0.04, IF=0.45). Birth attendance by a student nurse had low sensitivity (16%) and high specificity (99%), indicating that a large proportion of women who had a student nurse as their primary birth attendant during delivery did not report it.

Although complete validity analysis for provider types other than nurse/midwife and student nurse was limited by low variability in the types of provider present, descriptive cross-tabulation results suggest that women tend to have difficulty discerning among the narrower categories of providers, such as nurses/midwives versus student nurses (**Table 5**).

TABLE 5. Cross-tabulation: Main provider during delivery.

Self-Report (Number)	Observer Report (Number)							Total
	Doctor (Obgy)	Medical resident	Medical Intern	Nurse/ Midwife	Clinical Officer	Student Nurse	Other	
Doctor (Obgyn)	16	46	6	46	2	7	0	123
Medical resident	0	1	0	0	0	0	0	1
Medical intern	0	0	0	1	0	1	0	2
Nurse/midwife	2	7	1	450	3	17	3	483
Clinical officer	1	0	0	12	0	1	0	14
Student nurse	0	1	0	8	0	5	0	14
Other	0	2	0	5	0	0	0	7
DK	1	3	0	0	0	0	0	4
Total	21	60	7	525	5	31	3	648

Main provider who assisted in labor

Women’s self-reported coverage of skilled attendance in labor was 90%. Similar to reporting of skilled birth attendance during delivery, reporting of the main provider during labor as skilled (i.e., doctor, medical resident, or nurse/midwife) had high sensitivity (91%) and low specificity (17%), indicating that a high proportion of women who were not observed to receive skilled birth attendance during labor, reported that they had a skilled provider. Like skilled attendance during delivery, this indicator met the IF criteria only (AUC=0.5359± 0.04, IF=0.97), suggesting that an acceptable estimate of skilled attendance may be produced at the aggregate level.

Labor was most commonly assisted by a nurse/midwife as the primary provider (94%). To a lesser extent, the main provider was a student nurse (3%), medical intern (2%), personal support companion

(2%), or other type of provider such as clinical officer (<1%), doctor (<1%), or medical resident (<1%). In contrast to delivery, the indicator on whether the main provider during labor was a nurse/midwife did not meet both acceptability criteria (AUC=0.5555± 0.04, IF=0.86). In particular, this indicator had low specificity (30%), indicating that women not attended by a nurse/midwife had low accuracy in reporting they were not attended by a nurse/midwife at an individual level.

While insufficient variation in the type of provider who assisted during labor limits the ability to conduct validity analysis with other types of providers, cross-tabulation results (**Table 6**) show that women both overreported and underreported the skill level of their provider during labor. For example, of the 62 women who reported that their main provider during labor was a doctor, none of these cases were observed to have been a doctor and in fact labor was attended by a nurse/midwife (n=56), medical intern (n=2), clinical officer (n=2) or student nurse (n=2). However, 26 women also reported that no one attended their labor, while no observers reported this. These results suggest women may have low reporting accuracy on the skill level of their main provider during labor at the individual level. In contrast, women were able to validly report some aspects of skilled birth attendance during delivery. Potential explanations for the discrepancy between women’s and observers’ reports could be differences in how women defined the time period of ‘labor’ and in what information was used to assess provider type (e.g., type of dress or gender).

TABLE 6. Cross-tabulation: Main provider during labor.

Self-Report (Number)	Observer Report (Number)								Total
	Doctor (Obgyn)	Medical resident	Medical Intern	Nurse/ Midwife	Clinical Officer	Student Nurse	Support Person	Other	
Doctor (Obgyn)	0	0	2	56	2	2	0	0	62
Medical resident	0	0	0	0	0	0	0	0	0
Medical intern	0	0	0	0	0	1	0	0	1
Nurse/midwife	2	1	9	485	2	11	9	5	522
Clinical officer	0	0	0	12	0	1	1	0	14
Student nurse	0	1	0	12	0	2	0	0	15
Support person	0	0	1	3	0	1	0	0	4
No one	0	0	0	25	0	0	1	0	26
Other	0	0	0	5	0	0	0	0	5
DK	0	0	0	1	0	0	0	0	1
Total	2	1	12	599	4	18	11	3	650

Additional providers during delivery

In addition to identifying their primary provider during labor and delivery, women were asked to list all additional providers who were present both in labor and delivery. For delivery, the next most common additional provider present was a nurse-midwife (48% self-report), followed by a doctor (10%), student nurse (5%), clinical officer (5%), and medical resident or medical intern (<1%, respectively).

Results of the validity analysis reflect the same trend as the primary provider findings, in that the majority of women have difficulty distinguishing among narrower categories of providers (**Annex A**). An indicator on whether a doctor was present as an additional provider during delivery met one of the validity criteria, as did the presence of a nurse/midwife (**Annex B**). No indicators of additional providers during delivery or labor met both validity criteria.

Number of providers

Nearly half of women were attended by 1 provider (46% true prevalence), followed by 2 providers (34% true prevalence). Fewer women were attended by 3 (8%) or 4 or more (11%) providers. The indicator “*Did more than one provider assist with the birth?*” (constructed in analysis) did not meet both study validity criteria (AUC: 0.5640 ± 0.04 , IF: 1.18), nor did indicators on whether 1, 2, or 3 providers assisted with the birth. Women were, however, able to recall with high accuracy whether 4 or more providers assisted with the birth (AUC: 0.7682 ± 0.04 , IF: 1.16).

SKILLED BIRTH ATTENDANT: SUMMARY

Women may be able to accurately report that a nurse/midwife was the main provider during delivery in areas where nurse/midwives are common. Indicators on whether the main provider during labor or delivery was ‘skilled’ (i.e., a doctor, medical resident, or nurse/midwife) met the IF criteria only, suggesting that while valid results were not produced at the individual-level, this indicator may be of use when approximating population-based coverage of skilled birth attendance.

Women were less able to report with accuracy on whether a student nurse was their provider during any stage of labor. Robust analysis was not possible for other types of providers who were less common. Women are able to recall whether 4 or more providers assisted with the delivery with accuracy, but not birth attendance by fewer than this number of providers.

Immediate Newborn Care

Box 5A details immediate postnatal newborn health indicators with sufficient variation for validity analysis. Only one indicator met both study validity criteria: whether the newborn was low birthweight <2,500 g (AUC: 0.8490 ± 0.04, IF: 0.87).

Box 5A. Descriptive frequencies: Newborn health indicators assessed by validity analysis. (>5 reports per cell)			
	Woman, %	Observer, %	
Baby given to mother immediately			
after birth	59.90	57.61	
Immediate skin-to-skin contact	78.90	16.28	
Breastfeeding initiated within first			
hour	76.41	52.99	
Low birthweight baby (<2,500 g)	7.77	8.64	

In the exit interview, women were asked to give the gram weight of their newborn at birth. The low birthweight indicator was constructed in analysis from maternal reports of the gram weight of the newborn. Women’s reports of the gram weight of their newborn was moderately correlated with observer reports ($r=0.6$, $p<0.001$). Feedback from project field staff indicates that nearly all women were given a card with their newborn’s weight listed. It is possible that even if women did not read directly from the card during the interview (analysis only conducted for respondents who self-reported their newborn’s weight, rather than reading directly off the card), having the card may have increased the salience of this indicator and accuracy of reporting. There was an insufficient number of women whose babies were heavy birthweight (>4,500 grams) for analysis.

The indicator of breastfeeding within the first hour after birth did not meet both validity criteria (**Box 5B**). This indicator had high sensitivity (88%) and low specificity (37%), suggesting that while most women who initiate breastfeeding in the first hour after birth

		Box 5B. Cross-tabulation: Breastfed within first hour after birth.		
		Observer Report (Number)		Total
Self-Report (Number)		No	Yes	
	No	96	34	130
	Yes	163	258	421
Total	259	292	551	

correctly reported doing so, nearly two-thirds of women who did not breastfeed in the first hour after birth falsely reported doing so (63%). An IF of 1.44 suggests women would overreport on this indicator nearly 1.5 times the true prevalence in a population-based survey. Confusion over whether ‘the first hour after delivery’ meant the time immediately following delivery or the time after labor procedures ended may have contributed to higher positive reporting by women. For example, in some cases newborns were placed in a warmer and subsequently breastfed. In such instances women may have conceptualized the first hour after birth as the first hour after labor room procedures ended, rather than the first hour after birth as used by observers.

Newborn thermal care

Neither whether the 'newborn was given to the mother immediately after birth' (AUC: 0.5776 ± 0.04 , IF: 1.04) nor whether the 'newborn was placed on the mother's chest, against her skin, immediately after delivery' (AUC: 0.4980 ± 0.04 ; IF: 4.85) met both validation criteria. However, the indicator of whether the 'newborn was given to the mother immediately after birth' did meet IF validity criteria, suggesting it may produce an acceptable estimate of this type of care at the population level.

Results suggest that the way indicators are constructed has important implications for women's reporting accuracy. Specifically, we found that a two-item indicator about whether the newborn was placed skin-to-skin on the mother's chest immediately after delivery reduced women's overestimation of the practice compared to a one-item indicator. The one-item indicator that read "*Did someone place the baby on your chest, against your skin, immediately after delivery of the baby*" had a reported prevalence by women of 79%, and observer prevalence of 16%. This indicator was characterized by a high false positive rate: 79% of women whose newborn was *not* placed skin-to-skin reported receiving the intervention (sensitivity: 79%, specificity: 21%). A two-item indicator that included only women who positively answered "*Did someone place the baby on your chest, against your skin, immediately after delivery*" and subsequently reported that the baby was lying naked against the skin to the question "*Was your baby wrapped in a towel while lying against your chest or lying naked against your skin*" had fewer false positive reports. Self-reported prevalence on the two-item skin-to-skin indicator was 29%, compared to the observer report of 16% (sensitivity: 27%, specificity, 70%, AUC: 0.4857 ± 0.04 , IF: 1.80).

Nearly all babies breathing at birth were covered with a towel. This occurred regardless of whether the newborn was placed directly on the mother's skin and then draped with a towel, or wrapped in a towel first and then placed with the mother. It was more common to wrap the baby in a towel before placing the newborn with the mother (84% of all births). Because the practice of wrapping the newborn after birth was nearly universal, it was not possible to assess the accuracy of women's reporting on this indicator. Cross-tabulation showed that 91% of women whose baby was wrapped in a towel and then placed with the mother reported that this practice took place (compared to 92% observed prevalence).

Composite indicator of essential newborn care

A composite indicator of three essential elements of newborn care (whether the newborn was immediately dried, placed skin-to-skin on mother's chest, and breastfed in first hour) was constructed in analysis. This indicator had fairly high sensitivity (70%) and low specificity (28%) and did not meet either study validity criteria. In particular, the indicator had a high false positive rate, where 72% of women

who did not receive all three elements of newborn care falsely reported receiving the interventions. Using the two-item 'skin-to-skin' item in this composite indicator greatly reduced the reporting of false positives to 11%, but also lowered sensitivity (7%), indicating fewer true cases were accurately captured. Neither indicator met both study validity criteria.

IMMEDIATE NEWBORN CARE: SUMMARY

Among indicators of immediate newborn practices that had sufficient variability to be assessed, only the indicator for low birthweight met both study validity criteria. Question wording and order may have significant implications for the validity of indicators of newborn thermal care. The indicator 'whether the newborn was placed with the mother immediately following birth' met only the IF validity criteria.

Results for the reporting of newborn birthweight should be interpreted with caution. Although the analysis was restricted to women who self-reported the newborn weight, women were given a card with the newborn's birthweight, which may have enhanced the salience of this indicator. Women tended to overreport other types of practices such as breastfeeding in the first hour, whether their baby was placed skin-to-skin on the mother's chest, and composite indicators of newborn care.

Maternal Outcomes

Complications

Women were asked whether they experienced any of the following symptoms either during or immediately following delivery: (1) high blood pressure, seizures, blurred vision, severe headaches, (2) swelling in hands or face, (3) baby was in distress/ too large, (4) long labor (more than 12 hours), (5) excessive bleeding, (6) infection (fever), (7) another type of complication (asked to specify), or (8) no complications. Nearly half (45%) of women reported experiencing some type of complication, exceeding the observed prevalence (11%) nearly five times. Women's report of experiencing any type of complication did not meet study validity criteria (**Annex B**). Self-reports of experiencing any complication had a sensitivity of 63%, indicating that over one-third of women who had experienced a complication did not report it. The indicator also had lower specificity (57%), reflecting a high rate of false positive reports by women. Specifically, nearly half of women (43%) who had not experienced a complication falsely reported experiencing one.

Women most commonly reported experiencing prolonged labor (lasting >12 hours) (24% self-reported prevalence), followed by excessive bleeding (11%) (indicative of hemorrhage) and either high blood pressure, blurred vision, severe headache, or swelling of hands or face (11%) (indicative of pre-

eclampsia/ eclampsia) (**Box 6**). There was sufficient variation to assess the accuracy of women’s report of two complications: hemorrhage (APH + PPH) and prolonged labor. Neither indicator met both study validity criteria (**Annex B**).

Box 6. Descriptive frequencies: Complications.					
Did you experience any of the following complications during or after your delivery? (List all)			Record whether mother had any of the following complications. (Select all)		
	Number	Percent		Number	Percent
High blood pressure, seizure, blurred vision, severe headache, swelling in hands/face	72	10.81	Pre-eclampsia/ Eclampsia	2	0.30
Excessive bleeding	76	11.41	Hemorrhage	30	4.54
Baby was in distress/ too large	64	9.61	Obstructed labor	11	1.66
Long labor (more than 12 hrs)	157	23.57	Prolonged labor	24	3.63
Infection (fever)	22	3.30	Puerperal infection	0	0
Other	22	3.30	Other	16	2.42
No complications	343	51.50	No complications	589	89.11
Don't Know	30	4.50			
Total	666	100	Total	661	100

The indicator for hemorrhage had a sensitivity of 33%, indicating that two-thirds of true hemorrhage cases were not reported by women. Over 10% of women who reported experiencing excessive bleeding were not observed hemorrhage cases (10%). The indicator for prolonged labor had a sensitivity of 50%, indicating that one-half of prolonged labor cases were not reported by women. Simultaneously, nearly one-quarter of women falsely reported experiencing prolonged labor (23%). Taken together, there is misclassification of the complications of hemorrhage and prolonged labor due to false reporting, while true complications are also underreported.

While not able to be robustly assessed due to lack of variation in the data, a greater proportion of women reported symptoms of preeclampsia/eclampsia (11%) compared to observer reports of the complication (<1%). It is possible that women experienced these symptoms and providers and/or observers were not aware of them. If women are conscious of signs of preeclampsia/eclampsia and make providers aware of their symptoms, this may draw attention to danger signs indicative of the complication, a leading cause of maternal death among women.

Cesarean section

14% of women reported receiving a cesarean section operation, which closely approximated the true prevalence of 13%. Validity analysis shows that women are able to accurately report whether they

received a cesarean operation (sensitivity: 93%, specificity: 99%). This indicator met both study validity criteria (**Annex B**).

Given lack of variation in the timing of the decision for cesarean operation and in who made the decision, we were not able to fully assess these indicators. However, of women who received a cesarean operation after labor had started, 91% reported receiving the operation at this time. Similarly, of women whose decision for cesarean operation was made after labor started, 91% of women reported the decision was made at this time. According to the observers, all decisions for cesarean operations were made by a provider. Over four-fifths of women (83%) whose cesarean operation was decided on by a provider indicated that this was the case.

Indication for cesarean section

Given the relatively small number of cesarean operations (n=89), there was insufficient variation to assess the accuracy of women’s report of their indication for cesarean operation. Observers’ and women’s self-reported frequencies are listed in **Box 7**.

Box 7. Descriptive frequencies: Cesarean operation indication.					
What was the reason for your cesarean operation? (List all)			What was the indication for the cesarean operation? (Select all)		
	Number	Percent		Number	Percent
I was bleeding	0	0	Excessive bleeding	0	0
The baby was stuck	17	19.10	Obstructed labor	16	19.05
I was in labor for a long time	9	10.11	Prolonged labor	30	35.71
Baby not in the right position	11	12.36	Malpresentation	8	9.52
Problems with the baby	13	14.61	Fetal distress	18	21.43
Doctor/nurse told me I had to	37	41.57	Previous scar	12	14.29
Don't Know	2	2.25	No medical reason	0	0
			Don't Know	0	0
			Other	13	15.48
Total	89	100	Total	84	100

We were able to collapse two indications for cesarean operation for obstructed labor and prolonged labor (>12hrs) to assess the accuracy of women’s report of receiving a cesarean operation because “the baby was stuck” or the woman was “in labor for a long time”. This indicator did not meet validity criteria (**Annex B**). The indicator had particularly low sensitivity (39%), meaning that approximately two-thirds of women who received cesarean operation because of obstructed or prolonged labor did not report this reason.

Post-delivery health checks for mother

Nearly all (96%) women reported that the provider performed at least one of the following health checks in the first physical examination after delivery: (1) check for bleeding, (2) perineal exam, (3) check for involution (to see whether womb was becoming firm), (4) took blood pressure, or (5) took temperature (**Box 8**). Observers indicated that 95% of women experienced at least one of these checks.

While there was insufficient variation in results to conduct validity analysis of this indicator, we were able to assess the accuracy of women’s report of receiving specific post-delivery checks including: 1) bleeding, 2) examine the perineum, 3) temperature, 4) blood pressure, and 5) involution.

No indicators on specific health checks post-delivery met both study validity criteria (**Annex B**). Specifically, a high proportion of women falsely reported that a provider either took their temperature (50%) or blood pressure post-delivery (62%), although these indicators

had fairly high sensitivity. Provider checks to examine the perineum, check for involution, or check for bleeding were characterized by moderate to low sensitivity and specificity, reflecting both underreporting of true cases and false reporting by women. The discrepancy in observer and women’s reports of postnatal maternal health checks may be due to differences in timing. For example, study observers continued observation up to one hour after the birth of the baby. In the exit interview, women were asked whether the health checks occurred in the first physical exam following the delivery of the placenta, as this time period best correlated with the time period completed by study observers. The higher false positive rate among women may also indicate that some health checks (i.e. blood pressure and temperature) were also done upon admission to the post-labor ward, in which case they would not have been captured by observers.

Box 8. Descriptive frequencies: Post-delivery health checks.		
In the first physical examination after delivery did the provider look, ask for, or examine for: bleeding, perineal exam, involution, take blood pressure, or take temperature?		
	Woman, %	Observer, %
Yes to any of following checks in first physical examination post-delivery	95.99	94.92
Check for bleeding	62.04	90.59
Did a perineal exam	56.14	87.36
Checked whether womb was becoming firm (involution)	64.23	78.21
Took blood pressure	74.61	48.29
Took temperature	60.03	40.28

MATERNAL OUTCOMES: SUMMARY

Women were able to report with high accuracy that they received a cesarean section. Women were less able to report on the reason for cesarean operation, whether complications were experienced, and whether post-partum provider checks for specific aspects of maternal health took place following delivery.

Newborn Outcomes

Stillbirth delivery

Given the small number of stillbirth deliveries (n=9, observed prevalence), we were unable to analyze the validity of an indicator of stillbirths. While results should be interpreted with caution, cross-tabulation results (**Box 9**) indicate that more than half of stillbirth deliveries were incorrectly

classified by women (5 of 9). Additionally, two mothers reported their newborn was a stillbirth while the observer reported the newborn to be alive. These newborns were born alive and died within an hour after delivery. The short newborn survival times suggest that mothers may have thought their newborn was not born alive.

Box 9. Cross-tabulation: Stillbirth delivery.

		Observer Report (Number)		
		Alive	Stillbirth	Total
Self-Report (Number)	Alive	640	5	645
	Stillbirth	2	4	6
	Total	642	9	651

Discussion

51 out of 98 possible indicators had sufficient numbers for validity analysis. Of these, 5 indicators met both acceptability criteria ($AUC > 0.6$ and $0.75 < IF < 1.25$) (Table 7). One additional indicator (receiving episiotomy) was very close to meeting both criteria ($AUC = 0.8659$, $IF = 1.26$). A total of 26 indicators met only one of the acceptability criteria: 13 met acceptability criteria for individual-level classification ($AUC > 0.60$), and 13 met acceptability criteria for valid approximation of population-based coverage.

Skilled Birth Attendance Indicators

A key objective of the study was to assess the validity of women's reports of 'skilled attendance at birth', i.e., what category of provider assisted with their deliveries. Study results suggest that women can accurately report on the most common type of provider in this setting, a nurse/midwife.

Specifically, an indicator on whether a nurse/midwife was the main provider who assisted during delivery met both validity criteria. While the

TABLE 7. Indicators that met both validation criteria.

Main provider during delivery was a nurse/midwife
More than 4 providers assisted with the birth
A support companion was present during the birth
Newborn with low birthweight (<2,500 grams)
Cesarean section performed

low variability in types of providers observed may have limited our analysis of other types of providers, study findings suggest women are less likely to correctly categorize types of providers who are less common, such as student nurses. Cross-tabulation results suggest that women have greater difficulty in distinguishing among less prevalent types of providers (e.g., medical residents, medical interns, student nurses), although the low prevalence of these types of providers in the facility setting limits these findings. Some skilled birth attendant indicators such as whether the main provider during labor and or delivery was 'skilled' (defined as a doctor/obgyn, nurse/midwife, or medical resident), met only the IF validity criteria. While the AUC calculated for these indicators did not meet criteria for accurate classification at the individual level, discrepancies in sensitivity and specificity balance out to produce an acceptable estimate for monitoring coverage purposes at the population level.

Taken together, results suggest women can accurately report that a nurse/midwife was their main provider during delivery in facility-based settings where nurse/midwives are common (>94%). Other, less common types of providers during labor and delivery in facility-based settings may be less accurately reported. We hypothesize that differences in how women understood key terminology, such the time period of 'labor' and how to identify the 'main provider,' may have contributed to discrepancies

in reporting at the individual level. Skilled birth attendant indicators that met only the IF criteria may be acceptable for use in surveys intended to estimate acceptable coverage of maternal and newborn care at the population-level.

Validated Indicators that Met Both Study Validity Criteria

Study findings also suggest that indicators of observable aspects of care are more likely to be validly reported than indirect indicators such as facility policies or justification for an intervention received. For example, women were able to recall whether a surgery such as a cesarean operation or episiotomy was performed (the latter approached validity criteria), but not the indication for receiving a cesarean operation. Similarly, women were able to report on observable aspects of provider respectful care, such as whether a support person was present during the birth, but not policies such as whether the woman was allowed to eat/ drink or was encouraged to assume different positions during labor. More indirect aspects of care may have been less explicitly stated and therefore less clearly understood by women. Validation of observable aspects of care, such as the presence of a support companion during birth and cesarean section delivery, confirms results of a separate study conducted in Malawi.⁹

The indicator “newborn was low birthweight (<2,500 grams)” also met both study validity criteria. However, these results should be interpreted with caution since, although analysis was restricted to women who reported their newborns’ weight from recall (not read from a card), all women were given a card listing their newborn’s birthweight. This practice may have increased the salience of the birthweight indicator. Of note is that the low birthweight indicator was constructed in analysis. Women were asked to give their newborn’s birthweight in grams, which corresponds to the measurement of newborn birthweight in DHS and MICS. Women’s report of broader categories of newborn birthweight (i.e., low birthweight or not) were reported with accuracy, while reports of the newborn’s weight in grams were only moderately correlated with observer records. These findings may have implications for indicators used in DHS and MICS, such as “*When (NAME) was born was s/he very large, larger than average, average, smaller than average, or very small*” and could be explored in future studies, including population-based studies used for DHS and MICS data collection.

Validated Indicators that Met One Validity Criteria

Thirteen indicators met AUC acceptability criteria of greater than 0.60 (**Annex B**). The most accurately reported responses were whether labor was induced with a uterotonic, a combined indicator on whether the main provider during delivery was a doctor or medical resident, and whether episiotomy was performed (AUC of 0.80 or greater). The other validity criteria was an IF score between 0.75 and 1.25.

An additional 13 indicators met this criteria. These were: whether the woman was referred to the facility because of a problem, whether the woman's blood pressure was taken during her initial assessment, whether she was encouraged or assisted to ambulate during labor, whether the woman was allowed to have a support person during labor and delivery, whether she was attended by a skilled provider (doctor, medical resident, or nurse/midwife) during labor and delivery, whether the main provider during labor was a nurse/midwife, whether a nurse/midwife was present as an additional provider during delivery, whether multiple providers assisted with the birth, whether two providers assisted with the birth, whether the baby was placed with the mother immediately following birth, whether the infant received three essential elements of newborn care (using composite 2-item skin-to-skin indicator) and whether, in the first examination following delivery, the provider checked for involution. The fact that these indicators met only the IF criteria suggests that although low levels of indicator sensitivity and/or specificity prevented the AUC from reaching an acceptable level, under- and overreporting at the individual level canceled out to generate acceptable estimates for monitoring coverage of these indicators at the population level. However, we recommend caution with regard to estimating low-prevalence indicators since, without near perfect negative classification (i.e., specificity), these indicators are likely to be overestimated.

Validated Indicators that Did Not Meet Either Study Validity Criteria

Also of importance are core indicators assessed by robust analysis which did not meet either validity criteria. We define 'core' indicators as those with high potential for practical use, given their inclusion in DHS and MICS surveys, or those under consideration by WHO-PMNCH for measuring quality of care in health facilities, or those that reflect essential aspects of obstetric and immediate newborn care. Core assessed indicators which did not meet either validity criteria were: whether the newborn was placed skin-to-skin on the mother immediately following delivery, and whether the newborn received three key elements of essential newborn care. The three elements of essential newborn care comprised a proxy composite indicator of essential newborn care that measured whether the newborn was: immediately dried following delivery, immediately placed skin-to-skin with mother, and breastfed in the first hour after birth. Although delayed cord clamping has also been included as an element of essential newborn care, this indicator was not assessed in the present study. None of these core validated indicators met both study validity criteria. The composite indicator of newborn care was notably overestimated by women.

Findings also highlight the importance of indicator wording on women's reporting accuracy. For example, a two-item indicator about whether the newborn was placed skin-to-skin on the mother's chest immediately after delivery greatly reduced women's overestimation of the practice compared to a one-

item indicator. These results suggest that women's conceptualization of key terms used in survey items (e.g., understanding of directly 'on skin' as opposed to wrapped in a towel and placed on the skin) and standard of care practices should be further explored to refine indicators. The skin-to-skin indicator contributed to overreporting of the composite indicator of essential newborn care among women. While adjusting the composite indicator by using a two-item skin-to-skin indicator reduced overreporting, neither indicator (one or two-item) met both validation criteria. Difficulty in women's understanding of indicators related to newborn thermal care and the timing of events after delivery has also been reported in previous studies of newborn and postnatal care in Bangladesh and Malawi.¹¹

Other indicators are not recommended, given notably high 'Don't Know' responses reported by women, including measures of provider hygiene (hand-washing/ hand sanitization practices), aspects of immediate postnatal care (whether baby was fed anything other than breastmilk or bathed in the first hour after birth), and whether the provider conducted a perineal health-check post-delivery (**Table 4**).

Indicators to be Further Researched

Because institutional delivery is a widely used proxy for whether women have access to comprehensive services such as emergency obstetric care and lifesaving commodities (e.g., uterotonic, magnesium sulfate, antibiotics), and is included in DHS and MICS, we want to know whether women can accurately report on the type of facility where they delivered. The present study was not designed to assess whether women can accurately report on this indicator. However, cross-tabulation results show that nearly all women (98%) who responded to this indicator correctly classified the type of facility as a public/ government hospital. These results may suggest that women in Kisumu and Kiambu districts of Kenya who deliver in a public, government facility are able to identify the location as such. This finding should be explored further in a study that includes multiple facility types (private and public).

Also of interest was whether a uterotonic for the prevention of postpartum hemorrhage was received immediately (up to 3 minutes) after birth, since this intervention is critical in preventing postnatal hemorrhage, a leading cause of maternal death. Although robust analysis for this indicator was limited because of insufficient variation in the data, preliminary results suggest that women may be able to report on whether a prophylactic uterotonic was received, but not the timing of the intervention. These results should be further explored in another study, particularly since this study did not assess whether the woman knew the purpose of the injection or medication she received.

Literature Cited

1. Adegoke AA, van de Broek N. Skilled birth attendance- lessons learnt. *BJOG*. 2009;116(Suppl 1):33-40.
2. Hodgins S. Achieving better maternal and newborn outcomes: coherent strategy and pragmatic, tailored implementation. *Global Health: Science and Practice*. 2013;1(2):146-153.
3. Filippi V, Ronsmans C, Gandaho T, Graham W, Alihonou E, P. S. Women's reports of severe near-miss)obstetric complications in Benin. *Studies in Family Planning*. 2000;31(4):309-324.
4. Seoane G, Castrillo M, O'Rourke K. A validation study of maternal self reports of obstetrical complications: implications for health surveys. *International Journal of Gynecology & Obstetrics*. 1998;62(3):229-236.
5. Sloan NL, Amoafu E, Arthur P, Winikoff B, Adjei S. Validity of women's self-reported obstetric complications in rural Ghana. *Journal of Health, Population and Nutrition*. 2001;19(2):45-51.
6. Stewart MK, Festin M. Validation study of women's reporting and recall of major obstetric complications treated at the Philippine General Hospital. *International Journal of Gynecology & Obstetrics*. 1995;48(Supplement)(53-66).
7. Stewart MK, Stanton CK, Festin M, Jacobson N. Issues in measuring maternal morbidity: lessons from the Philippines Safe Motherhood Survey Project. *Studies in Family Planning*. 1996;27(1):29-35.
8. Tuncalp O, Stanton C, Castro A, et al. Measuring coverage in MNCH: Validating women's self-report of emergency cesarean sections in Ghana and the Dominican Republic. *PLOS One*. 2013;8(5):e60761.
9. Stanton SK, Rawlins B, Drake M, et al. Measuring coverage in MNCH: Testing the validity of women's self-report of key maternal and newborn health interventions during the peripartum period in Mozambique. *PLOS One*. 2013;8(5):e60694.
10. Liu L, Li M, Yang L, et al. Measuring coverage in MNCH: A validation study linking population survey derived coverage to maternal, newborn and child health care records in rural China. *PLOS One*. 2013;8(5):e60762.
11. Yoder P, Rosato M, Mahmud R, Fort A, Rahman F, al. e. *Women's recall of delivery and neonatal care: a study of terms, concepts and survey questions*. Calverton, Maryland: ICF Macro;2010.
12. Hill Z, Tawiah-Agyemang C, Wickenden M, Okyere E, Ten-Asbroek G, Kirkwood B. *Postnatal care in rural Ghana: Mothers' experiences and recall*. Institute for Child Health: University of London;2010.
13. Moran AC, Kerber K, Sitrin D, et al. Measuring coverage in MNCH: Indicators for global tracking of newborn care. *Plos Medicine*. 2013;10(5):e1001415.

14. Kenya National Bureau of Statistics (KNBS) and ICF Macro. *Kenya Demographic and Health Survey 2008-09*. Calverton, Maryland: KNBS and ICF Macro;2010.
15. Eng J. Receiver operating characteristic analysis: a primer. *Academic Radiology*. 2005;12:909-916.
16. Metz CE. Practical issues in ROC studies. *Radiology*. 1989;143:29-36.
17. Pepe MS. *Statistical evaluation of medical tests for classification and prediction*. Oxford: Oxford University Press; 2003.
18. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;201:621-625.
19. Vecchio TJ. Predictive value of a single diagnostic test in unselected populations. *New England Journal of Medicine*. 1966;274(21):1171-1173.
20. Campbell H, Biloglav Z, Rudan I. Reducing bias from test misclassification in burden of disease studies: use of test to actual positive ratio- new test parameter. *Croatian Medical Journal*. 2008;49:402-414.

Annexes

Annex A.

TABLE 1. Full List of Indicators Assessed and Measured Coverage,^{a,b} Kenya.

Indicator	Self-Report Prevalence (Matched data)	True Prevalence (Matched data)	At least 5 counts/cell?
Initial Client Assessment			
Type of facility where gave birth- public hospital	98.00	100.00	N
Referred to facility because of a problem	8.24	7.94	Y
HIV status checked	25.19	94.67	N
Offered HIV test	8.33	1.67	N
Receives HIV test	8.56	9.33	N
Provider washes hands with soap and water or uses antiseptic before any initial examination	73.09	26.55	Y
Takes blood pressure	93.43	87.00	Y
Takes urine sample	5.66	1.38	N
Checks fetal heart rate with fetoscope/ ultrasound	95.75	99.7	N
Wears high-level disinfected or sterile gloves for vaginal examination	99.85	99.85	N
Provider Respectful Care			
Encourages/assists woman to ambulate during labor	86.96	77.48	Y
Woman allowed to drink liquids/eat	66.83	41.99	Y
Woman allowed to have a support person present during labor and delivery	8.78	9.10	Y
Encourages/assists woman to assume different positions in labor	14.33	58.24	Y
A support person is present at birth	3.73	4.81	Y
First Stage of Labor			
Induces labor by uterotonic (IV, IM, tablet) ^d	10.79	4.60	Y
(Of women whose labor was induced) Uterotonic route for induction of labor - Tablet (oral or vaginal) ^d	80.95	100.00	N
Augments labor with uterotonic (by IV line, IM injection, or tablet) ^d	39.20	22.40	Y
(Of women whose labor was augmented) Augmentation of labor by IV line (Push, Drip, Drip plus IM) ^d	92.45	100.00	N
Uterotonic received (to induce or augment labor) ^d	43.78	27.14	Y
Membranes ruptured (to induce or augment labor) ^d	3.08	42.31	Y
Skilled Birth Attendance- Main Provider			
Skilled main provider labor ^{c, d}	89.98	92.60	Y
Main provider labor- doctor or medical resident ^d	9.55	0.46	N
Main provider labor- doctor (ob-gyn)	9.72	0.31	N
Main provider labor- medical resident	0.00	0.16	N
Main provider labor- medical intern	0.16	1.88	N

Indicator	Self-Report Prevalence (Matched data)	True Prevalence (Matched data)	At least 5 counts/cell?
Main provider labor- nurse/midwife	80.41	93.73	Y
Main provider labor- clinical officer	2.04	0.63	N
Main provider labor- facility support/ staff aide	0.15	0.31	N
Main provider labor- student nurse	2.31	2.77	N
Main provider labor- support companion	0.62	1.71	N
Skilled main provider delivery ^{c,d}	94.25	92.86	Y
Main provider delivery- doctor (ob-gyn) or medical resident	19.25	11.80	Y
Main provider delivery- doctor (ob-gyn)	19.10	3.00	N
Main provider delivery- medical resident	0.15	8.85	N
Main provider delivery- medical intern	0.31	1.09	N
Main provider delivery- nurse/midwife	75.00	81.06	Y
Main provider delivery- clinical officer	2.17	0.78	N
Main provider delivery- student nurse	2.17	4.81	Y
Skilled Birth Attendant- Other Providers Present			
Other provider(s) labor			
Other provider labor- doctor (ob-gyn)	9.02	1.99	N
Other provider labor- medical resident	0.15	3.98	N
Other provider labor- medical intern	0.92	28.29	N
Other provider labor- nurse/midwife	50.46	69.88	Y
Other provider labor- student nurse	5.35	12.23	Y
Other provider labor- clinical officer	3.82	9.79	N
Other provider(s) delivery			
Other provider delivery- doctor (ob-gyn)	9.52	3.07	Y
Other provider delivery- medical resident	0.15	4.92	N
Other provider delivery- medical intern	0.61	14.59	N
Other provider delivery- nurse/midwife	48.15	64.31	Y
Other provider delivery- student nurse	5.26	8.96	Y
Other provider delivery- clinical officer	4.79	5.41	N
More than one provider assisted with birth ^d	64.55	54.69	Y
One provider assisted with birth	35.02	46.06	Y
Two providers assisted with birth	34.86	34.07	Y
Three providers assisted with birth	16.88	8.04	Y
Four or more providers assisted with birth	12.93	11.20	Y
Second & Third Stage of Labor			
Episiotomy performed	22.94	18.17	Y
Uterotonic administered within few minutes of delivery (via injection, IV medication, or oral/rectal tablets)	96.80	98.75	N
Uterotonic received 1-3 mins after birth	96.92	81.52	N
Uterotonic received after delivery of placenta	59.06	2.36	Y
Applies controlled cord traction	97.50	98.93	N
Performs uterine massage after delivery of placenta	88.35	98.57	N
Position of mother at birth- on back	94.73	99.84	N

Indicator	Self-Report Prevalence (Matched data)	True Prevalence (Matched data)	At least 5 counts/cell?
Health provider wore gloves during delivery of baby	100.00	99.82	N
Immediate Newborn Care (babies breathing at birth)			
Baby immediately dried with towel/cloth	96.13	99.49	N
Baby given to mother immediately after birth	59.90	57.61	Y
Baby placed immediately skin to skin on mother's abdomen	78.90	16.28	Y
Baby immediately skin to skin on mother (2 item indicator) ^e	29.19	16.20	Y
Babies on skin covered with dry towel on mothers abdomen	42.31	100.00	N
Babies not on skin wrapped with towel	90.70	91.86	N
Breastfeeding within first hour of birth	76.41	52.99	Y
Something other than breastmilk given to baby within first hour of delivery	1.92	1.05	N
Baby bathed within the first hour after birth ^d	2.81	0.05	N
Baby weighed	99.84	100.00	N
Low birth-weight baby (<2,500 g) ^d	6.74	7.77	Y
High birth-weight baby (>=4,500 g) ^d	1.03	1.03	N
3 elements of newborn care (immed. dried + on skin + breastfed in first hour) ^d	71.54	9.29	Y
3 elements of newborn care (immed. dried, 2 item skin-to-skin ^e , breastfed in first hour) ^d	29.4	9.18	Y
Immediate Postnatal Care			
Palpates uterus 15 minutes after delivery of placenta	88.33	70.20	Y
Provider did at least one post-delivery health check ^d	95.99	94.92	N
In first post-delivery exam, provider checks for bleeding	62.04	90.59	Y
In first post-delivery exam, provider examines perineum	56.14	87.36	Y
In first post-delivery exam, provider takes temperature	60.03	40.28	Y
In first post-delivery exam, provider takes blood pressure	74.61	48.29	Y
In first post-delivery exam, provider checks for involution	64.23	78.21	Y
Maternal and Infant Outcomes			
Cesarean section (C/S) performed	13.52	13.36	Y
Decision for C/S taken after labor started	90.79	100.00	N
C/S performed after labor started	90.79	100.00	N
Provider decided C/S would be done	82.50	100.00	N
Reason for C/S- prolonged/obstructed labor	32.89	67.11	Y
Complications- Any ^d	44.80	11.00	Y
Eclampsia	10.86	0.31	N
Hemorrhage	11.16	4.59	Y
Prolonged labor (>12 hours)	23.70	3.67	Y
None	51.53	88.99	Y
Blood products given	15.28	18.06	N
Woman asked for pain relief medication while at facility	32.13	10.50	Y
Woman received pain relief medication	59.38	17.50	Y
Stillborn delivery ^d	0.92	1.38	N

^aText in blue notes indicators where there was not sufficient cell counts for robust analysis (n<5 per cell).

^bExcludes 'Don't Know' responses.

^c Skilled provider is doctor (ob-gyn), nurse/midwife or medical resident

^d Indicator constructed in analysis to dichotomize women's responses to related question.

^e Indicator constructed from two skin-to-skin items: (1) baby placed against mother's chest after delivery and (2) baby was naked against the mother's chest.

Annex B.

TABLE 1. Validation Results for All Indicators With at least 5 Counts per Cell, Matched Data*, Kenya.

Indicator	N Matched data	Reported Prev (%) Matched data	True Prev (%) Matched data	Sensitivity of Self Report	Specificity of Self-Report	Population Survey Estimate	AUC (>0.60)	IF (0.75 to 1.25)	Recommend? (Y/N) <i>List Criteria</i>
Initial Client Assessment									
Woman referred to facility because of a problem	655	8.24	7.94	25.00	93.20	8.24	0.5910	1.04	IF
Provider washes hands with soap and water or uses antiseptic before initial examination	467	73.09	26.55	83.87	32.94	71.52	0.5841	2.69	
Takes blood pressure	654	93.43	87.00	87.73	23.26	86.30	0.5549	0.99	IF
Provider Respectful Care									
Woman allowed to drink liquids or eat	624	66.83	41.99	72.52	37.29	66.83	0.5491	1.59	
Encourages/assists woman to ambulate during labor	644	86.96	77.48	90.18	24.14	86.96	0.5716	1.12	IF
Encourages/assists woman to assume different positions in labor	649	14.33	58.24	19.05	92.25	14.33	0.5565	0.25	
Woman allowed to have a support person during labor and delivery	648	8.80	9.10	23.73	92.70	8.80	0.5821	0.97	IF
Support companion present during birth	644	3.73	4.81	48.39	98.53	3.73	0.7346	0.77	Yes
First Stage of Labor									
Induces labor with uterotonic	630	10.79	4.60	68.97	92.01	10.80	0.8049	2.35	AUC
Augments labor with uterotonic	625	39.20	22.40	72.86	70.52	39.20	0.7169	1.75	AUC
Uterotonic received (labor induction or augmentation)	619	43.78	27.14	77.98	68.96	68.96	0.7347	1.61	AUC
Membranes ruptured (labor induction or augmentation)	650	3.08	42.31	4.00	97.60	3.08	0.5080	0.07	
Skilled Birth Attendance									
Skilled main provider ^a labor	649	89.98	92.60	90.52	16.67	89.99	0.5359	0.97	IF
Main provider labor nurse/midwife	638	80.41	93.73	81.10	30.00	80.28	0.5555	0.86	IF
Other provider labor nurse/ midwife	654	50.46	69.88	48.36	59.90	45.87	0.5413	0.66	
Other provider labor student nurse	654	5.35	12.23	13.75	95.82	5.35	0.5478	0.44	
Skilled main provider ^a delivery	644	94.25	92.86	94.98	15.22	94.25	0.5510	1.02	IF
Main provider delivery doctor (ob-gyn)/ medical resident	644	19.25	11.80	82.89	89.26	19.25	0.8608	1.63	AUC
Main provider delivery nurse/ midwife	644	75.00	81.06	86.21	72.95	75.00	0.7958	0.93	Yes

Indicator	N Matched data	Reported Prev (%) Matched data	True Prev (%) Matched data	Sensitivity of Self Report	Specificity of Self-Report	Population Survey Estimate	AUC (>0.60)	IF (0.75 to 1.25)	Recommend? (Y/N) <i>List Criteria</i>
Main provider delivery student nurse	644	2.17	4.81	16.13	98.53	2.18	0.5733	0.45	
Other provider delivery doctor (ob-gyn)	651	9.52	3.07	35.00	91.28	9.53	0.6314	3.10	AUC
Other provider delivery nurse/ midwife	650	48.15	64.31	49.52	54.31	48.15	0.5192	0.75	IF
Other provider delivery student nurse	647	5.26	8.96	13.79	95.59	5.25	0.5469	0.59	
Second and Third Stage Labor									
Episiotomy performed	545	22.94	18.17	82.83	90.36	22.94	0.8659	1.26	AUC
Uterotonic received following delivery of placenta	552	59.06	2.36	53.85	40.82	59.05	0.4733	25.1	
Multiple providers assisted with birth	629	64.55	54.69	70.35	42.46	64.55	0.5640	1.18	IF
One provider assisted with birth	634	35.02	46.06	42.81	71.64	35.02	0.5722	0.76	
Two providers assisted with birth	634	34.86	34.07	37.50	66.51	34.86	0.5200	1.02	IF
Three providers assisted with birth	634	16.88	8.04	29.41	84.22	16.88	0.5682	2.10	
4+ providers assisted with birth	634	12.93	11.20	60.56	93.07	12.94	0.7682	1.16	Yes
Immediate Newborn Care									
Baby given to mother immed. after birth	611	59.90	57.61	66.48	49.03	59.91	0.5776	1.04	IF
Baby placed immed. skin to skin on mother	602	78.90	16.28	78.57	21.03	78.90	0.4980	4.85	
Baby placed immed. skin to skin on mother (2 item) ^b	596	29.19	16.20	26.80	70.34	29.20	0.4857	1.80	
Breastfeeding within first hr of birth	551	76.41	52.99	88.36	37.07	76.41	0.6271	1.44	AUC
3 elements of essential newborn care (immed. dried, on mother's skin, breastfed within first hr)	506	71.5	9.29	70.21	28.32	71.54	0.4927	7.70	
3 elements of essential newborn care (immed. dried, 2 item on mother's skin ^b , breastfed within first hr)	501	29.44	9.18	7.14	90.13	9.62	0.4864	1.05	IF
Low birthweight newborn (<2,500g)	579	6.74	7.77	71.11	98.69	6.73	0.8490	0.87	Yes
Immediate Postnatal Care									
Palpates uterus 15 minutes after delivery of placenta	557	88.33	70.20	88.75	12.65	88.33	0.5070	1.26	
First post-delivery exam, provider ask/checks for bleeding	627	62.04	90.59	59.86	16.95	62.04	0.3840	0.68	
First post-delivery exam, provider examines perineum	554	56.14	87.36	57.85	55.71	56.14	0.5678	0.64	
First post-delivery exam, provider takes temperature	638	60.03	40.28	75.10	50.13	60.03	0.6261	1.49	AUC
First post-delivery exam, provider takes blood pressure	642	74.61	48.29	88.06	37.95	74.61	0.6301	1.55	AUC
First post-delivery exam, provider checks for involution	615	64.23	78.21	64.03	35.07	64.23	0.4955	0.82	IF

Indicator	N Matched data	Reported Prev (%) Matched data	True Prev (%) Matched data	Sensitivity of Self Report	Specificity of Self-Report	Population Survey Estimate	AUC (>0.60)	IF (0.75 to 1.25)	Recommend? (Y/N) <i>List Criteria</i>
Woman asked for pain relief medication during stay	638	32.13	10.50	35.82	68.30	32.13	0.5206	3.06	
Woman received pain relief medication	640	59.38	17.50	85.71	46.21	59.38	0.6596	3.39	AUC
Maternal Outcomes									
Cesarean section (C/S) performed	651	13.52	13.36	93.10	98.76	13.52	0.9593	1.01	Yes
Reason for C/S- prolonged/obstructed labor	76	32.89	67.11	39.22	80.00	32.90	0.5961	0.49	
Complications (any)	654	44.80	11.00	62.50	57.39	44.80	0.5994	4.07	
Hemorrhage	654	11.16	4.59	33.33	89.90	11.17	0.6162	2.43	AUC
Prolonged labor	654	23.70	3.67	50.00	77.30	23.70	0.6365	6.46	AUC
None	654	51.53	88.99	53.78	66.67	51.53	0.6022	0.58	AUC

Notes: Recommended indicators meet both AUC and IF validation criteria.

* Excluding 'Don't Know' responses

^a Skilled provider includes doctor (ob-gyn), medical resident or nurse/midwife

^b Indicator constructed from two skin-to-skin items: (1) baby placed against mother's chest after delivery and (2) baby was lying naked against the mother's chest.