2014

Validating indicators of the quality of maternal health care: Final report, Mexico

Karla Berdichevsky

Claudia Diaz
Population Council

Katharine McCarthy
Population Council

Ann K. Blanc
Population Council

Final Report

VALIDATING INDICATORS OF THE QUALITY OF MATERNAL HEALTH CARE: FINAL REPORT, MEXICO

Karla Berdichevsky
Claudia Diaz-Olavarrieta
Katharine McCarthy
Ann Blanc

POPULATION COUNCIL
Ideas. Evidence. Impact.

2014

The Population Council confronts critical health and development issues—from stopping the spread of HIV to improving reproductive health and ensuring that young people lead full and productive lives. Through biomedical, social science, and public health research in 50 countries, we work with our partners to deliver solutions that lead to more effective policies, programs, and technologies that improve lives around the world. Established in 1952 and headquartered in New York, the Council is a nongovernmental, nonprofit organization governed by an international board of trustees.

Population Council
Insurgentes Sur #2453 Piso 9
Obregónn Benito Juárez, Mexico City
Mexico, D.F., 1090

Tel: +52 55 5999 8645
Fax: +52 55 5999 8631

popcouncil.org

Suggested citation: Berdichevsky, K., Diaz-Olavarrieta, C., McCarthy, K., and Blanc, A. 2014. “Validating Indicators of the Quality of Maternal Health Care: Final Report, Mexico.” Mexico City: Population Council.

Table of Contents

Executive Summary 3

List of Abbreviations 6

Tables and Figures 7

Study Background 9

Study Overview 11

Ethical Clearance 12

Methods 13

Research Site & Study Population 13

Recruitment and Data Collection Procedures 13

Results 17

Study Sample and Descriptives 17

Indicator Validity Results 20

Discussion 45

Literature Cited 51

Appendices

Annex A, Table 1	List of Assessed Indicators and Measured Coverage, Matched Data
Annex A, Table 2	Near Universal and Rare Indicators, Unmatched Data
Annex A, Table 3	Indicators with High 'Don't Know' Responses, Unmatched Data
Annex B, Table 1	Indicator Validation Results, Matched Data

Acknowledgments

We would like to thank and acknowledge all those who contributed to this project and who are dedicated to improving the measurement of the quality of maternal health care. We particularly thank Saumya RamaRao and Charlotte Warren who were co-investigators on the project. These individuals provided leadership, technical support, and insights throughout the project. We acknowledge Hannah Taboada who conducted a comprehensive literature review of indicators to be assessed in the study and developed the initial tools used in the study. We also acknowledge Tahilin Karver who provided assistance to study implementation. We appreciate the advice of the maternal health experts who advised on the study and are grateful to the women and providers who allowed us to share their experiences. The study was supported by the Maternal Health Task Force through a grant to the Harvard School of Public Health from the Bill & Melinda Gates Foundation.

Executive Summary

Despite widespread use, the majority of indicators proposed as measures of the quality of maternal health services have not been sufficiently validated. To help accurately track progress towards national and global maternal health goals, the present study sought to validate and identify a set of maternal health indicators that can be practically applied in facility and population-based surveys. To evaluate the indicators, the study employed a facility-based design. The study was conducted in public /government hospital facilities in Kenya and Mexico. Participants included women aged 15-49 who underwent labor and delivery at participating study facilities and the providers who attended them. Women's self-report of obstetric and immediate postnatal maternal and newborn care received was compared against a "gold standard" of observations by a trained third party observer during labor and delivery.

This report presents results of the Mexico study. Data collection took place between November 2013 and April 2014. A total of 600 births were observed and the mothers participated in an exit interview prior to hospital discharge. A large proportion of assessed indicators were either routinely practiced or rarely occurred. The lack of variation in observed interventions limited the ability to conduct full validity analysis for some indicators. Of the 108 indicators assessed, 48 had sufficient variation for validity analysis using area under the receiver operator characteristic curve (AUC) and inflation factor (IF) analysis.

Of assessed indicators, 5 met both acceptability criteria for both AUC (AUC>0.6) and IF (0.75<IF<1.25). These were: whether a urine sample was taken upon hospital admission, whether an injection or IV medication was received at some time during labor (before the birth of the baby), episiotomy, hemorrhage, and receipt of blood products. Findings suggest events that caused pain, concern or were considered 'important information' by mothers, were particularly salient for women and may have enhanced recall. Important to note is that although a urine screening test, hemorrhage and receipt of blood products each met both validity criteria, these indicators had moderate or low sensitivity (50% or lower). As the lower sensitivity and specificity for these indicators balances out at the aggregate level, these indicators may be more appropriately applied to estimate the population-based coverage of these events. Also of note is that while an indicator of whether an injection of IV medication was received at some time during labor met both validity criteria, additional findings indicate that women are unable to report on the indication for the medication (if received for induction or augmentation of labor). Taken together, these indicators are recommended dependent on the objective of their use (i.e., for accurately

classifying ‘true positive’ and ‘negative’ cases or for obtaining an approximate population-based prevalence).

A total 28 indicators met acceptability criteria for one of the validity measures; 19 indicators met the IF criteria and 9 met the AUC criteria only. While having high sensitivity and specificity for indicators is important in ascertaining which women received care at an individual level, approximating the broader population-based prevalence can also provide actionable data on the coverage of maternal and newborn health care received. For example, in some cases, underreporting of true cases and false positive reporting of negative cases (i.e., low indicator sensitivity and specificity) equaled out to generate acceptable estimates for monitoring coverage at the population level. We recommend caution with regard to low-prevalence indicators that, without near-perfect specificity (i.e., true negative classification), are likely to be overestimated.

Of key objective of the present study was to validate the skilled birth attendance indicator. Attendance by a skilled provider (defined as any doctor- specialized or a general practitioner, medical resident or nurse) was near universal and unable to be robustly analyzed. Cross-tabulation results, however, indicate that the majority of women observed to have skilled birth attendance, correctly reported receiving skilled care (high sensitivity). We also assessed if women could accurately ascertain what category of provider was responsible for the majority of care. We found a combined doctor (any type) and medical resident indicator could be reported with accuracy at the population-level, but was not suitable for individual-level classification given the tendency for overreporting by women. Given the little distinction between skilled doctor categories however, the inability of women to distinguish between finer distinctions of providers may be less programmatically meaningful than the ability to report on the coverage of the indicator at the aggregate level.

Although it was not possible to validate indicators with near universal practice, descriptive cross-tabulation results suggest that women may be able to report on some aspects of routine care with accuracy. For example, women reported on their newborn’s birthweight with near perfect classification, although so few women incorrectly reported their newborn birthweight that robust analysis was not possible. An additional indicator of potential high use not able to be assessed in the present study was the type of facility where women delivered. Although the present study was not designed to evaluate this indicator, descriptive results suggest a high proportion (85%) of women correctly identified the type of institution where they delivered. Results also suggest that the two-part question methodology to identify institution type, also used in DHS and MICS surveys, is important. These indicators should be explored further in studies with multiple facility types (private and public sector), or different practices for recording newborn weight.

In sum, study findings suggest women in a facility-based setting validly report 5 indicators of maternal health services and immediate newborn care. An additional 19 indicators met the IF criteria only and 9 met AUC criteria only. We recommend the use of these indicators with caution and dependent on whether the purpose is to identify coverage at the population level, or to distinguish among true positive and negative cases at an individual level. Results suggest indicators related to timing, technical terminology, and the sequence of events may be reported with difficulty. Taken together, findings suggest the validity of a number of indicators may be highly dependent on context and question wording. Future studies should explore how key terms and questions related to timing and order of events are understood by women in order to enhance indicator accuracy.

List of Abbreviations

AMTSL	Active management of the third stage of labor
APH	Antepartum Hemorrhage
C/S	Cesarean section
DHS	Demographic and Health Surveys
HGM	Hospital General de México, Dr. Eduardo Liceaga
IM	Intramuscular
IRB	Institutional Review Board
IV	Intravenous
MDG	Millennium Development Goal
MICS	Multiple Indicator Cluster Surveys
MNCH	Maternal, Newborn and Child Health
MOH	Ministry of Health
OB-GYN	Obstetrics and Gynecology
PPH	Postpartum Hemorrhage
SBA	Skilled Birth Attendant
SRH	Sexual and Reproductive Health

Tables and Figures

Report

Figure 1	Participant Response Rates, Mexico	17
Table 1	Sample Background Characteristics	18
Box 1A-B	Descriptive Frequencies: Type of Facility	21
Box 2A	Cross-tabulation: Urine Sample Taken at Admission	22
Box 3A	Cross-tabulation: Induction of Labor by Uterotonic	24
Box 4A	Cross-tabulation: Augmentation of Labor by Uterotonic	25
Box 4B	Cross-tabulation: Uterotonic for Induction or Augmentation of Labor (General Indicator)	26
Box 4C	Cross-tabulation: Uterotonic for Induction or Augmentation of Labor (Composite Indicator)	26
Box 5A	Cross-tabulation: Received Prophylactic Uterotonic	27
Box 5B	Cross-tabulation: Timing of Prophylactic Uterotonic (1-3 Minutes Post-Delivery)	28
Box 5C	Cross-tabulation: Oxytocin Received After Birth	29
Box 5D	Cross-tabulation (Self-Report): Prophylactic Uterotonic and Oxytocin	30
Box 6	Cross-tabulation: All 3 Components of AMSTL Received	31
Table 2	Cross-tabulation of Main Provider During Delivery	33
Box 7	Cross-tabulation: Main Provider during Delivery- Doctor/Ob-gyn or Medical Resident	33
Table 3	Cross-tabulation of Main Provider During Labor	34
Figure 2	Scatterplot of Woman and Observer Report on Number of Providers Who Attended the Birth	36
Box 8	Cross-tabulation (Self-Report): Newborn Skin to Skin Indicators	38
Box 9A	Descriptive Frequencies: Complications	40
Box 9B	Cross-tabulation: Complication of Hemorrhage	40

Box 9C	Descriptive Frequencies: Post-Delivery Health Checks	41
Box 10	Cross-tabulation: Stillbirth delivery	44
Table 4	Indicators that Met Both Validation Criteria	45

Appendices

Annex A, Table 1	List of Assessed Indicators and Measured Coverage, Matched Data
Annex A, Table 2	Near Universal and Rare Indicators, Unmatched Data
Annex A, Table 3	Indicators with High ‘Don’t Know’ Responses, Unmatched Data
Annex B, Table 1	Indicator Validation Results, Matched Data

Study Background

Global monitoring of the percentage of women who have received quality maternal health services is crucial to guide the scale-up and allocation of resources to reduce 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 widely used indicators to measure 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.

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 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, identifying the actual interventions that a woman receives is necessary to provide a more 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 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 Republic[8], 2) indicators of care received by women and their newborns during labor, delivery, and the postnatal period in Mozambique,[9] and (3) indicators of care received by women and newborns in rural China.[10] In these studies, women's reporting of events during 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 previous 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 provides insight into factors (e.g., participant variables, type of delivery, instances of complications or other events) that may influence the accuracy of recall. The results of the study inform the recommendation of a select number of indicators that have the potential for valid and reliable measurement and for integration into routine population-based and facility-based data collection systems.

Mexico and Kenya were chosen as study sites in light of variations in the status of maternal health and the coverage and organization of maternal health services. This report presents results from the Mexico study.

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:

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. We conducted additional searches of publications from 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 on the basis 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 indicators for validity testing (see Annex A, Table 1). 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 the 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.

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 the participating hospital facility, *Hospital General de México "Dr. Eduardo Liceaga",* Mexico City, Mexico.

ETHICAL CLEARANCE

The protocol was approved by the Population Council's Institutional Review Board in May 2013 (IRB Protocol 594) and by the Ethics and Research Committees of the Hospital General de México "Dr. Eduardo Liceaga" in October 2013. No participants were enrolled in the study until ethical approval was obtained from both ethics committees.

Ethics and Research Committees Research Division
Hospital General de México "Dr. Eduardo Liceaga"
Dr. Balmis 148
Colonia Doctores
Mexico City, 06726
Mexico
Tel. +52 55 2789200 ext. 1164

Methods

RESEARCH SITE & STUDY POPULATION

Study Population:

The study population consisted of women whose births were documented by study data collectors at the participating study facility between November 2013 and April 2014. Women aged 15-49 who were admitted for labor at the study facility and who consented to study participation were eligible for inclusion. The study population also included the providers who attended participating women in labor and delivery and whose labor and delivery care was observed by study data collectors.

Study Location: *Hospital General de México Dr. Eduardo Liceaga* (HGM), Mexico City

All women were recruited from the above hospital, a public facility in Mexico City. HGM provides comprehensive obstetric care to women with normal pregnancies who are self-referred for admission, in combination with high-risk pregnancies that are referred from other public primary or secondary health care institutions. The hospital population in Mexico City tends to have a lower-than-average socioeconomic status and characteristically lacks health insurance. Patients may travel large distances to HGM. Slightly more than half of the patients who receive health care at HGM live in Mexico City, with a significant proportion from the neighboring State of Mexico (37%) and 5% from the rest of the country.[14]

On average in Mexico, hospital-based deliveries are proportionately high. Data from the most recent national health survey indicate that 94% of women delivered with a medical doctor.[15] HGM is one of the public hospitals in Mexico City with the highest volume of obstetric patients and in 2013 provided delivery care for 4169 women, of which 2235 (54%) were vaginal births.

RECRUITMENT PROCEDURES

Data Collectors

Two types of data collectors were involved in this study: observers, who registered the quality and content of obstetric care provision; and interviewers, who applied questionnaires to women prior to hospital discharge. Observers were general medical practitioners or nurses with sufficient clinical training to accurately observe delivery room practices and procedures and to know how to conduct themselves professionally. Interviewers were social workers and psychologists with qualitative research experience. We recruited only female interviewers in order to facilitate rapport with participating women. All data collectors were external to HGM and not personally known to hospital staff in order to minimize observer bias and to ensure respondents' anonymity. In addition, we selected a study coordinator from the pool of observer applicants. The study coordinator had previous research and managerial experience and provided supervision and support to fieldwork teams.

Job openings for data collectors were posted on university websites and list serves. Interested applicants who fulfilled our criteria were interviewed and asked for professional references. Most of our observers held part-time jobs elsewhere or were studying for admission into medical residency programs. Their shifts were thus scheduled around their availability to give them flexibility to continue with those activities.

Data Collector Training

Data collection training for observers and interviewers took place over two consecutive days in October 2013. Training included detailed study protocol description, practice using data collection instruments, and key ethical aspects of research, stressing the importance of informed consent. Sessions included teaching tools such as role-playing dynamics and question-and-answer sessions to clarify technical concepts. All field staff were required to complete and submit a certificate of the NIH online training course "Protecting Human Research Participants" (https://phrp.nihtraining.com/users/login.php) to ensure that they fully understood and adhered to human subject research ethics. Training manuals were developed and distributed to all data collectors, who were instructed to use them as a reference during study implementation.

Study interviewers were trained to listen and observe intently, without displaying judgmental attitudes towards information they received. They were instructed to read out questions as worded in the questionnaires and to offer additional explanations only when strictly necessary. Observers received training on procedures for being unobtrusive and for locating themselves toward the head of the client rather than the foot when possible. Observers also received information on how to conduct themselves in the event that they witnessed unacceptable provider behavior that put clients at risk of morbidity and

how and when to intervene. Both interviewers and observers were instructed to keep all data collection instruments in strict confidentiality.

Because all but two of our observers and our study coordinator took new jobs or began their medical residencies in January 2014, we conducted new recruitment and training to replace these staff. The second training was a one-day workshop that covered the same content as the original training. Contrary to what occurred with the observers, the interviewer group remained throughout the study period. Although no refresher trainings were conducted as a group, our study coordinators gave constant feedback to observers and interviewers and were available to respond to their questions and solve problems as they arose.

Process for Participant Recruitment and Informed Consent

All women who were undergoing labor and delivery at HGM and met the inclusion criteria were eligible for study participation. Inclusion criteria were: 1) women aged 15-49 years old (the same age range as in the DHS), 2) admitted for delivery at HGM and able to provide consent, and 3) the health care providers who attended the woman in labor and delivery. Women were excluded if they were unable to provide consent (including unconsciousness or presenting with a complication) or if their stage of labor was considered too advanced by the attending medical personnel. Women were eliminated from the study if a cesarean section was indicated at any point during labor.

When women were admitted to the labor ward, health providers asked them whether they were willing to meet with a member of the research team. If the woman gave her permission to speak with an observer, the health provider identified an appropriate time (relative to her clinical status) to meet with her. The observer then explained the study procedures, including the fact that a researcher would stand in the corner of the room during her labor and delivery and use a checklist to record the actions of the doctors or nurses.

Written informed consent was then obtained from interested participants and, in accordance with local ethical guidelines, consent was also requested from their spouses/common law partner or parents when applicable. For participants who were minors, information about the study was also provided to, and written consent sought from, their parent or spouse/ common law partner (as responsible parties). Everyone who participated in the informed consent process was given a copy of the consent form with the study information and signature page. The main informed consent process included both study

activities: observation and exit interview. Before each data collection activity, study interviewers confirmed that women still wished to participate and requested their verbal consent.

Provider consent was obtained to observe health care workers in the labor and delivery ward. Before study implementation began, several meetings were organized by the OB-GYN director for our research team to explain the study to nurses and doctors, answer their questions, and obtain their informed consent. Once the study began, the study coordinator and observer team individually approached providers who had not attended the larger staff meetings to obtain their consent. This was especially important since new hospital personnel were constantly recruited.

A list of providers who gave their consent was available for observers to check against before their observations. Consultations were observed only if provider consent had been obtained. Very few providers did not agree to participate in the study and, in those cases, we did not conduct observations of the services they provided.

Data Collection Procedures

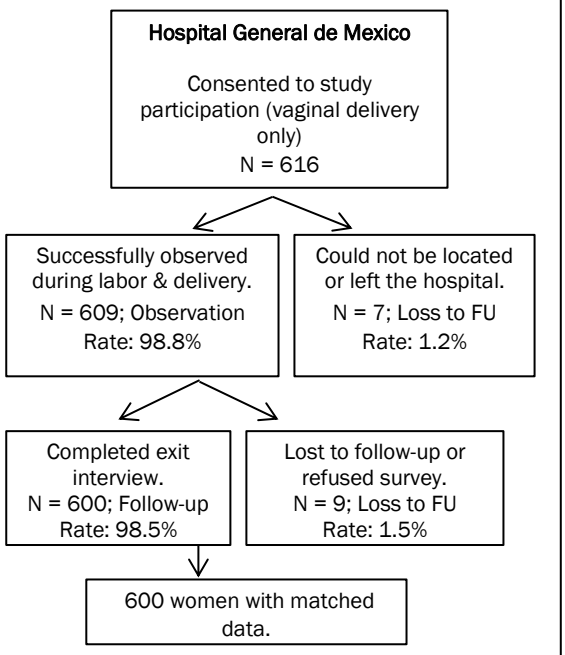
Data collection was carried out from 1 November 2013 to 23 April 2014. Data were collected continuously, seven days per week, 24 hours per day. There were 3 shifts for staff carrying out observations, with two data collectors per shift. On average, observers worked 3 to 4 shifts per week, while interviewers worked individually and were assigned an average of two shifts per week. Interviewer shifts (10 am to 5 pm) were designed to cover the period when patients were discharged from the hospital. Data collection was briefly interrupted at the beginning of February, as our first group of observers left the study and we trained a new team. Data collection was also suspended during national holidays.

With the objective of gaining better insight into data quality, we conducted debriefing sessions with the fieldwork teams two weeks after concluding data collection. We carried out two separate group interviews: one per observer and interviewer teams. The information collected was used to interpret our quantitative findings.

Results

STUDY SAMPLE

Figure 1. Participant response rates (vaginal delivery), Mexico.



A total of 779 women admitted for labor at Hospital General were recruited to participate. Because of hospital policies that allowed the complete observation only of women who underwent vaginal deliveries, data collection for women who became indicated for cesarean sections was discontinued at the time of indication. **Figure 1** depicts a flow chart of data collection for women who could be observed throughout their labor and delivery (i.e., women who delivered by vaginal birth). Specifically, 616 women who consented to participate in the study delivered by vaginal birth. Of these women, 609 (99%) were successfully observed during labor and delivery. A total of 7 (1%) women were not observed as a result of being sent home because they did not progress into labor, because they could not be located after recruitment, or because they requested to be discharged from the hospital prior to delivery. Of the women whose labor and delivery was observed, 600 completed an exit interview prior to hospital discharge. Exit interviews were not completed for 9 women (2%) who either refused participation at the time of interview or because they received an early hospital release. 600 observer reports and client exit interviews were accurately matched and could be analyzed. In the following sections, all data refer to women who delivered by vaginal birth only, unless otherwise specified.

SAMPLE DESCRIPTIVES

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. ±6) and ranged between 15 and 42 years. The majority of women (56%) were living with a partner (*en unión libre*) or were currently married (18%), whereas one-quarter were single and had never been married (26%). Approximately half of women had given birth previously (52%). Of women who had previously delivered, most of the women had one prior birth. The highest parity among women was seven prior births. Nearly all of the women reported secondary school education or higher as their highest level of completed education (92%). Of these women, approximately 40% had completed or obtained at least some post-secondary education (42%)

or tertiary/ preparatory school (41%), and less than 10% had completed or obtained at least some education at the university or graduate level (8%). Of observed vaginal deliveries, 3% were assisted with forceps.

Study participants who became indicated for cesarean operation after study enrollment (and for which observation was discontinued at the time of indication) did not differ from all other women on key sociodemographic characteristics such as age, marital status, or education level. Women who underwent a cesarean operation had, on average, fewer previous births than women who delivered vaginally (mean difference: 0.3± Margin of Error 0.1, $p<0.001$).

Maternal and Newborn Outcomes

A total of 604 deliveries were observed, including 2 sets of twins. Less than 1% of births resulted in a stillbirth (n=2) or neonatal death (n=1) (i.e., newborn died within the first hour of delivery).

The majority of participants reported coming to the facility because they had planned to deliver there (72%), while 28% reported coming because they had a problem, such as needing to be referred to an alternate facility for medical care or an inability to pay. Complications were observed among 14% of women. Most complications were observed following delivery (44%), while approximately 40% occurred during delivery (38%), and 19% occurred prior to delivery. The most common type of complication was hemorrhage (APH+PPH) (8% of women), followed by eclampsia (2%), and prolonged labor (>12 hours), (2% of women). Less than 1% of women with complications experienced more than one.

TABLE 1. Sample background characteristics*.

Age	Percentage (%)
15-19	27.2
20-24	36.3
25-29	19.2
30-34	9.7
35-39	5.9
40+	1.6
Prior parity (total # live births)	
0	47.7
1	29.4
2	14.7
3	5.6
4 or more	2.6
Educational level	
None	0.2
Primary	8.4
Secondary	42.3
Higher	49.2
Marital status	
Single, never married	25.5
Married	17.8
Living together	55.7
Separated	1.0
Widowed	0.0
Type of vaginal delivery	
Unassisted	97.5
Assisted forceps	2.5

Note: *Data refer only to women who delivered vaginally.

Service Delivery Coverage

The indicators selected for validity testing are presented in **Annex A, Table 1**. The table 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 of indicators ranged from extremely rare (<5%) to near universal (>95%). The broad range indicates that some preventive interventions were almost always implemented, while other harmful practices rarely occurred. Indicators meeting these criteria are listed in **Annex A, Table 2**.

A total of 18 indicators had an observed prevalence >95%, while 13 indicators had an observed prevalence <5%. Universal or rare practices reflected standard hospital practices or policies. For example, it was the standard practice for newborns to be wrapped in a towel and then placed with the mother, rather than being placed directly naked against the mother's chest (skin-to-skin) and then covered with a towel or cloth. Other standard practices were that women were generally not allowed to eat or drink during labor or to have a support companion present during labor or delivery.

Of note is that few HIV tests were offered (2%). This was in part because few HIV tests were available at the facility, and because the practice of first checking the woman's HIV status by either asking the woman or consulting her records was high (70%) (n=169). Similarly, few women were observed to receive HIV testing (1%). A greater proportion of women who were offered HIV testing were observed to receive the test (3 of 10 women; 30%). Provider hand-washing practices were also notably low. In contrast to the low level of observed hand-washing, there was a near universal practice of providers wearing high-level disinfected or sterile gloves during examinations of the woman.

For all indicators, women and observers were given the option to respond, "Don't Know". The proportion of women who responded "Don't Know" to indicators was minimal (<5%). Indicators where the proportion of women who responded "Don't Know" exceeded 5% are reported in **Annex A, Table 3**.

The greatest proportion of women responded "Don't' Know" to the indicator, "*Did anyone give you a medication called 'oxytocin' to make your womb contract or become firm?*" (37% "Don't Know"). Potential explanations for this high percentage are that women were not informed if they were given oxytocin or that women did not know what the drug was for and its name was not salient. Other practices with a high self-reported "Don't Know" prevalence were also related to indicators of care received in the immediate postnatal period. This included the oxytocin indicator, as well as whether a uterotonic for the prevention of postpartum hemorrhage (PPH) was received following delivery– whether in the first few minutes after delivery, "*anyone [gave] medication intravenously through a tube in your arm*" (7% "Don't Know"). At the same time, observer reports indicate that nearly all women received the prophylactic uterotonic oxytocin following delivery, and that it was administered via an IV line in the arm. Taken together, these descriptive findings suggest that many women have difficulty reporting on

indicators related to receiving a uterotonic for the prevention of postpartum hemorrhage. High "Don't Know" responses were also given for other postnatal practices related to immediate newborn care and postnatal health checks.

The percentage of women who were able to report on whether anyone offered them an HIV test was also notably low (23% "Don't Know"). Study data collectors noted that many women did not know what HIV was, or confused it with human papillomavirus (HPV), which may account for the high "Don't Know" percentage. Indicators on whether the provider washed his or her hands before examining her or whether the newborn was given anything to drink besides breastmilk in the first hour after delivery may reflect uncertainty due to the practice occurring outside of the woman's view.

Observer responses of "Don't Know" were also minimal across indicators and, where they occurred, were generally less than 2% of responses (**Annex A, Table 3**). Six indicators had an observer report of "Don't Know" of 3% or greater. These indicators generally related to practices that may have occurred outside of the labor and delivery room, such as whether the baby was given anything other than breastmilk to eat within the first hour after birth (6% "Don't Know"), whether the provider checked the woman's HIV status, or whether the provider washed his or her hands before examining the woman (4% and 3% "Don't Know").

VALIDITY OF QUALITY OF CARE INDICATORS

Analytic Approach

The general approach for the validity analysis relied on comparing women's self-report of each quality of care indicator to its "true" classification according to the observer report (i.e., reference standard) at the time of facility-based delivery. Using these data, sensitivity and specificity for each indicator were calculated using two-by-two tables. An indicator's sensitivity is the proportion of actual positives (women who received care) that were correctly identified as such (i.e., "true positive rate"). An indicator's specificity is the proportion of negatives (women who did not receive care) that were correctly identified as such (i.e., "true negative rate"). For indicators meeting the criteria of at least 5 counts per cell, we also calculated the area under the receiver operating curve (AUC) and the inflation factor (IF).

Receiver operating curve analysis is a valuable method to describe the accuracy of diagnostic tools by plotting the tradeoff between sensitivity (true positive rate) against 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 test's sensitivity and specificity by a single number.[16-18] An AUC of 1.0 can be interpreted as a test with perfect accuracy, while an AUC of 0.5 represents a random guess.[19] 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.[20] By comparing the ratio of the estimated survey-based prevalence to its true population prevalence (observer report or 'reference standard'), we estimate the degree to which each indicator would be over- or under-estimated if assessed using a population-based survey.[21]

A priori benchmark criteria for 'valid' indicators were an AUC>0.6 and an IF between <1.25 and >0.75 and were informed by criteria previously used in the literature.[9,10] Use of the two methods provides complementary data to inform 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 refer to matched indicator data. In the results section that follows, the AUC and its margin of error are also reported.

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	380	85.2
Govt. clinic/health center	7	1.6
Govt. health dispensary	0	0
Other public sector	11	2.5
Private Sector		
Private hospital	45	10.1
Private clinic	3	0.7
Private maternity home	0	0
Other private sector	0	0
Total Reported in Categories	**446**	**100**

Box 1B: Woman not able to determine whether private or public but specified facility details.

	Number	Percent
Facility name (Hospital General de Mexico)	120	83.3
Hospital and public/ govt. type	5	3.5
Hospital (other)	13	9.0
Other facility detail (location, level, etc.)	6	4.2
Total Specified	**144**	**100**

Facility Arrival & Initial Assessment

Type of facility

Of interest is whether women can accurately report on the type of facility where they delivered. Replicating the methodology of DHS and MICS surveys, women were asked to identify where they gave birth by first indicating whether they delivered in a public or private sector institution. Women were then asked to specify the facility type as a hospital, health clinic/center, health post, or other location. If women were unable to determine whether the facility was public or private, they were asked to name the facility. Since the participating study facility was a public/government hospital (100% true or '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 women who classified the facility as public or private sector and then specified the level of facility, 85% of women correctly classified the type of facility as a public hospital (**Boxes 1A-B**).

24% of women (n=144) were not able to classify *Hospital General* as a private or public sector facility but did specify the facility name or other details. 83% of these women correctly specified the facility name as "*Hospital General*" or "*Hospital General de México*". An additional 9% of women who could not identify the type of facility were able to report it was a hospital of some kind (but not a public sector or governmental hospital). Notably, 4% of women were able to report the facility was a hospital and part of the public or government sector (e.g., by reporting "*Hospital público*", "*Hospital federal*", or "*Hospital de asistencia pública*"). The fact that women could identify the facility as a government or public sector hospital by being asked to report the facility name in an open-ended question but were not able to classify the facility using pre-existing categories suggests that more women are able to accurately report on the facility type than are captured in initial questioning. Some women (4%) also provided additional correct details about the facility, such as its level, location, or ability to offer specialized care, but not whether the facility was in the public sector or a hospital.

TYPE OF FACILITY: SUMMARY

Descriptive results suggest that both parts of the DHS and MICS indicator question methodology – categorical responses and specific facility names - are important in capturing self-reported information on the type of facility. These results should be interpreted with care as they are descriptive only.

Initial assessment practices

Two indicators of the initial client assessment phase could be assessed. Other practices were near universal or rarely occurred and did not meet criteria for robust analysis. Assessed indicators in the initial client assessment phase were: (1) whether the woman's HIV status was checked and (2) whether the woman's urine sample was taken upon admission.

Box 2A: Cross-tabulation: Urine sample taken at admission.

Did someone ask you to give them a urine sample at or near your admission to the facility?

		Observer Report (Number)		
		No	Yes	**Total**
Self-report (Number)	No	361	72	433
	Yes	67	72	139
	Total	428	144	572

Only one indicator, whether the woman's urine sample was taken, met both study validity criteria (AUC: 0.6717± 0.04, IF: 0.97). A urine sample screen is an early test for preeclampsia/ eclampsia, or

elevated blood pressure in the mother which can lead to deadly seizures, one of the leading causes of maternal mortality and morbidity. Despite meeting both study validity criteria, the urine screen indicator had moderately low sensitivity (50%), indicating that only half of women whose urine sample was taken reported that the intervention took place (**Box 2A**). However, the indicator did have high specificity (84%), indicating that a large proportion of women who did not have their urine sample taken accurately reported that the intervention did not occur. Since urine screening test can identify a potentially life-threatening complication, the low sensitivity of the indicator signifies it would not capture critical information.

The indicator "HIV status checked" met only the IF validity criteria. The underreporting of true cases (55% sensitivity) and false positives among negative cases (1-specificity, 52%), however, balances out, so the IF approximates the population-based prevalence that would be obtained via household survey (AUC: 0.5129± 0.04, IF: 0.78).

Initial assessment practices that we were unable to assess by robust analysis because of near universal or rare occurrence of the intervention included: (1) provider hand washing or antiseptic use before initial examination of the woman, (2) whether high-level disinfected or sterile gloves were worn for vaginal examination, (3) whether the woman's blood pressure was taken, (4) whether the woman was offered or (5) received an HIV test, and (6) whether the fetal heart rate was checked with a fetoscope/ultrasound (**Annex A, Tables 1-2**). Descriptive results for these indicators show that nearly all women received and positively reported receiving these standard practices of care. However, of the relatively few women who did not receive such interventions, nearly all falsely reported that these interventions took place, resulting in low specificity rates, <10%. As a result, women overestimated indicators of provider hand washing or antiseptic use (96% self-report prevalence compared to 2% observed prevalence), being offered an HIV test (14% self-report prevalence; 2% observed prevalence), and receiving an HIV test (16% self-report prevalence; 1% observed prevalence).

INITIAL ASSESSMENT: SUMMARY

1 of 2 assessed indicators of facility-based care received met study criteria for valid measurement— whether a urine sample was taken. While this indicator had high accuracy in accurately classifying women who did not receive the intervention, its sensitivity of 50% suggests that nearly half of women who received the intervention were not identified. An indicator on whether HIV status was checked met the IF criteria only.

Respectful Care

Five indicators reflected aspects of women-centered care (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 or birth.

In general, the majority of proxy measures of respectful or woman-centered care were not observed to take place (e.g., being able to eat or drink during labor, being allowed or having a support companion during labor and delivery). A correspondingly low proportion of women reported that such interventions took place, limiting the ability to conduct full analysis. Only one indicator, whether the woman was encouraged or assisted to ambulate during labor, had sufficient numbers for robust analysis. Specifically, being encouraged or assisted to ambulate in labor was observed to occur among 22% of women, while it was reported by only 8%. This indicator did not meet study validity criteria (AUC: 0.4966± 0.04, IF: 0.37) and was underreported by women.

PROVIDER RESPECTFUL CARE: SUMMARY

Only 1 of 5 indicators could be assessed by robust analysis— whether the woman was encouraged to ambulate during labor. This indicator did not meet study validity criteria and was particularly underreported by women.

Induction / Augmentation of Labor

Induction of labor

Box 3A. Cross-tabulation: Induction of labor by uterotonic.

2-item indicator constructed: 1) someone intervened to bring on labor and 2) medication through IV was given to bring on labor. (Proxy for observer report of uterotonic given).

		Observer Report (Number) *Induces labor with uterotonic*		
		No	Yes	**Total**
Self-report (Number)	No	440	47	487
	Yes	54	18	72
	Total	494	65	559

Approximately 17% of all women reported "*someone did something*" to bring on their labor, rather than their labor starting spontaneously. Of women who reported that someone intervened (n=91), the majority (79%) reported receiving medication through an IV line in their arm; while 21% reported their membranes were ruptured. We were unable to assess the validity of receiving membrane rupture for induction of labor, since observers

did not record the indication for receiving this intervention (i.e., whether performed for induction or augmentation of labor).

To assess the validity of 'received a uterotonic for labor induction', we constructed a two-item indicator of women who reported "*something was done to bring on labor*" and who subsequently reported receiving a uterotonic (i.e., medication through an IV line, as no option for intramuscular injection was given) to bring on labor (13%). This indicator was compared to observer reports of whether the woman received a uterotonic to induce labor (12%) (**Box 3A)**. This indicator did not meet the AUC validity criteria (AUC: 0.5838± 0.04, IF: 1.11). In particular, this indicator had low sensitivity (28%), indicating that not all women who received a uterotonic for labor induction reported receiving the intervention (**Annex B**). However, the indicator did meet the IF criteria, indicating that it may be acceptable to approximate the population-based prevalence.

We also sought to assess whether women could accurately report on the method by which a uterotonic for labor induction was received. Robust analysis, however, was not possible since all observers indicated that women received the uterotonic through an IV line.

Augmentation of labor

More than half (52%) of all women reported "*someone did something to strengthen or speed up*" their labor. The most common method of uterotonic administration reported for augmentation of labor was an IV line inserted in the arm (both womens' and observers' report).

A two-item indicator in which women who reported "*something was done to speed up or strengthen labor*" and who subsequently reported receiving a uterotonic (i.e., an injection or medication through an IV line) to strengthen labor (42%) was compared to the observer report of the woman receiving a uterotonic to augment labor (75%) (**Box 4A**). This indicator did not meet the IF criteria (AUC: 0.6539± 0.04, IF: 0.56). Similar to receiving a uterotonic for induction of labor, this indicator had relatively low sensitivity (50%), indicating that not all women who received a uterotonic for the augmentation of labor reported the intervention.

Box 4A. Cross-tabulation: Augmentation of labor by uterotonic.

2-item indicator constructed: 1) someone intervened to strengthen labor and 2) received medication through an injection or IV line to speed up or strengthen labor.

(Proxy for observer report of uterotonic given).

		Observer Report (Number) *Augments labor with uterotonic*		
		No	Yes	**Total**
Self-report (Number)	No	95	179	274
	Yes	22	176	198
	Total	117	355	472

All women who received a uterotonic for labor augmentation received it by IV line, so robust analysis for this indicator was not possible. Cross-tabulation results suggest that women are not able to accurately report on the method of augmentation of labor. Specifically, only slightly more than half of women (57%) who were observed to receive medication via IV line for labor augmentation reported receiving the intervention via this route.

Injection or IV medication during labor

In addition to being asked to specify the method by which their labor was brought on or strengthened, women were asked a general indicator on whether "*before the birth of the baby, did you receive an injection or IV medication during labor?*" This indicator was used as a proxy measure for receiving a uterotonic for the induction or augmentation of labor (no providers administered uterotonic in tablet form). In contrast to the observer report on receiving a uterotonic (by injection or IV line) for labor induction or augmentation, this indicator met both study validity criteria (AUC: 0.6349± 0.04, IF: 0.84) (**Box 4B**).

Box 4B. Cross-tabulation: Uterotonic for induction or augmentation of labor (general indicator).

Before the birth of your baby, did you receive any injection or IV medication during labor?

Self-report (Number)	Observer Report (Number) *Composite indicator of whether observer reported IV or injection for induction or augmentation of labor.*		
	No	Yes	**Total**
No	78	130	208
Yes	59	304	363
Total	137	434	571

A comparison indicator (which was constructed in analysis to combine women's responses to two survey questions) that compared women's self-report of whether an injection or IV medication was received for either induction or augmentation of labor did not meet either study validity criteria (AUC: 0.6947± 0.04, IF: 0.65). Specifically, this indicator had lower sensitivity (59%), indicating that not all women who received a uterotonic by injection or IV medication reported receiving either intervention for the purposes of bringing on or strengthening labor (**Box 4C**).

Box 4C. Cross-tabulation: Uterotonic for induction or augmentation of labor (composite indicator).

Composite indicator of whether women reported receiving an injection or IV medication for induction or augmentation of labor (constructed in analysis).

Self-report (Number)	Observer Report (Number) *Composite indicator of whether observer reported IV or injection for induction or augmentation of labor.*		
	No	Yes	**Total**
No	91	148	239
Yes	23	214	237
Total	114	362	476

Taken together, these results suggest that at an individual classification level women can report whether they received an injection or IV medication at some point during labor before delivery, but not the indication for the medication (whether received for induction or augmentation of labor). An indicator on whether a uterotonic was received for the induction of labor may be appropriate for measuring population-based coverage of this practice.

Membrane rupture (induction or augmentation of labor)

27% of all women reported receiving membrane rupture (either for induction or augmentation of labor), compared to 57% of observers reporting this intervention. Given the relatively low sensitivity of the indicator (38%), the indicator meet only the AUC validity criteria (AUC: 0.6333± 0.05, IF: 0.47). The low sensitivity of the indicator indicates that the majority of women who received membrane rupture did not report it.

INDUCTION & AUGMENTATION OF LABOR: SUMMARY

Results suggest women may be able to report on whether an injection or IV medication was received at some point during labor before delivery, but not the indication for the medication (for induction or augmentation of labor). Women did not report accurately on whether membrane rupture was received for the induction or augmentation of labor at the population level.

Uterotonic for Prevention of Postpartum Hemorrhage (PPH) (Post-delivery)

Box 5A. Cross-tabulation: Received prophylactic uterotonic (Y/N).

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? *(Select all)*

Self-report (Number)	Observer Report (Number) *Uterotonic administered (Y/N)*		
	No	Yes	**Total**
No	0	232	232
Yes	4	357	361
Total	4	589	593

Active management of the third stage of labor (AMTSL), which includes administration of a uterotonic following delivery, is a critical intervention to prevent postpartum hemorrhage, a leading contributor to maternal mortality and morbidity.[22] Nearly all women received a uterotonic following delivery (99%), and nearly two-thirds (64%) received the medication within 3 minutes of birth. All uterotonics were administered via IV line, most often by IV drip (93%), a bolus injected into an IV line (6%), or an IV drip combined with intramuscular injection (<1%).

To assess women's ability to accurately report whether a uterotonic for the prevention of postpartum hemorrhage was administered following delivery, women were asked, "*In the first few minutes after the delivery of your baby, did anyone give you an injection in your thigh or buttock, medication intravenously (through a tube in your arm) or tablets to swallow, hold in your mouth, or placed tablets in your rectum?" (option to select all applicable interventions received).*

It was not possible to conduct robust analysis of this indicator, given that nearly all women received the intervention. Cross-tabulation results (**Box 5A**) indicate that of women who received the intervention, less than two-thirds reported receiving it (sensitivity: 61%). These results suggest that the composite proxy indicator constructed to assess whether women were administered a uterotonic for PPH did not capture a substantial proportion of women who received the intervention.

Timing of uterotonic administration for PPH

To examine whether women could accurately report on the timing in which the uterotonic was received, we assessed women's self-report of whether any IV medication, injection in the thigh or buttocks, or tablets (oral, vaginal or rectal) were received immediately (in the first few minutes) following birth. This composite proxy indicator was compared to the observer report of whether the uterotonic was administered up to 3 minutes following birth. This indicator met only the IF study validation criteria (AUC: 0.5106± 0.04, IF: 0.95). Specifically, of women who were observed to receive the prophylactic uterotonic within this time frame, 61% reported it. The moderate sensitivity level indicates that a considerable proportion of women did not report the intervention when it in fact took place. The indicator was also marked by a high degree of false positive reporting (i.e., low specificity). Over half of women (59%) who did not receive the intervention within 3 minutes of delivery falsely reported that it took place (**Box 5B**). These results suggest that at an individual level, this indicator is unlikely to be validly reported, as both underreporting of true cases and false positive reporting among negative cases occurs. However, at a population-level, IF results indicate the low sensitivity and specificity cancel out and the indicator produces an acceptable estimate of the population-based prevalence.

Box 5B. Cross-tabulation: timing of prophylactic uterotonic (1-3 minutes post-delivery).

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?

Self-report (Number)	Observer Report (Number) *Uterotonic administered 1-3 minutes following birth*		
	No	Yes	**Total**
No	85	143	228
Yes	124	228	352
Total	209	371	580

We also examined the accuracy of women's self-reports of whether the prophylactic uterotonic was received following the delivery of the placenta. For this comparison, we assessed women's report on whether any IV medication, injection in her thigh or buttocks, or tablets (oral, vaginal or rectal) were given in the first few minutes after the delivery of the placenta. We compared this proxy indicator to the observer report on whether the uterotonic was received up to 3 minutes following the delivery of the placenta. A high proportion of women (94%) reported receiving IV medication, an intramuscular injection, or tablets following the delivery of the placenta. In contrast, only one-fifth of women (20%) were observed to receive a prophylactic uterotonic following the delivery of the placenta. The proxy for women's self-report on receiving prophylactic uterotonic following the delivery of the placenta had high sensitivity (96%), demonstrating nearly all women who received the intervention were able to report it. However, the indicator was marked by a high false positive rate (94%) and did not meet either study validation criteria criteria (AUC: 0.5097± 0.04, IF: 4.71).

Method of uterotonic administration

All observers reported that the prophylactic uterotonic was administered by IV line (either by IV bolus, IV drip, or IV drip plus IM injection), compared with 61% of women. 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 uterotonic by IV line (any IV method) (n=498), less than two-thirds report having received medication by this administration route (specificity: 61%).

Type of uterotonic

The most common type of uterotonic administered was oxytocin (97%), followed by ergonovine and carbetocin. Since oxytocin was the standard medication administered, robust analysis was not possible. However, cross-tabulation results demonstrate that of women who received oxytocin, 50% of women accurately reported receiving the medication by its brand name (**Box 5C**).

Box 5C. Cross-tabulation: Oxytocin received after birth.

Immediately after the birth of your baby, did anyone give you medication called oxytocin to help your womb become firm?

		Observer report (Number)		
		No	Yes	**Total**
Self-report (Number)	No	1	178	179
	Yes	2	177	179
	Total	3	355	358

Since oxytocin was the uterotonic administered to nearly all women and its route of administration was by IV line (97%), we compared women's self-report of receiving a uterotonic by administration route (assessed by a proxy composite measure of 'yes' to whether the woman received any of the following: (1) an injection in the thigh or buttock, (2) medication intravenously through a tube in arm, (3) or tablets given orally or rectally in the first few minutes following birth), and their reporting 'yes' to the question "*after the birth of your baby, did anyone give you a medication called oxytocin to help your womb become firm or contract?*".

Box 5D. Cross-tabulation (Self-Report): Prophylactic uterotonic and oxytocin.

Q1. 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?

Q2. Immediately after the birth of your baby or sometime before the delivery of the placenta, did anyone give you a medication called oxytocin? (*constructed in analysis*)

Self-Report (Number)

Self-report (Number) Q1.	Q2. No oxytocin	Q2. Yes oxytocin	Total
Injection	9	21	30
IV line	52	125	177
Tablet	0	1	1
Total	61	147	208

The cross-tabulation results of women's self-report on these two indicators are presented in **Box 5D**. While only descriptive analysis was possible, these results demonstrate that 85% of women (125 out of 147) who reported receiving medication by IV line following delivery also reported receiving oxytocin. Of note is that 85% (52 out of 61) of women who reported receiving IV medication following delivery did not report receiving oxytocin. To determine whether these women had in fact received oxytocin (highlighted cell, **Box 5D**), we conducted a cross-tabulation with the observer report. Data show that all but one of these women (99%) received oxytocin by IV line. Therefore, 51 of the 52 women were correct in noting they had received an IV line, but incorrect in not indicating that they had received oxytocin. These results suggest not all women can report on receiving a uterotonic by name.

UTEROTONIC FOR THE PREVENTION OF POSTPARTUM HEMORRHAGE: SUMMARY

Nearly all women received oxytocin via IV line following delivery, meaning there was insufficient variation to assess whether a woman can report receiving a uterotonic for the prevention of postpartum hemorrhage, the route of administration, or the name of the drug. Cross-tabulation results, however, suggest that a large proportion of women who received a uterotonic would not be captured using a composite study indicator. Descriptive results also suggest that women are unlikely to be able to report on receiving oxytocin by name. The timing in which the uterotonic was received (in the first few minutes of birth or after the delivery of placenta) was robustly analyzed and did not meet criteria for valid

reporting at the individual level. IF results suggest the indicator on whether a uterotonic was received within 3 minutes after delivery could approximate the population-based prevalence.

Active management of the third stage of labor

Active management of the third stage of labor for the prevention of postpartum hemorrhage includes 3 components: (1) administration of a uterotonic for PPH, (2) controlled cord traction, and (3) uterine massage immediately following delivery of the placenta. Robust analysis was possible only for whether uterine massage was performed immediately following delivery of the placenta. This indicator had moderately high sensitivity (80%) and low specificity (20%), indicating that approximately 20% of women who received uterine massage following delivery of the placenta were not captured by the indicator and 80% of women who did not receive the intervention reported that it had taken place. Given that nearly all women received and reported receiving prophylactic uterotonic and controlled cord traction, it was not possible to assess these indicators using robust analysis. However, cross-tabulation results indicate that a substantial proportion (38%) of women did not report receiving a uterotonic following delivery, although they were observed to do so. Controlled cord traction was also a near universal practice, which limited the ability to conduct robust analysis. Most women who received controlled cord traction reported it (sensitivity: 94%), but of the few cases where controlled cord traction was not implemented, there was a high degree of false reporting (95% false positive reports).

A composite indicator of the 3 essential components of the active management of the third stage of labor (AMTSL) was assessed. This indicator did not meet the AUC study validity criteria (AUC: 0.4986± 0.04, IF: 0.93) Cross-tabulation results in **Box 6** indicate that approximately one-quarter of women who received all 3 components of care were not captured by the indicator (sensitivity: 75%). Additionally, about three-quarters of women who did not receive all 3 elements of care reported that they had (1- specificity: 76%). The AMTSL composite indicator, however, did meet IF criteria, suggesting it may be suitable for assessing population-based coverage of the practice.

Box 6. Cross-tabulation: All 3 components of AMTSL received.

Prophylactic uterotonic + controlled cord traction + uterine massage following delivery of placenta received.

		Observer Report (Number)		
		No	Yes	**Total**
Self-report (Number)	No	24	109	133
	Yes	76	340	416
	Total	100	449	549

ACTIVE MANAGEMENT OF THE THIRD STAGE OF LABOR: SUMMARY

A composite indicator on whether 3 essential elements for active management of the third stage of labor—uterotonic for the prevention of postpartum hemorrhage, controlled cord traction, and uterine massage following delivery of the placenta—did not meet study validity criteria. The indicator was both under- and overreported by women. The composite AMTSL indicator did meet the IF criteria, indicating a close approximation of the population-based prevalence of the practice.

Skilled Birth Attendance

Skilled provider during labor and delivery

Both observers' and women's self-reported coverage of skilled birth attendance [defined as a general "doctor" category, doctor subcategories (identified as an ob-gyn specialist, a medical resident or general practitioner), or a nurse] was high for the primary birth attendant during labor and delivery and for the provider who delivered or 'caught' the newborn during birth (>90%). Given the near universal occurrence of this practice, it was not possible to conduct robust analysis. Cross-tabulation data, however, indicate that nearly all women observed to have a skilled provider 'catch' their newborn at birth correctly reported it as such (specificity: 97%). A similarly high proportion of women observed to have a skilled birth attendant as their primary provider in labor (97%) or delivery (93%) also identified the level of their provider as such. However, the vast majority of women not observed to receive skilled birth attendance at each of these stages falsely reported receiving skilled attendance, indicating low indicator specificity (>90% false positive reporting across birth phases). These results should be interpreted with caution given the small number of women who did not receive skilled birth attendance (n=19 in labor; 42 in delivery), since the lack of variation in the data prevented full analysis.

Main provider who assisted in delivery

Most deliveries were attended by medical residents (90%) as the primary provider, followed by medical interns (7%), general practitioners or ob-gyns (2%), and nurses (<1%) (observed prevalence, excluding 'Don't Know' responses). The vast majority of women reported a doctor as the primary provider during their delivery (93%), without being able to define the category to which the medical provider belonged to (i.e., ob-gyn, medical resident, and general practitioner or medical intern). Specifically, of the 542 women who reported a doctor as their main provider during delivery 40 (7%) providers were actually medical interns (**Table 2**).

TABLE 2. Cross-tabulation: Main provider during delivery.

Self- Report (Number)	**Observer Report** (Number)						
	Doctor /Ob-gyn	Medical resident	Medical Intern	Nurse	Student Nurse	Other	**Total**
Doctor /Ob-gyn	14	487	40	1	0	0	542
Medical resident	0	17	0	0	0	0	17
Medical intern	0	11	2	0	0	0	13
Nurse	0	11	0	1	0	0	12
Student nurse	0	0	0	0	0	0	0
Don't Know	0	9	2	0	0	0	11
Total	14	535	44	2	0	0	595

Box 7. Cross-tabulation. Main provider during delivery—doctor /ob-gyn or medical resident.

Self-report (Number)	Observer report (Number)		
	No	Yes	**Total**
No	3	22	25
Yes	41	510	511
Total	44	532	576

In Mexico, there is little distinction between doctor categories; however, while medical residents are considered skilled attendants, medical interns are not. We assessed a combined indicator of whether the main provider during delivery belonged to any doctor subcategory of skilled provider (i.e., ob-gyn, medical resident, or general practitioner) (**Box 7**). While robust analysis was not possible, cross-tabulation results show the indicator would have high sensitivity (96%). Cross-tabulation results in **Table 2** suggest women may be unable to distinguish the presence of less prevalent, unskilled providers, such as medical interns. The discrepancy in reporting may reflect differences in how the 'main' provider was conceptualized by women in comparison to observers. For example, it is possible that women identified the 'main' provider as the highest-ranking attendant present during delivery. Observers, instead, identified the 'main' provider as the provider who spent the most time, or administered the majority of care received by the woman. Additionally, given the low prevalence of instances when the main provider was not a doctor/ob-gyn or medical resident, without near-perfect specificity these indicators are likely to be overreported.

Main provider who 'caught' the newborn

In addition to identifying their main provider during delivery ("*Who was the main provider assisting you during delivery?*"), women were asked, "*Who was the main provider who actually delivered your baby (caught the baby)?*" The majority of providers who 'caught the newborn' following delivery were either doctors/ob-gyns or medical residents (90% self-report compared to 91% observed prevalence). Of the

485 women who reported that the provider who caught their newborn was a doctor, 9% were either medical interns or nurses.

It was possible to validate a combined indicator on the category of birth attendant who 'caught' the newborn during delivery as a doctor/ob-gyn or medical resident. This indicator did not meet the AUC criteria (AUC: 0.5203± 0.04). Specifically, at an individual classification level, the combined doctor/ob-gyn/medical resident indicator was marked by high sensitivity (90%) and low specificity (14%). These results indicate women had low accuracy in reporting when the main provider did *not* belong to one of the skilled provider doctor subcategories (doctor or medical resident), and instead was some other less common type of provider (e.g., nurse, student nurse, or medical intern). However, at the aggregate level for this indicator, differences in sensitivity and specificity canceled out to produce an acceptable IF level (IF: 0.98). The IF result of close to 1 highlights the potential of the indicator to obtain an acceptable population-based coverage estimate of the provider category.

Main provider labor

Similar to delivery, the most prevalent providers during labor were medical residents or doctors/ob-gyn (82% self-report and 96% observed prevalence). Of the 501 women who reported a doctor was the main provider during labor, 12% were either medical interns, nurses, or student nurses (**Table 3).** The combined indicator on whether the provider was a doctor/ob-gyn or medical resident during labor was robustly analyzed. The indicator did not meet study AUC validity criteria (AUC: 0.5477± 0.04), with high sensitivity (82%) and notably low specificity (27%). The indicator did meet IF criteria, suggesting the potential for this indicator to obtain acceptable population-based estimates of coverage of a doctor/ob-gyn during labor (IF: 0.85) at an aggregate level.

TABLE 3. Cross-tabulation: Main provider during labor.

Self-report (Number)	**Observer Report** (Number)						
	Doctor/ ob-gyn	Medical resident	Medical Intern	Nurse	Student Nurse	Other	**Total**
Doctor /ob-gyn	7	423	12	56	3	0	501
Medical resident	3	26	1	0	0	0	30
Medical intern	0	19	2	0	0	0	21
Nurse	0	76	4	0	0	0	80
Student nurse	0	4	0	0	0	0	4
Don't Know	0	10	0	0	0	0	10
Total	10	558	19	56	3	0	646

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 was a doctor (80% self-report), followed by a nurse (81%), medical student or intern (26%), and medical resident (14%). Women also specified an 'other' option, and listed specific types of doctors such as a pediatrician (39%) or anesthesiologist (10%) as also present. Because women and observers were asked to report all types of providers present, rather than specifying one main provider, there were sufficient numbers to assess the accuracy of women's self-report of multiple categories of providers.

Results of the validity analysis reflect the same trend as the primary provider findings, in that the majority of additional providers noted to be present during labor were doctors/ob-gyns or medical residents (82% self-report, 88% observed prevalence). Although the combined doctor/ob-gyn /medical resident indicator had a similar prevalence rate reported by women and observers, it was marked by low specificity (21%) and met only the IF validity criteria (IF: 0.93). Other indicators of additional providers during delivery or labor that met the IF validity criteria alone were whether additional providers were a nurse or nursing student.

Additional providers—skilled

The near universal presence of a skilled provider (doctor of any type, or nurse) as an additional provider in both labor and delivery limited our ability to conduct robust analysis. It is noteworthy; however, that almost all of the handful of women who were not observed to have a skilled provider as an additional provider present during labor or delivery falsely reported the presence of one. This finding should be interpreted with caution since there were few cases in which an additional skilled provider was not present during labor and delivery.

Skilled birth attendant: summary

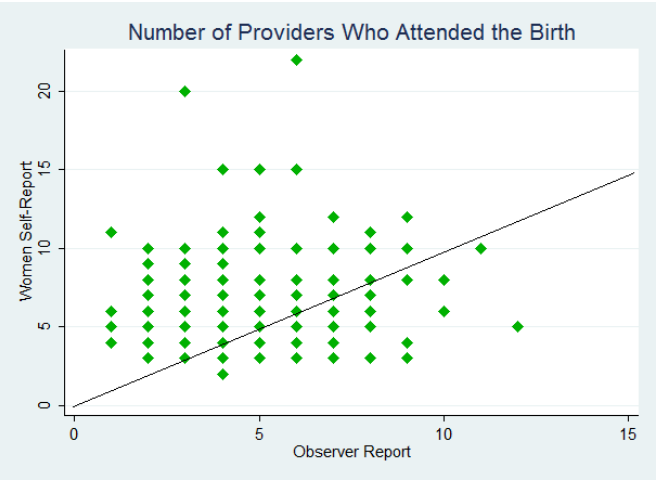
In summary, across birth phases and provider roles, the category of provider most commonly observed to be present during labor and delivery were skilled provider doctors (i.e., ob-gyns, general practitioners, or medical residents) either as the primary provider or as an additional provider present. Since there is little distinction in Mexico between these doctor subcategories, including medical residents, a combined

indicator was assessed. No composite doctor/ob-gyn /medical resident indicator met both validity criteria. Across birth phases the combined doctor indicator had low specificity, indicating that women were likely to report any doctor category as the main provider or as an additional provider who was present when they were not. Results from 'other providers' present during labor or delivery, for which robust analysis was possible, also reflected these findings. Given the low prevalence of other types of providers present during labor and delivery, without near-perfect specificity these indicators are likely to be overestimated by women at the individual level. However, across birth phases and provider roles, 8 birth attendant indicators met the IF criteria alone. They included whether the main provider who caught the newborn during delivery was a doctor/ob-gyn or medical resident. Birth attendant indicators that met the IF criteria alone may be useful for population-based coverage of these provider types. Given that the fine distinction between subcategories of doctors (i.e., ob-gyn, general practitioner, or medical resident) is likely less programmatically useful than assessing the prevalence of skilled attendance, the composite doctor indicator may be of particular use.

Number of providers who attended birth

The number of providers who attended the woman's birth was observed to range from 1 to 12 (mean: 4 ± 2). Approximately half of women were attended by more than 4 providers (45% true prevalence) during their delivery, while 75% of women were attended by six providers or fewer. The average number of providers who attended the birth by women's self-report was higher (mean: 6 ± 3) and ranged from 2 to 22, reflecting the trend for women to overreport the number of providers present (**Figure 2**).

FIGURE 2. Scatterplot: Women's and observers' report on the number of providers who attended the birth.



Women's self-reports and observers' reports on the number of providers present were weakly correlated (r=0.2, p<0.001). No indicators of the number of providers who attended the birth met both study validity criteria. In general, women's self-report of the number of providers present at their birth were marked by high rates of false positive reporting (i.e., low specificity) (**Annex B**).

SKILLED BIRTH ATTENDANT: SUMMARY

In contrast to other settings, in Mexico doctors and medical residents attend the vast majority of births in hospital settings, and both are considered skilled providers. An indicator that combined these skilled doctor sub-categories did not meet both validity criteria, irrespective of delivery phase and provider role

(main provider or additional provider present). Across birth phases the combined doctor indicator had low specificity, indicating that women were likely to report any doctor category as the main provider or as an additional provider who was present when they were not. Given the low prevalence of other types of providers, without near-perfect specificity (i.e., true negative classification) these indicators are likely to be overestimated by women at the individual level. However, across birth phases and provider roles, 8 birth attendant indicators did meet the IF validation criteria, including whether the main provider who caught the newborn during delivery was a doctor/ob-gyn/medical resident. These indicators may be useful for estimating population-based coverage of these provider types.

Immediate Newborn Care

Three indicators of immediate newborn care had sufficient variation for validity analysis: (1) whether the baby was given to the mother immediately after birth, (2) whether the newborn was breastfed within the first hour after birth, and (3) whether something other than breastmilk was given to the baby to eat or drink in the first hour after birth.

None of these indicators met both study validity criteria. For all 3 indicators, women's self-reported prevalence exceeded the observed prevalence. While the indicator of breastfeeding in the first hour after birth had fairly high sensitivity (81%), it was also marked by only moderate specificity (44%), indicating that nearly 60% of women not observed to breastfeed in the first hour after birth falsely reported that they had. The indicators of whether the baby was immediately placed with the mother after birth and whether the baby was given something other than breastmilk to eat or drink in the first hour after birth were marked by both moderate sensitivity and specificity (**Annex B**), indicating both the underreporting of true cases and false reporting of true negative cases**.** The reporting discrepancies may reflect differences in how women understood the terms 'immediately' or the 'first hour' after delivery.

Newborn thermal care

A high proportion of newborns were immediately dried following delivery (94% observed prevalence). Since nearly all mothers also reported that their newborns were immediately dried (99%), robust analysis was not possible.

Box 8. Cross-tabulation (Self-Report): Newborn skin-to- skin indicators.

Q1. Did someone place the baby on your chest, against your skin immediately after delivery?

Q2. Was your baby wrapped in a towel while lying against your chest or naked against your skin?

Q1. Self-Report (Number)	Q2. Self-Report (Number)		
	Wrapped in towel	Naked on Skin	**Total**
Not skin-to-skin	0	0	0
Yes, skin-to-skin	54	8	62
Total	54	8	62

Although validity analysis was also not possible for other indicators of newborn thermal care, results suggest indicator wording has important implications for women's reporting accuracy. Specifically, almost no observers (<1%) reported the newborn was placed skin-to-skin against the mother's chest following delivery. However, 11% of women reported that this practice took place, reflecting a substantial degree of false reporting (**Box 8**). A two-item indicator which included women who positively answered "*Did someone place the baby on your chest, against your skin immediately after delivery*" and subsequently reported the baby was lying naked against the skin to the item "*Was your baby wrapped in a towel while lying against your chest or lying naked against your skin,*" reduced false positive reporting. Women's reported prevalence to the two-item skin-to-skin indicator (1%) better approximated the observer report (<1%).

Instead of being placed naked skin to skin with the mother and then draped with a cloth, the standard practice at the Mexico facility was for babies who were breathing at birth to be first wrapped with a towel or cloth (99%). Babies were then brought near the mother for her to kiss on the cheek. This contact with the baby may have confused women when asked about skin-to-skin contact. In addition, because the practice of wrapping the newborn after birth was near universal, it was not possible to assess the accuracy of women's reporting on this indicator. However, cross-tabulation results show that almost all of the women whose baby was not placed naked against their skin and was instead wrapped in a towel accurately reported these circumstances (99%). These results should be interpreted with caution since full validity analysis was not possible, and all babies not placed directly on the mother's skin were covered with a towel (in matched data).

Because the practice of bathing the baby within the first hour after birth rarely occurred, we were unable to fully assess this practice (3% observed prevalence; <1% self-report prevalence). However, descriptive

cross-tabulation results indicate high specificity for this indicator, that is, of the 429 newborns not bathed within the first hour, 426 women correctly reported their newborn was not bathed (99% specificity).

Composite indicator of essential newborn care

A composite indicator of 3 essential elements of newborn care (newborn was immediately dried, placed skin-to-skin on mother's chest, and breastfed in first hour) was constructed in analysis. This indicator had a 10% self-reported prevalence by women and 0% observed prevalence. The zero observed prevalence of the composite measure is due to the fact that only one newborn was observed to be placed naked against the mother's skin immediately following delivery. Using the two-item skin-to-skin item in this composite indicator greatly reduced the reporting of false positives (<1% self-reported prevalence).

IMMEDIATE POSTPARTUM NEWBORN CARE: SUMMARY

We were able to assess 3 indicators of immediate newborn practices: whether the baby was placed with the mother immediately following birth, whether breastfeeding was initiated within the first hour of birth, and whether something other than breastmilk was given to the baby to eat or drink within the first hour of birth. No indicators met both study validity criteria. Question wording and order may have significant implications for the validity of indicators of newborn thermal 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 (specify), or (8) no complications. Over one-third (38%) of women reported experiencing some type of complication, exceeding the observed prevalence (14%) by nearly 3 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 61%, indicating that nearly 40% of women who had in fact experienced a complication did not report it. The indicator also had moderately low specificity (65%), reflecting a high

rate of false positive reports by women. Specifically, over one-third of women (35%) who had not experienced a complication falsely reported experiencing one.

Box 9A. Descriptive frequencies: Complications.

Did you experience any of the following complications during or after your delivery? (List all)	Number	Percent	Record whether mother had any of the following complications. (Select all)	Number	Percent
High blood pressure, seizure, blurred vision, severe headache, swelling in hands/face	150	25.05	Pre-eclampsia/ eclampsia	14	2.32
Excessive bleeding	25	4.17	Hemorrhage	50	8.31
Baby was in distress/ too large	44	7.35	Obstructed labor	5	0.83
Long labor (more than 12 hrs)	51	8.51	Prolonged labor	13	2.13
Infection (fever)	33	5.51	Sepsis	0	0
Other	45	7.51	Other	6	0.99
No complications	371	61.94	No complications	520	86.24
Don't Know	0	0	Don't Know	1	1.2

Women most commonly reported experiencing eclampsia (i.e., high blood pressure, seizures, blurred vision, severe headache and swelling in the hands or face) (25% self-reported prevalence), followed by long labor (more than 12 hours) (9%), the baby was in distress or too large (7%) (indicative of fetal distress or obstructed labor), infection (fever) (6%), and excessive bleeding (4%) (indicative of hemorrhage) (**Box 9A**). An additional 8% of women reported some other type of complication, including indications of gestational diabetes (high sugar), preterm or overdue newborn, nausea or vomiting, or not becoming dilated enough to progress into labor.

There was sufficient variation to assess the accuracy of women's report of two complications: pre-eclampsia/ eclampsia and hemorrhage (APH + PPH). The second indicator met both study validity criteria (AUC: 0.7261 ± 0.04, IF: 1.04). It is noteworthy, however, that while this indicator had high specificity (95%), its sensitivity was moderate at 50%. These results suggest that while the hemorrhage indicator would have a very low false positive rate, it would also miss half of women who in fact experienced hemorrhage (**Box 9B**). However, the indicator did meet IF criteria, indicating that while it may not be suitable for the individual classification of women, it would generate an acceptable approximation of the population-based prevalence of the complication.

Box 9B. Cross-tabulation: Complication of hemorrhage.

Did you experience the following complication during or after your delivery? *Excessive bleeding.*

		Observer Report (Number)		
		No	Yes	**Total**
Self-Report (Number)	No	519	24	543
	Yes	26	24	50
	Total	545	48	593

The indicator of eclampsia did not meet study validity criteria and was in general overreported by women. Specifically, if assessed by a population-based survey, the indicator would exceed the observed prevalence of pre-eclampsia/eclampsia by over 9 times (IF: 9.1).

Post-delivery health checks

All women (100%) 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 the womb was becoming firm), (4) took blood pressure, or (5) took temperature (**Box 9C**). Observers indicated that 99% 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 a specific post-delivery check, including: 1) bleeding, 2) examine the perineum, 3) temperature, 4) blood pressure, and 5) involution.

No indicators of specific post-delivery health checks on the mother met both study validity criteria (**Annex B**). While each indicator had fairly high sensitivity (>90%) apart from a slightly lower sensitivity for the provider check on involution, each indicator was also characterized by low specificity (i.e., high rates of false positive reporting).

Box 9C. Descriptive frequencies: Post-delivery health checks.

In the first physical examination after delivery did the provider look, ask about, 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	100.00	99.30
Check for bleeding	87.46	80.51
Did a perineal exam	90.10	75.09
Checked to see whether womb was becoming firm (involution)	83.45	81.65
Took blood pressure	93.38	95.59
Took temperature	93.34	66.89

These findings reflect a tendency for most women to report that the specific health checks took place, while observers did not report that the examinations occurred in all instances. Indicators of whether the provider took the woman's blood pressure in the first physical examination following delivery and whether the provider took the woman's temperature had notably low specificity levels (<10%), indicating that the vast majority of women who did not receive this care reported that they had. The low prevalence of instances when post-delivery health checks did not take place also indicates that without near-perfect negative classification by women, these indicators are likely to be overestimated. Another potential explanation for the discrepancy between observers' and women's report of postnatal maternal health checks may be due to differences in timing. For example, data collectors anecdotally noted that the first physical exam after birth would sometimes be delayed up to 2-3 hours following

delivery, after study observation had concluded, or could have been conducted outside of the delivery ward, in which case they would not have been seen by observers.

All indicators of immediate postnatal health checks on the mother (apart from whether the provider took the woman's temperature in the first physical examination following delivery) met the IF criteria, suggesting that coverage of these indicators would approximate the population-based prevalence.

Other maternal outcomes

Other maternal outcomes assessed included whether the mother received an episiotomy, was given any blood products during her stay in the facility, and whether she asked for or received pain medication at any time. Each of these 4 indicators was assessed by robust analysis. Indicators of episiotomy and blood products given met both study validity criteria.

Women reported receiving episiotomy with generally high accuracy (sensitivity: 98%, specificity: 63%, AUC: 0.8041± 0.03, IF: 1.15), and this indicator met both validation criteria. Although receiving blood products during the stay in the facility also met study validity criteria (AUC: 0.6706± 0.04, IF: 1.00), this indicator was marked by low sensitivity (36%), indicating that approximately 65% of women who had received blood products at some time during their stay did not report it. A woman requiring a blood transfusion is likely to be in critical health, possibly asleep or unconscious, which might cause women who received the intervention not to recall it.

Neither asking for nor receiving pain relief medication during the facility stay met both study validity criteria. In particular, asking for pain relief medication was underreported by women (sensitivity: 33%, specificity: 79%), while receiving pain medication tended to be overreported by women (sensitivity: 90%, specificity: 20%), with a higher false positive rate. Whether the woman received pain relief medication meet only the IF criteria.

MATERNAL OUTCOMES: SUMMARY

Women's report of whether any complication occurred did not meet both study criteria and would be overestimated nearly 3 times if assessed through a population-based survey. Indicators for two specific complications had sufficient variation for robust analysis: whether hemorrhage or eclampsia was experienced. The indicator of hemorrhage (whether excessive bleeding was experienced) met both study validity criteria, but was marked by low sensitivity, suggesting nearly half of hemorrhage cases would not be reported by women. However, the IF factor of approximately 1 indicates the indicator would produce an acceptable population-based estimate of hemorrhage. The indicator of eclampsia did

not meet the IF criteria and would likely be overestimated in a population-based survey. No indicator of postpartum health checks met both validity criteria. In general, women tended to report that health checks took place in all circumstances (regardless of whether they were observed to occur). The low prevalence of instances when health checks did not take place increases the potential for overreporting by women.

Other maternal outcomes examined included receiving episiotomy, receiving blood products at some time during their facility stay, and asking for and receiving pain medication. Two indicators, receiving episiotomy and blood products, met both validity criteria. Receiving blood products, however, was notably underreported by women, and approximately 65% of women who had received blood products did not report receiving the intervention.

Newborn Outcomes

Newborn birthweight

Observers' and women's self-reports of newborn gram weight were very highly correlated (r= 0.97). The high degree of correlation may partially reflect the standard facility practice of providing each mother with a bracelet that listed her newborn's birthweight, which may have increased the salience of this indicator. To account for this, validity analysis was restricted to women who reported their newborn's weight themselves, rather than taking the weight from a card (which occurred in 3 cases). We classified newborn birthweight into 3 categories: low birthweight (<2,500 grams), normal birthweight (2,500-4,499 grams), and high birthweight >4,500 grams). Most newborns were normal weight (88%), 12% were low birthweight, and none were high birthweight.

Women's self-report of low and normal newborn birthweight was near perfect. Specifically, all women who delivered a low birthweight newborn accurately reported their newborn's weight to be less than 2,500 grams (100% sensitivity); the indicator also had 99% specificity. Similarly, normal newborn birthweight was reported with 99% sensitivity and 100% specificity. Given the lack of variation in the data, we were unable to analyze these indicators fully. However, women's near-perfect reporting accuracy indicates that this information was highly memorable. While the high accuracy may be in part due to the practice of providing women with a bracelet that listed their newborn's weight, study interviewers also agreed that information related to the baby's weight seemed to be important to women and was easily recalled.

Stillbirth delivery

Given the small number of stillbirth deliveries (n=3 observed, matched data n=2 due to missing data in client exit interview), we were unable to analyze the validity of an indicator of stillbirth deliveries. While results should be interpreted with caution, descriptive cross-tabulations (**Box 10**) indicate that both stillbirth deliveries were correctly classified by women.

Box 10. Cross-tabulation: Stillbirth delivery.

		Observer Report (Number)		
		Alive	Stillbirth	**Total**
Self-Report (Number)	Alive	589	0	589
	Stillbirth	0	2	2
	Total	589	2	591

NEWBORN OUTCOMES: SUMMARY

Newborn birthweight was nearly perfectly reported by women, and likely reflected the fact that all women were given a bracelet with the newborn's birthweight listed. Very few newborns were stillborn, but the 2 stillborn cases that occurred were accurately reported.

Discussion

Overview:

48 out of 108 possible indicators had sufficient variation in numbers for validity analysis. Of these, 5 indicators met both acceptability criteria (AUC>0.6 and 0.75<IF<1.25) (**Table 4**). 28 of these indicators met only one of the acceptability criteria. Few indicators met both study validity criteria, in part because many quality of care interventions were either routinely practiced or rarely occurred (see **Annex A, Table 2**). As a result, there was often insufficient variation in women's and observers' reports for robust validity analysis.

Validated Indicators that Met Study Validity Criteria

TABLE 4. Indicators that met both validation criteria.

Provider takes a urine sample at or near facility admission
Injection or IV medication received at some time during labor, before the birth of the baby
Episiotomy was performed
Complication of hemorrhage was experienced
Blood products were given

Study findings suggest women's report of indicators that 'became significant for them' were most likely to be validly reported. Salient events included aspects of care that had caused pain, such as receiving an episiotomy, or were cause for concern, such as experiencing hemorrhage or receiving a blood transfusion. It is noteworthy, however, that although indicators on urine sample screening, experiencing hemorrhage, and receiving blood products met study validity criteria, each of these indicators had low sensitivity (50% for urine sample and hemorrhage; 36% for blood products). These findings indicate that of women who experienced these events, half or more did not report it. Because postpartum hemorrhage is a leading cause of maternal death, and blood transfusion and a urine screen for pre-eclampsia/eclampsia are potentially lifesaving interventions, the low sensitivity of these indicators would not capture critical information regarding the women who experienced these events. If the goal of using these indicators is not to obtain an accurate record of true positive and true negative cases, but rather to approximate the prevalence of the indicators in the population, the IF of nearly 1 suggests that the indicators may be suitable to estimate coverage at an aggregate level.

Receiving an injection or IV medication at some time during labor, before the birth of the baby (proxy for uterotonic for labor induction or augmentation), also met both validity criteria. The high sensitivity (70%) of this indicator may reflect the observation that upon admission to the labor /delivery ward it was universal practice for women to have an IV line set, which also likely decreased the use of alternative drug administration routes. Taken together, these results suggest women may be able to report whether

an injection or IV medication was received, but not on the indication for the intervention (whether the medication was received for induction or augmentation of labor).

Indicators on Skilled Birth Attendance

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 labor and delivery. Skilled birth attendance (i.e., doctor/ob-gyn, medical resident, or nurse) was common, with greater than 90% observed prevalence regardless of delivery phase or provider role (main provider or additional provider present). Cross-tabulation results (additional provider indicator) reflect high sensitivity and low specificity of women's self-reports on skilled birth attendance. However, given there were few instances when a skilled provider was not present, these results demonstrate that unless women had near perfect specificity in reporting, this indicator is likely to be overestimated.

Of specific relevance to this context was to test women's accuracy in identifying when the main provider belonged to a skilled category. The skilled provider most commonly observed was a medical resident (>92% across delivery phases). Of note is that in Mexico, there is little distinction between the skill level of a medical resident and other doctor categories (general practitioners, ob-gyn specialists) since they would all be classified as skilled providers. We therefore assessed the validity of a composite doctor/ob-gyn/medical resident indicator. This composite indicator did not meet both validity criteria, irrespective of delivery phase or provider role (main provider or additional provider present). Similar to 'skilled birth attendance', the doctor/ob-gyn/medical resident indicator was marked by high sensitivity and low specificity. As with the 'skilled provider' indicator, our results are limited by the lack of variation in provider type, since the tendency to overreport low-prevalence scenarios was common when specificity levels were not high.

The discrepancy in reporting on 'skilled birth attendance' and doctor/medical resident indicators may also reflect differences in how women and observers identified the provider with the 'main' role. Observers report identifying the primary provider as the staff member who spent the most time or conducted the majority of actions during labor or delivery. It is possible that women instead conceptualized the 'main' provider as the birth attendant who had the highest rank and was deemed to be 'in charge' of care. The high number of attendants at each birth (mean: 4 ± 2) and the setting of a teaching hospital setting may also have contributed to women's difficulty in identifying the 'main' provider. Future qualitative analysis is recommended to explore how women understand key terms in birth attendant indicators and women's understanding of the existence and differences between the doctor cadres.

While the low prevalence of instances when a skilled doctor was not present limited our ability to analyze whether women could distinguish among finer categories of providers, it is important to note that 8 birth attendant indicators met the IF criteria alone. This included whether the main provider who caught the newborn during delivery was a doctor/ob-gyn or medical resident. Birth attendant indicators that met the IF criteria alone may be useful for assessing the population-based coverage of these provider types. Given that broad categorizations of provider levels (i.e., whether the provider is skilled or some type of doctor) are likely to be more programmatically meaningful than the ability of women to delineate between specific provider types, the composite doctor indicator may be of particular use at the population level.

Validated Indicators that Met At Least One Study Validity Criteria

In addition to skilled birth attendance, other core indicators have high potential for practical use. We define 'core' indicators as those included in DHS and MICS surveys, or that reflect essential aspects of obstetric and immediate newborn care. Core assessed indicators that met at least one validity criteria were: HIV status checked, receiving a uterotonic up to 3 minutes following the birth and whether 3 essential components of AMTSL were received. Each of these indicators met the IF validation criteria alone, suggesting they may produce an acceptable estimate of indicator prevalence at the population level. Because these indicators did not meet the AUC criteria, they are not recommended for classifying which individual women received a specific intervention. However, at the aggregate level, the tendency for underreporting of true cases and overreporting of false cases cancel out to produce an acceptable population-based estimate.

Indicators Reported with Difficulty

Other indicators were reported with difficulty and had notably high 'Don't Know' responses by women. These include measures of specific medications or injections received, HIV testing, provider hygiene (hand sanitization or sterile glove use), aspects of immediate postpartum care (whether baby was immediately dried after birth, when the baby was first bathed, whether the baby was given anything else to eat or drink in the first hour after birth), and provider health checks for involution (whether the womb was becoming firm) in the first health exam post-delivery (see **Annex A, Table 3**).

The high 'Don't Know' responses on these indicators reflect study interviewer notes for indicators where they had to provide additional clarification to explain questions or terms to women. For example, HIV

was commonly confused with HPV, and interviewers frequently explained the difference between the two distinct sexually transmitted diseases. Interviewers also had to provide lengthy explanations about differences between receiving an epidural and other types of injections or IV medications. Interviewers commonly noted using body language to demonstrate the place where an epidural relative to other injections was administered. Other indicators where interviewers felt they had to provide additional clarification related to questions using the term 'immediate'. While allowing for an initial spontaneous response from the woman, the interviewer would then specify that the question referred to an 'immediate action' and sometimes asked the woman to estimate the number of minutes that had passed to enable her to respond more accurately to the question.

Women also had difficulty reporting some aspects of interventions related receiving uterotonics. This result is supported by women's high 'Don't Know' responses on indicators of uterotonic administration for PPH—whether oxytocin was received following delivery or whether anyone gave medication intravenously through a tube in the first few minutes after delivery (37% and 7%, respectively). The fact that some women accurately reported receiving medication by IV line following delivery, but did not report receiving oxytocin, despite being observed to receive oxytocin via IV line in this time period (**Box 5D**), could suggest that women recall having an IV line set because it was standard practice, but did not have specific knowledge of uterotonics received. Taken together, these results suggest that women can report some aspects of care routinely received, but may have difficulty with specific details (timing, name, route of administration) of medications received. This is likely because intervention procedures, in general, and specifically those regarding the use of medications were rarely explained to women, as noted by observers. Indicators regarding receiving uterotonics should be used with consideration of whether the purpose is to classify individual women who received the intervention, or to obtain an acceptable coverage estimate of the intervention at the population level. IF results suggest that one indicator— whether a uterotonic was received up to 3 minutes following delivery—may produce valid results only at the aggregate level.

In addition to data collector notes, quantitative findings also highlight the importance of indicator wording on women's reporting. 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. Quantitative findings and observer notes also suggest that women have difficulty reporting details related to the specific timing, routes, and names of medications received. 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 terms such as 'immediate' or 'within a few minutes') should be further explored to better refine indicators.

Indicators to be Further Researched

Because institutional delivery is a widely used proxy of whether women have access to comprehensive services such as emergency obstetric care and lifesaving commodities (e.g., uterotonics, magnesium sulfate, antibiotics), it is of interest 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 a high proportion of women (85%) who responded to this indicator correctly classified the type of facility as a public/ government hospital. Additionally, of women not able to classify the study facility as private or public, the majority (83%) correctly named the facility and 4% specified details including whether the facility was a hospital and part of the public or governmental sector. These results suggest that both aspects of question methodology, also used in DHS and MICS surveys, are important. These findings should be explored further in a study that includes multiple facility types (private and public sector).

An additional indicator, which we were unable to assess because of near-perfect correlation between observers' and women's responses, but which suggests the potential for valid reporting, was newborn birthweight. Specifically, women reported normal (2,500-4,499 grams) and low (<2,500 grams) newborn birthweight with near 100% accuracy. While the hospital practice of giving women bracelets listing the newborn birthweight, may have increased the salience of this indicator, analysis was restricted to women who recalled the weight from memory. Study interviewers also anecdotally corroborated the seeming importance of newborn birthweight, as women readily recalled the number. Occasionally women were able to recall only the amount in kilos and confirmed the grams after looking at their bracelet. Interviewers also noted that although not questioned on it, women spontaneously reported on the time of birth as another memorable event.

Another general trend observed was that some universal practices, particularly related to the client's initial assessment or the mother's post-partum health checks, were reported with high sensitivity by women. Although the high prevalence of these indicators limited robust analysis, it is noteworthy that in the few cases where women did not receive these preventative interventions, a high proportion of women falsely reported receiving them, suggesting low specificity. Because there were too few instances of women not receiving this type of care to robustly assess these indicators, this trend should be further explored. These findings may suggest that women in the Mexico setting may be positively biased towards reporting on aspects of preventative care received in the initial assessment and on health checks for the mother in the post-partum maternal period. It may also reflect the fact that, with

regard to low- prevalence scenarios, lack of high specificity in women's self-report may led to overestimation of the indicator.

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, Amoaful 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. Hospital General de México January-December 2013 Statistics Yearbook. Hospital General de México Anuario Estadístico Enero a Diciembre 2013. Accessed 10 July 2014.
15. Gutiérrez JP, Rivera-Dommarco J, Shamah-Levy T, et al. *Encuesta Nacional de Salud y Nutrición 2012.* Cuernavaca, México: Instituto Nacional de Salud Pública;2013.
16. Eng J. Receiver operating characteristic analysis: a primer. *Academic Radiology.* 2005;12:909-916.
17. Metz CE. Practical issues in ROC studies. *Radiology.* 1989;143:29-36.
18. Pepe MS. *Statistical evaluation of medical tests for classification and prediction.* Oxford: Oxford University Press; 2003.
19. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;201:621-625.
20. Vecchio TJ. Predictive value of a single diagnostic test in unselected populations. *New England Journal of Medicine.* 1966;274(21):1171-1173.
21. 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.
22. Qureshi Z, Lubano K. *Uterine timing of prophylactic uterotonics for the third stage of labour after vaginal birth: RHL commentary.* Geneva: World Health Organization;2011.

Annexes

Annex A.

TABLE 1. Full List of Indicators Assessed and Measured Coverage,[a,b] Mexico.

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	85.20	100.00	N
HIV status checked	54.13	69.77	Y
Offered HIV test	14.31	1.79	N
Received HIV test	16.32	1.17	N
Provider washes hands with soap and water or uses antiseptic before any initial examination	95.85	1.97	N
Takes blood pressure	98.80	97.77	N
Takes urine sample	24.30	25.17	Y
Checks fetal heart rate with fetoscope/ ultrasound	97.78	97.78	N
Wears high-level disinfected or sterile gloves for vaginal examination	100.00	99.65	N
Provider Respectful Care			
Encourages/assists woman to ambulate during labor	8.28	22.24	Y
Woman allowed to drink liquids/eat	2.88	0.51	N
Woman allowed to have a support person present during labor and delivery	0.35	0.00	N
Encourages/assists woman to assume different positions in labor	3.90	67.63	N
A support person is present at some point during labor	0.17	0.85	N
A support person is present at birth	0.17	1.02	N
First Stage of Labor			
Induces by uterotonic (IV line, IM injection, or tablet)[d]	12.88	11.63	Y
(Of women whose labor was induced) Uterotonic route for induction of labor by IV line[d]	27.69	100.00	N
Augments labor with uterotonic (by IV line, IM injection, or tablet)[d]	41.95	75.21	Y
(Of women whose labor was augmented) Augmentation of labor by IV line[d]	57.25	100.00	N
Receives injection during labor (to induce or augment labor)[d]	49.79	76.05	Y
Injection or IV medication received at some time during labor, before birth of baby (general)	63.57	76.01	Y
Membranes ruptured (to induce or augment labor)[d]	26.97	57.02	Y

Indicator	Self-Report Prevalence (Matched data)	True Prevalence (Matched data)	At least 5 counts/cell?
Skilled Birth Attendance- Main Provider			
Skilled main provider labor [c, d]	95.69	96.72	N
Main provider labor- doctor or medical resident [d]	81.90	96.21	N
Main provider labor- doctor (any)	76.72	1.72	N
Main provider labor- medical resident	5.17	94.48	N
Main provider labor- medical intern	3.62	3.28	N
Main provider labor- nurse	13.79	0.52	N
Main provider labor- nursing student /intern	0.69	0.00	N
Skilled main provider delivery [c, d]	97.77	92.81	N
Main provider delivery- doctor or medical resident†	95.66	92.36	N
Main provider delivery- doctor (any)	92.81	2.40	N
Main provider delivery- medical resident	2.91	90.07	N
Main provider delivery- medical student/ intern	2.23	7.19	N
Main provider delivery- nurse	2.05	0.34	N
Main provider delivery- nursing student /intern	0.00	0.00	N
Skilled main provider caught baby [c, d]	96.98	92.18	N
Main provider caught baby- doctor or medical resident [d]	89.70	91.12	Y
Main provider caught baby- doctor (any)	85.79	10.12	N
Main provider caught baby- medical resident	3.91	80.99	N
Main provider caught baby- medical student/ intern	2.31	7.82	N
Main provider caught baby- nurse	7.28	1.07	N
Main provider caught baby- nursing student /intern	0.17	0.00	N
Skilled Birth Attendant- Other Providers Present			
Skilled other provider labor [c, d]	92.05	96.95	N
Other provider labor- doctor or medical resident	70.39	84.60	Y
Other provider labor- doctor (any)	65.65	20.14	Y
Other provider labor- medical resident	11.76	77.65	Y
Other provider labor- medical student/intern	23.01	65.31	Y
Other provider labor- nurse	62.61	80.88	Y
Other provider labor- nursing student/intern	7.45	17.77	Y
Skilled other provider delivery [c, d]	97.14	97.98	N
Other provider delivery- doctor or medical resident	82.32	88.22	Y
Other provider delivery- doctor (any)	79.46	24.58	Y
Other provider delivery- medical resident	13.79	81.56	Y
Other provider delivery- medical student/intern	26.26	39.06	Y
Other provider delivery- nurse	81.14	81.99	Y
Other provider delivery- nursing student/intern	6.90	9.26	Y
Other provider delivery- pediatrician	39.40	6.84	Y
Other provider delivery- anesthesiologist	9.76	4.55	Y

Indicator	Self-Report Prevalence (Matched data)	True Prevalence (Matched data)	At least 5 counts/cell?
Number of Providers			
Four or more providers assisted with birth [d]	94.44	66.16	Y
More than six providers assisted with birth	37.54	13.64	Y
More than seven providers assisted with the birth	23.23	5.72	Y
Second & Third Stage of Labor			
Episiotomy performed	78.10	67.76	Y
Position of mother at birth- on back	93.30	99.66	N
Did health provider wear gloves during delivery of baby	99.82	99.29	N
Uterotonic administered within few minutes of delivery (IV line, IM injection, or tablets)	60.88	99.33	N
Uterotonic received 1-3 minutes after delivery	60.69	63.97	N
Uterotonic received after delivery of placenta	94.10	19.97	Y
Method of uterotonic administration- IV line	62.20	100.00	N
Oxytocin given following delivery	50.00	99.16	N
Applies controlled cord traction	93.83	96.91	N
Performs uterine massage after birth	82.37	83.07	Y
Palpates uterus 15 minutes after delivery of placenta	79.66	82.41	Y
3 AMTSL elements: prophylactic uterotonic, controlled cord traction, uterine massage post-placenta [d]	75.77	81.18	Y
Immediate Newborn Care (babies breathing at birth)			
Baby immediately dried with towel/cloth	99.27	93.58	N
Baby given to mother immediately after birth	59.83	10.34	Y
Baby placed immediately skin-to-skin on mother's abdomen	11.36	0.17	N
Baby immediately skin-to-skin on mother (2 item) [d,e]	1.05	0.17	N
Babies not on skin wrapped with towel	99.12	98.77	N
Breastfeeding within 1st hour of birth	64.81	34.74	Y
Something other than breastmilk given to baby within 1st hour of delivery	57.76	21.69	Y
Baby bathed within the first hour after birth [d]	0.68	2.72	N
Baby weighed	98.81	NA	N
Low birth-weight baby (<2,500 g) [d]	12.05	11.21	N
High birth-weight baby (>=4,500 g) [d]	0	0	N
3 elements of newborn care (newborn immediately dried + placed naked on skin + breastfed within 1st hr) [d]	9.66	0.00	N
3 elements of newborn care (newborn immediately dried + placed naked on skin (2 item) [e] + breastfed within 1st hr) [d]	0.74	0.00	N
Immediate Postnatal Care			
Provider did at least one post-delivery health check [d]	100.00	99.30	N
In first post-delivery exam, checks for bleeding	87.46	80.51	Y
In first post-delivery exam, examines perineum	90.10	75.09	Y

Indicator	Self-Report Prevalence (Matched data)	True Prevalence (Matched data)	At least 5 counts/cell?
In first post-delivery exam, takes temperature	93.34	66.89	Y
In first post-delivery exam, takes blood pressure	93.38	95.59	Y
In first post-delivery exam, checks for involution	83.45	81.65	Y
Maternal and Newborn Outcomes			
Cesarean section (C/S) performed	9.68	9.98	N
Decision for C/S taken after labor started	84.62	94.87	N
C/S performed after labor started	84.62	94.87	N
Provider decided C/S would be done	97.30	100.00	N
Reason for C/S- prolonged/ obstructed labor	21.62	29.73	Y
Complications (any)	38.15	13.61	Y
Eclampsia	21.38	2.36	Y
Hemorrhage	8.43	8.09	Y
Prolonged labor (>12 hours)	7.39	2.18	N
Prolonged/obstructed labor	10.76	3.03	N
None	61.85	86.39	Y
Blood products given	2.41	2.41	Y
Woman asked for pain relief medication while at facility	26.26	44.70	Y
Woman received pain relief medication	87.44	79.35	Y
Stillborn delivery [d]	0.34	0.34	N

[a] Text in blue notes indicators where there was not sufficient cell counts for robust analysis (n<5 per cell).
[b] Excludes '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 A.

TABLE 2. Near Universal and Rare Indicators, Unmatched Data

Indicator	True Prevalence
Near Universal practices (>95%)	
Provider takes woman's blood pressure	99.8
Fetal heart rate checked with fetoscope/Doppler ultrasound	97.8
Provider wore high-level disinfected or sterile gloves for vaginal examination	99.7
(Of those with induced labor), route for induction of labor by IV line	100.0
(Of those with augmented labor), route for augmentation of labor by IV line	100.0
Skilled main provider during labor (doctor, medical resident, or nurse)	96.7
Skilled additional provider during labor (doctor, medical resident, or nurse)	97.0
Skilled additional provider at delivery (doctor, medical resident or nurse)	97.1
Main provider during labor- medical resident	
Health provider wore gloves during delivery	99.3
Position of mother at birth – on back	99.7
Prophylactic uterotonic administered	99.3
Method of prophylactic uterotonic- IV line	100.0
Oxytocin given following delivery	96.2
Controlled cord traction	96.9
Newborn not placed skin-to-skin, wrapped in dry towel	98.8
Provider did at least one post-delivery health check	99.3
Of those who received cesarean section, provider made decision	100.0
Rare practices (<5%)	
Offered HIV test	1.8
Woman received HIV test	1.2
Provider washes hands with soap and water or uses antiseptic before each examination	2.0
Takes urine sample	1.4
Woman is allowed to drink liquids or eat during labor	0.5
Woman is allowed to have support person present during labor and delivery	0.0
A support person is present at some point during labor	0.9
A support person is present at birth	1.0
Main provider labor – nurse	0.5
Main provider labor – nursing student/ intern	0.0
Newborns placed immediately skin-to-skin on mother's abdomen	0.2
Newborns placed directly on mother's skin are covered with dry towel on mother's abdomen	0.0
Baby bathed within the first hour after birth	2.7

Annex A.

TABLE 3. Indicators with High 'Don't Know' Responses, Unmatched data

Indicator	% "Don't Know"
Woman's Self-Report (>5% DK)	
Did anyone give you a medication called 'oxytocin' to make your womb contract or become firm?	37.3
Did anyone give you medication or an injection called 'oxytocin' before you delivered the placenta?	35.8
While you were at the health facility for the birth of your baby, did anyone test you for HIV?	23.2
Did you or anyone else give anything to the baby to eat or drink within the first hour after delivery?	21.0
Did the health provider(s) wash their hands with soap and water or use antiseptic before examining you?	18.9
In the first few minutes after the delivery of your baby, did anyone give you medication intravenously through a tube in your arm?	6.9
Was your baby dried off with a towel or cloth immediately after birth, within a few minutes of delivery?	6.4
About how long after birth was your baby bathed for the first time?	5.9
In your first physical examination/check after delivery, did a health provider check your belly to see whether your womb was becoming firm?	5.7
Did the health providers wear rubber gloves to handle the placenta?	5.4
Observers' Self-Report (>2% DK)	
Was anything besides breast milk given to the baby to drink within the first hour after birth?	5.6
Was an HIV test done?	4.1
Provider performs artificial rupture of the membranes	4.1
Provider checks woman's HIV status (checks chart or asks woman)	3.3
Provider washed his/her hands with soap and water or used antiseptic before each examination of the woman	3.3
Provider ties or clamps cord when pulsations stop or by 2-3 minutes after birth (not immediately after birth).	3.0
Did the woman ask for any pain relief medication during labor, delivery, or immediately postpartum?	2.8
Provider takes urine sample to check for presence of proteins	2.6
Baby bathed within the first hour after birth.	2.5
What was the indication for the cesarean operation?	2.4
Provider encourages woman to have a support person present throughout labor and birth	2.1

Annex B.

TABLE 1. Validation Results for ALL Indicators With at Least 5 Counts per Cell, Matched Data*, Mexico.

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) *List Criteria*
Initial Client Assessment									
HIV status checked	569	54.13	69.77	54.91	47.67	54.13	0.5129	0.78	IF
Takes urine sample	572	24.30	25.17	50.00	84.35	24.30	0.6717	0.97	**Yes**
Provider Respectful Care									
Encourages/assists woman to ambulate during labor	580	8.28	22.24	7.75	91.57	8.28	0.4966	0.37	
First Stage of Labor									
Induces labor with uterotonic	559	12.88	11.63	27.69	89.07	12.88	0.5838	1.11	IF
Augments labor with uterotonic	472	41.95	75.21	49.58	81.20	41.95	0.6539	0.56	AUC
Receives injection for induction or augmentation of labor	476	49.79	76.05	59.12	79.82	49.79	0.6947	0.65	AUC
Injection or IV medication received at some time during labor, before birth of baby (general)	571	63.57	76.01	70.05	56.93	63.58	0.6349	0.84	**Yes**
Membranes ruptured (labor induction or augmentation)	356	26.97	57.02	38.42	88.24	26.96	0.6333	0.47	AUC
Skilled Birth Attendance									
Main provider									
Main provider labor- doctor or medical resident	580	81.90	96.21	82.26	27.27	81.90	0.5477	0.85	IF
Main provider delivery (who caught baby)– doctor or medical resident	563	89.70	91.12	90.06	14.00	89.70	0.5203	0.98	IF
Other providers present									
Other provider during labor was skilled [a]	591	92.05	96.95	92.15	11.11	92.05	0.5163	0.95	IF
Other provider during labor – doctor or medical resident	591	70.39	84.60	71.80	37.36	70.39	0.5458	0.83	IF
Other provider during labor – doctor (any)	591	65.65	20.14	72.27	36.02	65.65	0.5414	3.26	
Other provider during labor – medical resident	595	11.76	77.65	12.34	90.23	11.77	0.5128	0.15	
Other provider during labor – medical student /intern	591	23.01	65.31	25.65	81.95	23.01	0.5380	0.35	
Other provider during labor – nurse	591	62.61	80.88	64.02	43.36	62.61	0.5369	0.77	IF
Other provider during labor – nursing student/ intern	591	7.45	17.77	8.57	92.80	7.44	0.5068	0.42	
Other providers during delivery - doctor or medical resident	594	82.32	88.22	82.82	21.43	82.32	0.5213	0.93	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) *List Criteria*
Other providers during delivery - doctor (any)	594	79.46	24.58	81.51	21.21	79.46	0.5136	3.23	
Other providers during delivery - anesthesiologist	594	9.76	4.55	33.33	91.36	9.76	0.6235	2.15	AUC
Other providers during delivery - pediatrician	599	39.40	6.84	41.46	60.75	39.40	0.5111	5.76	
Other providers during delivery – medical resident	602	13.79	81.56	13.65	85.59	13.79	0.4962	0.17	
Other providers during delivery – medical student/ intern	594	26.26	39.06	31.03	76.80	26.26	0.5392	0.67	
Other providers during delivery – nurse	594	81.14	81.99	82.96	27.10	81.15	0.5503	0.99	IF
Other providers during delivery – nursing student/ intern	594	6.90	9.26	5.45	92.95	6.90	0.492	0.75	IF
Second and Third Stage Labor									
Episiotomy performed	580	78.10	67.76	97.71	63.10	78.10	0.8041	1.15	**Yes**
Uerotonic received 1-3 minutes after delivery	580	60.69	63.97	61.46	40.67	60.69	0.5106	0.95	IF
Uterotonic received following delivery of placenta	576	94.10	19.97	95.65	6.29	94.10	0.5097	4.71	
Palpates uterus 15 minutes after delivery of placenta	580	79.66	82.41	79.50	19.61	79.66	0.4955	0.97	IF
3 AMTSL elements: prophylactic uterotonic + controlled cord traction + uterine massage post-delivery of placenta	549	75.77	81.18	75.72	24.00	75.7	0.4986	0.93	IF
4+ providers assisted with the birth	573	82.37	83.07	82.14	16.49	82.37	0.4932	0.99	IF
7+ providers assisted with the birth	594	37.54	13.64	59.26	65.89	37.54	0.6257	2.75	AUC
8+ providers assisted with the birth	594	23.23	5.72	50.00	78.39	23.24	0.6420	4.06	AUC
Immediate Newborn Care									
Baby given to mother immediately after birth	580	59.83	10.34	63.33	40.58	59.82	0.5196	5.78	
Breastfeeding within 1st hour of birth	449	64.81	34.74	81.41	44.03	64.81	0.6272	1.87	AUC
Something other than breastmilk given to baby within 1st hour of birth	438	57.76	21.69	71.58	46.06	57.77	0.5882	2.66	
Immediate Postnatal Care									
First post-delivery exam, provider ask/checks for bleeding	590	87.46	80.51	87.37	12.17	87.46	0.4977	1.09	IF
First post-delivery exam, provider examines perineum	586	90.10	75.09	90.23	10.27	90.11	0.5025	1.20	IF
First post-delivery exam, provider takes temperature	586	93.34	66.89	94.39	8.76	93.35	0.5158	1.40	
First post-delivery exam, provider takes blood pressure	589	93.38	95.59	93.25	3.85	93.38	0.4855	0.98	IF
First post-delivery exam, provider checks for involution	556	83.45	81.65	84.58	21.57	83.45	0.5308	1.02	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) *List Criteria*
Maternal Outcomes									
Complication (yes to any)	595	38.15	13.61	60.49	65.37	38.15	0.6293	2.80	AUC
Eclampsia	594	21.38	2.36	57.14	79.48	21.38	0.6831	9.07	AUC
Hemorrhage	593	8.43	8.09	50.00	95.23	8.43	0.7261	1.04	**Yes**
None	595	61.85	86.39	65.56	61.73	61.84	0.6365	0.72	AUC
Blood products given	582	2.41	2.41	35.71	98.42	2.40	0.6706	1.00	**Yes**
Woman asked for pain relief medication at some time	575	26.26	44.70	32.68	78.93	26.26	0.5581	0.59	
Woman received pain relief medication	581	87.44	79.35	89.37	20.00	87.43	0.5469	1.10	IF

* Excluding 'Don't Know' responses.

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